Enhancement in Network Architectures for Future Wireless Systems

Aftab Ahmed

Ph.D Thesis

Electronic Engineering
University of York

July, 2018

Abstract

THIS thesis investigates innovative wireless deployment strategies for dense ultra-small cells networks. In particular, this thesis focuses on improving the resource utilisation, reliability and energy efficiency of future wireless networks by exploiting the existing flexibility in the network architecture. The wireless backhaul configurations and topology management schemes proposed in this thesis consider a dense urban area scenario with static outdoor users.

In the first part of this thesis, a novel mm-wave dual-hop backhaul network architecture is investigated for future cellular networks to achieve better resource utilization and user experience at the expense of path diversity available in dense deployment of base stations. The system-level performance is analysed and compared for the backhaul section using mm-wave band. Followed by the performance of the network model which is validated using a Markov Model.

The second part of the thesis illustrates a topology management strategy for the same dual-hop backhaul network architecture. The same path diversity is also utilized by the topology management technique to achieve high energy savings and improvement in performance. The results show that the proposed architecture facilitates the topology management process to turn-off some portion of the network in order to minimize the power consumption and can deliver Quality-of-Service guarantee.

Finally, the methodology to admit new users into the system, to best control the capacity resource, is investigated for radio resource management in a multi-hop, multi-tier heterogeneous network. A novel analytical Markov Model based on a two-dimensional state-transition-rate diagram is developed to describe system behaviour of a coexistence scenarios containing two different sets of users, which have full and limited access to the network resources. Different levels of restriction to access the network by specific groups of users are compared and conclusions are drawn.

Contents

List of Figures

List of Tables

Acknowledgements

First, I would like to express my sincere gratitude to my supervisor Prof. David Grace, without whom my work would not have been possible. He has helped me not only by providing lots of insightful ideas and valuable feedback, but also encouraging and supporting me when I had difficult times.

I would also like to thank Dr. Paul D. Mitchell, for his valuable comments and assistance on my research work.

My thanks also go to my nice colleagues in the Communication Technologies Research Group, who have provided me useful suggestions since I started my Ph.D studies.

Finally, my deepest gratitude goes to my family and wife, for their unconditional love and support.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Some of the research presented in this thesis has resulted in publications. These publications are listed as follows:

- **Conference Papers**

    i. Aftab Ahmed and David Grace, "A Dual-hop Backhaul Network Architecture for 5G Ultra-Small Cells Using Millimetre-Wave" published in *15th IEEE International Conference on Ubiquitous Wireless Broadband (2015, Canada)*.

    ii. Aftab Ahmed and David Grace, "Energy-Aware Topology Management for 5G Dual-hop Ultra-High Capacity Networks Exploiting Path Diversity" published in *8th IEEE International Conference on Ubiquitous and Future Networks (2016, Austria)*.

- **Journal Articles**

    i. Aftab Ahmed, David Grace and Paul Mitchell, "Grade of Service Enhancement in a mm-Wave Multi-hop, Multi-tier Heterogeneous 5G Network Architecture" under review in *IEEE Access Journal*.

Chapter 1. Introduction

Contents

1.1. Overview

WITH the escalation of next generation electronic gadgets, it is expected that 5G cellular networks will provide ubiquitous and seamless communication, supporting a diverse range of new applications. More and more people are seeking high speed Internet connectivity on the move, and it is becoming a requirement that similar data experience should be achieved both indoors and outdoors. However, currently the satisfaction with the outdoor connectivity experience is significantly lower than that of indoor [1]. This is even more prominent for applications such as Ultra High Definition (UHD) video streaming and cloud computing on smartphones and tablets. This switch to new data-intensive applications such as device-to-device (D2D) communications, UHD streaming media and cloud services is a crucial factor behind the development of new standards for the fifth generation (5G) of wireless systems [1, 2].

State-of-the-art technologies and drastic changes to the basic network architecture are needed to realize the perception of end user that they are connected to the network anywhere and anytime for anything. Although ultra-small cells are mainly deployed for indoor environments such as offices and shopping malls, operators have also started to investigate outdoor ultra-dense small cell deployments to accommodate the increase in

outdoor traffic [2, 3]. This access node densification needs to be supported by high capacity and low latency backhaul in order to guarantee the future promised user experience.

The desirable physical medium for 5G backhaul is fibre optic which can provide almost unlimited capacity and high reliability. However, installing adequate fibre in every small cell for future networks is not feasible and realistic due to the associated cost expenses as well as operational and maintenance challenges. Instead, wireless backhaul, especially using the mm-waves, is a promising approach to deliver the ubiquitous 5G network due to its higher deployment flexibility, limited signal propagation and cost effectiveness [2].

1.2. Motivation

With the proliferation of next generation smart phones and the success of social networking, forecasts indicate that the number of mobile subscribers worldwide accessing the Internet mainly for video streaming is increasing on annual basis from 25% to 50% and this trend is expected to continue to at least 2030. Also, it is predicted that mobile data traffic will escalate 11 times by 2020 with peak hour data traffic greater by 87% than the average [4, 6].

Some potential transmission technologies are emerging to support 1000x wireless traffic volume increase in future cellular communications. As discussed above, the mm-wave technologies for high capacity last mile is explored which have the opportunity to exploit large contiguous bandwidths in the frequency ranges 24 - 27 GHz, 71 - 76 GHz and 81 - 86 GHz that can be used for outdoor urban backhaul communications [7]. Massive Multiple-Input Multi-Output (MIMO) antenna technology has been shown to improve the spectrum efficiency by 10-20x using aggressive spatial multiplexing in the same frequency bandwidth [8, 9].

Considering wireless signal propagation characteristics, the mm-wave communication

technologies will reduce cell coverage in ultra-high capacity density networks [10, 11] as a result the signal interference between the neighbouring cells will be significantly reduced. Therefore, ultra-small cell networks with highly directional antennas are emerging in 5G networks.

In this case, the 5G network is not a simple upgrade of its predecessor, by adding additional spectrum and thus boosting the capacity, or replacing it with advanced radio technologies. It requires rethinking from the system and architecture levels down to the physical layer. In addition, we need to be able to answer the question of how to forward hundreds of gigabits backhaul traffic in ultra-dense cell networks with guaranteed QoS and affordable energy consumption by sustainable systems [3].

One way of delivering such high data requirements, is the ultra-small cell networks which tend to be self-organizing, low cost and low power small cell base stations (BSs). A series of techniques such as mm-wave communication, multihop relays and coordinated multipoint (CoMP) have been proposed to aid the ultra-small cell networks in achieving high capacity and spectrum efficiency [12]. These technologies present new tests for backhaul networks as additional traffic and coordination information is needed among the nodes. Therefore, innovative network architectures as well as networking mechanisms are needed to support the future service demands. A Macro-BS and an Ultra-Small Cells Network (USCN) can be used to construct a heterogeneous architecture as shown in Figure 1.1.

Additionally, research activity investigating mm-wave phased array antennas which exploit reflection has grown rapidly in the last year, but due to timing, the potential additional flexibility offered by such techniques have not been applied in this thesis. This offers a new direction for research in future network designs. Meanwhile, highly directional static beams with large angular separation between the nodes in densely deployed scenario implies that both signal interference and multipath should be negligible, and the backhaul links are largely unaffected by its environment [9, 13]. The street canyons with high buildings also add benefits, such as interference reduction and higher fre-
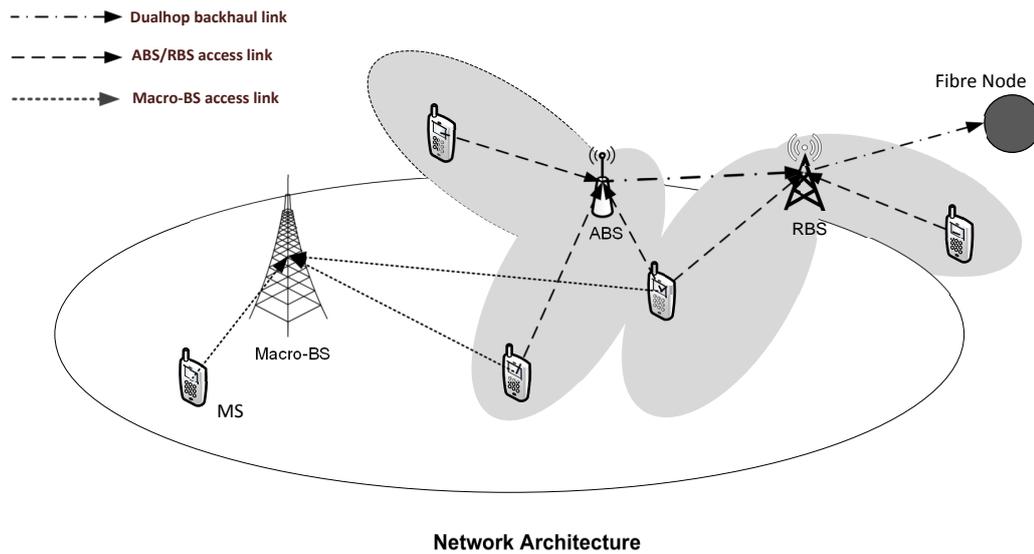
Figure 1.1: Multi-tier heterogeneous network for ultra-dense scenarios

quency reuse [9]. In theory, it is also possible to use Non line-of-Sight (NLOS) links by employing reflections from building walls and other objects. However, it is highly sensitive to node placement, antenna pattern and the local environment especially with the narrow beams [14]. Hence, Line-of-Sight (LOS) links are the most straightforward option to deliver the high speed backhaul connectivity. This ultra-dense small cell network architecture has a number of technical challenges. The major issues can be categorized as follows:

i. Resource Utilization

An ultra-small cell architecture brings significantly more infrastructure nodes into a wireless network. A major issue in this type of network is the backhaul architecture. Traditionally wired fibre or microwave links connect Macro/Micro BSs to the Core Network (CN). However, this approach will incur substantial deployment costs in a small cell network [15]. A wireless backhaul architecture is an effective solution for flexible deployment and cost reduction. Using this approach, a large number of small Base Stations (BSs) can be deployed in the locations that have capacity enhancement demands. Therefore, a mechanism is required to exploit the existing network diversity of multiple paths for effective resource

management and utilization.

ii. Energy Efficiency

Green communication is becoming vitally important in the future wireless networks. Analysis of energy consumption in typical cellular systems shows that the BSs consume most of the energy in a wireless network [16, 17]. It can be anticipated that the energy issue will be even more serious in ultra-small cell networks, because a large number of BSs are densely deployed. Therefore, a mechanism is required to intelligently control the number of activated BSs based on the dynamics of user traffic, as well as maintaining adequate QoS and capacity.

iii. Network Fairness

With coexistence of different types of nodes, a heterogeneous state with different levels of accessibility is expected to happen, where users with limited network choices and users with a more comprehensive set of connectivity options coexist in the same environment. This is primarily because different users often have dissimilar terrestrial locations, elevation/azimuth angles and/or antenna equipment choices [18, 19]. Meanwhile, when different types of users coexist in the same coverage area sharing a mutual resource pool, the resource utilization level of the two groups will affect each other. The presence of different types of users can worsen the performance of the disparate users and make the management of frequency spectrum challenging. Therefore, a mechanism is required for user admission control to provide an appropriate Grade-of-Service (GoS) for both types of users.

1.3. Thesis Hypothesis

The hypothesis guiding the research presented in this thesis is as follows:

> *"Introducing and exploiting an appropriate level of path diversity in wireless backhaul networks can enhance resource utilisation, energy efficiency and overall system fairness."*

In order to deliver this ultra-high capacity density networks with ultra-small size cells, it is essential to look into the detail layout about the backhaul network architectures. For this purpose, we need to analyse wireless networks having LOS mm-wave ultra-high capacity links which exploit the available path diversity. Path diversity denotes the availability of multiple alternative paths in the network to minimize the power consumption and can deliver Quality-of-Service (QoS) guarantee. While this work focusses on LOS path transmission aspects in the absence of interference, it would be generally compatible with more complex physical layers.

Meanwhile, network flexibility and path diversity have become a fundamental requirement for future wireless network and will have a significant impact on the design of new network architectures. The impact of increasing the flexibility in the backhaul network is assessed by evaluating the QoS performance of the proposed backhaul architectures in different large-scale small cell deployment scenarios.

1.4. Thesis Outline

The rest of the thesis is organised as follows:

Chapter 2 provides a literature review on the established work related to this thesis. The use cases and key requirements for 5G are discussed. Small cell networks and the associated challenges are then introduced as a key enabler to 5G. It then reviews the wireless backhaul technologies including sub 6-GHz and mm-wave. Subsequently, a

detailed state-of-the-art review in the active field of energy-aware topology and resource management is presented. Finally, the key aspects of heterogeneous network and the associated challenges are mentioned.

Chapter 3 introduces the wireless network architecture used in this work, including a dual-hop backhaul network and an ultra-small cell access network. Various simulation tools are discussed. The detailed modelling methodology and the system flow chart are presented, followed by parameters used for performance evaluation. Furthermore, an introduction to Markov analysis used for a theoretical proof later in this thesis is given.

Chapter 4 introduces an ultra-dense small cell scenario with static outdoor users. It proposes a mm-wave dual-hop backhaul architecture for future cellular generations, and exploits the available network diversity for the improvement in system performance. Finally, the same network model is analytically validated using simplified Markov Model.

Chapter 5 presents an energy efficient network topology management strategy which exploits the available network diversity in dual-hop backhaul architecture in order to lessen the power consumption by switching as many underutilized BSs as possible to a dormant mode. It is shown that topology management strategy is able to notably reduce the energy consumption. This scheme dynamically tunes the various states of the communicating nodes in a dual-hop layout to meet a trade-off between QoS and energy expenditure.

Chapter 6 discusses an analytical model based on a two-dimensional state-transition diagram is developed to help set the parameter values to control the issuance of resources in a multi-hop, multi-tier HetNet setup. A restriction mechanism is implemented to two different case studies in order to guarantee system fairness for each user group having different levels of network accessibility.

Chapter 7 presents the conclusions, summarises the orginal contributions and gives a number of recommendations for future work.

Chapter 2.   Literature Review

Contents

2.1. Introduction

THE responsibility of a backhaul network is to connect access networks to their core networks via a wired (e.g. fibre and copper) or wireless medium (e.g. microwave and mm-wave). Therefore, the design layout and optimization of wireless backhaul networks play a crucial role in future 5G cellular communication. The purpose of this chapter is to provide the background knowledge related to this thesis.

The rest of this chapter is organised as follows: the futuristic vision of wireless communication, its associated features are discussed first in Section 2.2. Next, the concept of mm-wave for 5G networks is introduced in Section 2.3. Followed by introduction to wireless ad-hoc networks in Section **??**. Latest research work carried out in the area of network deployment strategies and increase in energy efficiency are discussed in Section 2.4 and 2.5 respectively. Some key concepts related to network fairness are introduced in Section 2.6. Finally, conclusions are drawn in Section 2.7.

2.2. Wireless Communication Systems

Wireless systems and services have undergone a remarkable development, since the first cellular and cordless telephone systems were introduced in the early 1980s [20]. The cellular network which is a specific form of wireless network is divided into many overlapping geographical areas also termed as **Cells**. These cells are marginally overlapped at the edges to ensure that end users always remain within the coverage area. There are some other various forms of wireless networks like Bluetooth and Wi-Fi which support high speed communication but such techniques are only capable of providing short range links and with low mobility provision. The wireless network consists of two sections.

- **Access Network**

  The section of wireless network which provides the end user the actual connectivity to the network is called access network. It is also know as **Last Mile** which refers to the portion of the network that physically reaches the end-user's premises. The access network comprises of base stations and base station controllers. The mobile users get access to the network frequency resource through these BSs [21].

  For the access network, these BSs are equipped with antennas which have a wider beamwidth to provide network coverage and to establish a connection to the omnidirectional antenna fitted on the end user mobile [22, 23]. To cut down the installation

expenses, it is suggested that the small BS antennas can be mounted on street lamps, thus providing the necessary network infrastructure [24, 25].

- **<u>Backhaul Network</u>**

  The section of the wireless network which connects the access network to the backbone network is termed as backhaul network. For 5G networks, the backhaul needs to be more flexible to expose the potential of better, more effective and more flexible usage of immense spectrum. Further details about the backhaul network are given in section 2.3.

2.2.1. Degree of Deployment

Small cells that are needed to support a macrocell in a metropolitan area have to increase in number and complexity in order to catch up with the escalating 5G network throughput and coverage demand. To be assessed "**Covered**" in 5G, these small cells need to deliver a definite level of QoS.
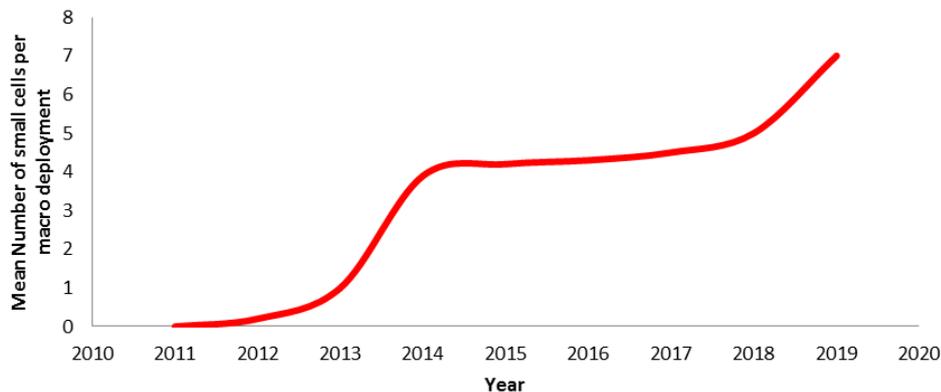
Figure 2.1: Evolution of small cells per macro in a metropolitan area [Redrawn from [1]]

One case study in [1] considers a metropolitan heterogeneous network scenario to define the number of small cells needed to complement the macro layer BSs in order to meet traffic demand growth from 2011-2019. This study considers a very realistic non-uniform BSs deployment as well as a non-uniform user data traffic distribution. The

study realizes that the recommended number of small cells would be 9 per macrocell by the end of 2019. These small cells are expected to be positioned outdoors at spacing between 50-200 metres (m) apart. As these small cells will provide network coverage to a very small geographical area, therefore it is termed as **Ultra-Small Cell Networks** [1].

It is anticipated that in the future mobile operators will combine various backhaul technologies to provide more efficient connectivity in these ultra-small cells. The network operators aim is to spend their capital on improving end-user experience, not on overhead-rigorous backhaul methods and techniques, and as a result are moving in a huge majority towards non-fibre based backhaul techniques [26]. Solutions which can transform their network capacity in a fraction of time and completely disarm the threat of imminent bandwidth crunch crisis are rigorously investigated. The operators are also very keen to use the same ecosystem in which there will be one unit for both the access and backhaul links, thus significantly simplifying the operation and maintenance and increase operational efficiency [27].

2.2.2. Backhaul Technologies

In a hierarchical telecommunications network, the backhaul portion of the network consists of the intermediate links between the core network or backbone network and the access network. At times, the term backhaul is used for the entire wired section of the network, whilst most often some networks have wireless (e.g. microwave and mm-wave) instead of wired backhaul (e.g. fibre and Copper).

Fibre has been considered as the ultimate backhaul approach that is able to deliver almost boundless capacity and high dependability. Installing adequate fibre to every ultra-small cell for future networks is not realizable due to the accompanying cost concerns as well as maintaining issues. Instead, new wireless backhaul approaches have become a major topic of research for mobile companies due to its higher deployment flexibility and cost effectiveness. Also, operators are considering making these small cells ultra-small to support the urban macrocell networks to expand their network coverage at not-spots and reduce the bottleneck congestion at hot-spots [13]. This extra layer of small cell BSs use the same macro cell network radio technology for higher throughput capacities in hotspot areas but also improve the coverage in certain remote cell edges.

Ultra-small cell deployments which utilise low transmit power BSs to provide localised coverage and capacity will be a key enabler to enhance the link capacity and coverage
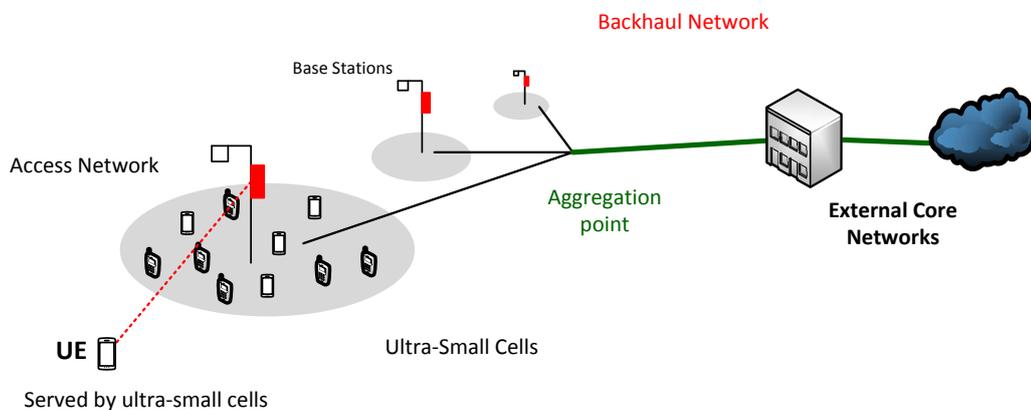
Figure 2.2: Abstract view of Ultra-Small Cell Network

of future networks especially in city centres and other high traffic areas [28]. Small cells usually have a coverage range of tens of metres to several hundred metres, and are referred to as femto-cells, pico-cells and micro-cells depending on the cell sizes and transmission power [29]. They can be deployed both indoors and outdoors. A mix of different cell sizes and radio technologies results in a **Heterogeneous Network** (HetNet) [30].

However, the cost per BS for an ultra-small cell will have to be significantly small than for macrocell BS but on the other hand the end user QoS cannot be compromised. All cell sites are independent whether it is macro, micro or pico require each of the underneath.

- User data traffic demand.

- Feasible location in terms of power availability and access for maintenance.

- Alternate wired backhaul infrastructure in case of emergency.

- Line-of-Sight propagation to backhaul nodes.

- **Wireless Topologies**

Using the orthodox LOS connectivity may limit the backhaul network coverage; therefore NLOS approach is being considered because there might be many BSs which cannot directly communicate with the macrocell BS via a single hop link because of the physical obstacles (buildings and towers) but can be reached via another hop which is known as a **Dual-hop Network** [31]. In such scenarios, more complex wireless topologies like mesh, stars, chains and trees could be used as shown in Figure 2.3.

A mesh topology whereby each node in the network is interconnected with one another can offer better robustness to traffic fluctuation and availability compared to the other two due to the path redundancy. However, these benefits come with associated costs, complexity of topology management, and scalability issues. Hence, there is no clear cut direction of which topology is superior as the deployments of the wireless
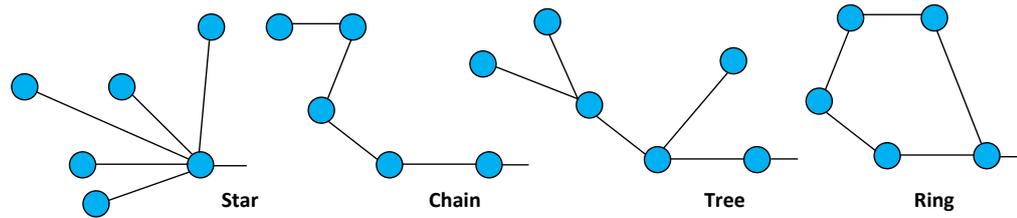
Figure 2.3: Backhaul types of topology configuration

backhaul are usually based on a combination of factors, such as geographical and business requirements.

Linking small cell BSs via chains or even trees may be a suitable approach when the BSs are deployed e.g. on lamp posts, bus shelters or any other place few meters above street level. In those situations it is adequate that only one of the small cell BSs is connected to the backhaul network and further connectivity is delivered among the small cell BSs themselves as shown in Figure 2.4. It is also possible that in some situations, the small cell BS is connected to two or more macrocell sites for topology flexibility known as **Dual Attachment** [32]. In such type of link attachment, if one link gets disconnected the second one is ready for link establishment but both of them are not used simultaneously.
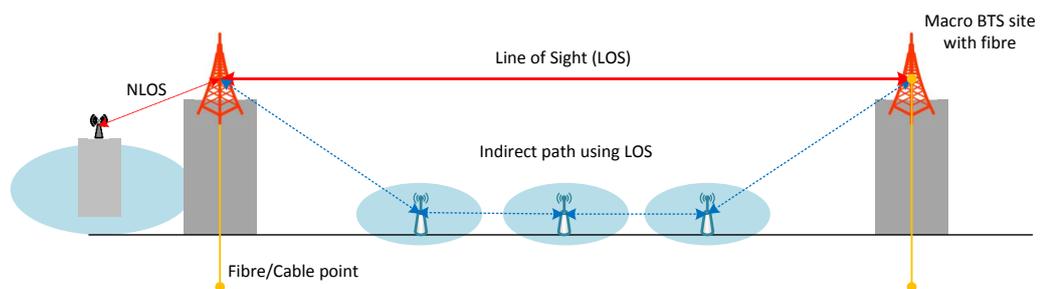
Figure 2.4: Potential wireless backhaul topology

- **Wire-line Topologies**

In case of connecting the small cell BSs to a wired backhaul, an underground cable (i.e. fibre or copper) is deployed for the connection between the sites. A layout diagram for wireline backhaul technologies is shown in Figure 2.5. The main benefit of wireline backhaul is the disposal of enormous frequency bandwidth as compared to

wireless backhaul. However, for the backhaul needs of ultra-small cells, fibre proves too expensive and impractical challenging to execute on a larger scale [9].
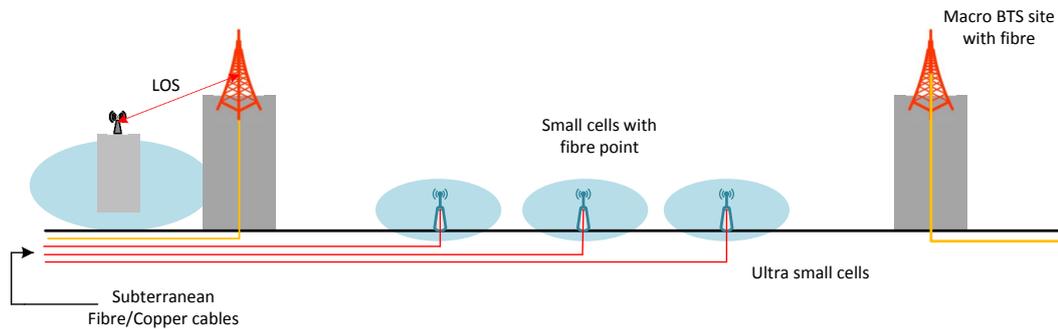
Figure 2.5: Potential wireline backhaul topology

2.2.3. Frequency Bands

There are comprehensive studies that describe the evolution of backhaul, including wired and wireless backhaul technologies such as [22, 33, 34]. With regard to the medium of backhauling, while fibre has been considered as an ideal backhaul solution that is able to provide almost unlimited capacity and high reliability, installing adequate fibre in every small cell for future networks is not feasible due to the associated cost issues as well as operational challenges. Instead, wireless backhaul technologies have become a major focus of attention due to their high deployment flexibility and cost efficiency. There are multiple candidate frequency bands for wireless backhaul, including sub-6 GHz band, and above 6 GHz band especially mm-wave bands. Besides, Free Space Optics (FSO) has also been considered as a promising technology for future wireless backhaul [34, 35]. Figure 2.6 clearly illustrates the amount of bandwidth available for each frequency band.

i. Sub-6 GHz Band

The existing generation of cellular systems around the world is based on sub-6 GHz band as it is very much suitable for NLOS Point-to-Point (P2P) and Point-to-multipoint (P2MP) situations as well as direct medium range LOS scenarios.

The largest contiguous spectrum allocated for International Mobile Telecommunications (IMT) is 200 MHz in the 3.5 GHz band [36]. This band is already significantly over utilized such that the available bandwidth is not enough to fulfil the needs of 5G.

In the FP7 BuNGee project [37], a dual-hop architecture is proposed in order to provide a capacity density of 1 $Gbps/km^2$. A series of emerging technologies have been applied to the backhaul network to improve the spectral efficiency and resource utilisation, including the use of advanced antenna array and in-band backhaul, whereby the access network spectrum can also be used by backhaul.

One of the key technologies to drastically increase the spectral efficiency is massive MIMO [9, 38, 39]. It proposes to use a much larger number of antennas at the BSs than its serving devices and aggressive spatial multiplexing techniques to achieve high capacity [38]. Although originally proposed for access networks, the authors of [39] applied massive MIMO technique to in-band wireless backhaul for small cells. The results demonstrated that higher link capacity can be expected with the massive MIMO approach.

ii. **Microwave Band**

Existing microwave backhaul techniques use highly directional antenna for their LOS connectivity. Latest research in high microwave regions indicates that these bands are also very capable for short range LOS/NLOS connections because of low path loss that can be easily compensated by using high gain antennas system [40]. Most often a LOS connection is not always available between the nodes, and this makes a necessity for near- and non-line-of-sight microwave backhaul. Results in [41] indicate that for NLOS scenarios microwave band beyond 20 GHz provides better tolerance to diffraction, reflection and penetration than sub-6 GHz. These bands also promise massive amount of unlicensed spectrum. It is convenient to think that the utilization of 28 GHz and 38 GHz bands will escalate in near future.
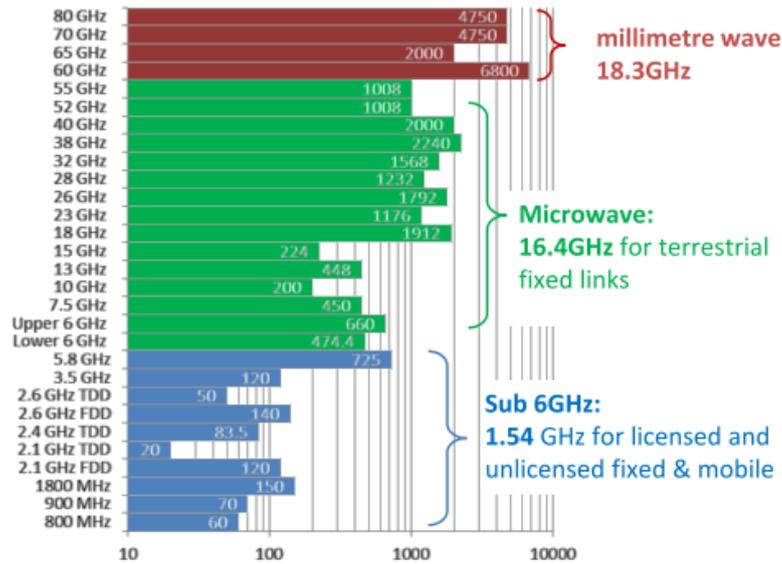
Figure 2.6: Spectrum allocations for terrestrial services in the UK [Taken from [1]]

iii. <u>Millimetre Wave Band</u>

Millimetre wave frequency band is the most promising band which is often referred to as the frequency range from 20 GHz to 300 GHz, the wavelength of which is between 10 mm to 1 mm [11]. It has abundant indoor and outdoor applications that include in-home applications like high-definition audio/video transmissions, desktop connections and portable devices. In addition, it is also very suitable for outdoor short range point-to-point communication links. That is why the mm-wave in bands 24 - 27 GHz is now being suggested as the 5G Pioneer band while 60 GHz and 70 - 80 GHz are also very exciting prospects for short range LOS wireless backhaul scenarios as the available bandwidth is so enormous that multi-gigabit throughput rates can be easily achieved [1].

There are also both **License Free** or **Light License** frequency bands available in mm-wave ranges which will not only heavily reduce the spectrum cost but also improve the spectrum reuse [7]. Also, the mm-wave at 60 GHz are highly vulnerable to oxygen absorption and rainfall [42] which obstructs the performance on long length links [13] which reduces the propagation range thus limiting the cell size as shown in Figure 2.7.

Due to high attenuation in free space and through walls, the mm-wave band is

a perfect choice for small range backhaul links [11]. This limited propagation in certain circumstances will restrict the signal from interfering with the nearby cells. The in-built safety and privacy provided by mm-wave is much superior due to lack of range and comparatively narrow size beam widths that can be achieved [42]. On a global scale, microwave and mm-wave are estimated to seize 61.5% wireless backhaul links by 2019 while mm-wave alone usage is expected to stretch from 3.2% in 2013 to 24% in 2019. Intensive research work on the propagation modelling of mm-wave through field measurements has been presented in [43]. Measurements showed that mm-wave in short distance LOS environments have almost identical path loss as free space [42].

Two types of mm-wave link failure are usually considered because of its propagation characteristics: weather based outages and beam misalignment. The current attempts to increase the reliability of mm-wave links include using a microwave link as a backup link, deploying multiple nodes to provide link diversity, and multi-hop relay backhauling [44]. Using hybrid links which combine microwave and mm-wave wave transceivers can improve reliability of the network at the expense of capacity when the mm-wave links suffer from outages caused by rain or obstacles. However, switching to microwave wave when a mm-wave link fails may not be able to support the capacity and QoS required in urban areas.

These existing methods also assume several candidate links/paths are already available and focus on the link/path selection algorithms. The feasibility of deploying the backhaul aggregation nodes is often neglected. Due to the narrow antenna beamwidth, mm-wave links are prone to misalignment outages caused by wind. The impact of wind sway on mm-wave beamforming misalignment has been investigated in [45, 46] for fixed beam antennas. It is concluded that beam tracking on the order of milliseconds is required to overcome outage. However, phased array antennas can exploit a more complex mm-wave channels with multipath, e.g. that caused by reflections. They show a promising way of overcoming these deficiencies [11, 47, 48, 49, 50].
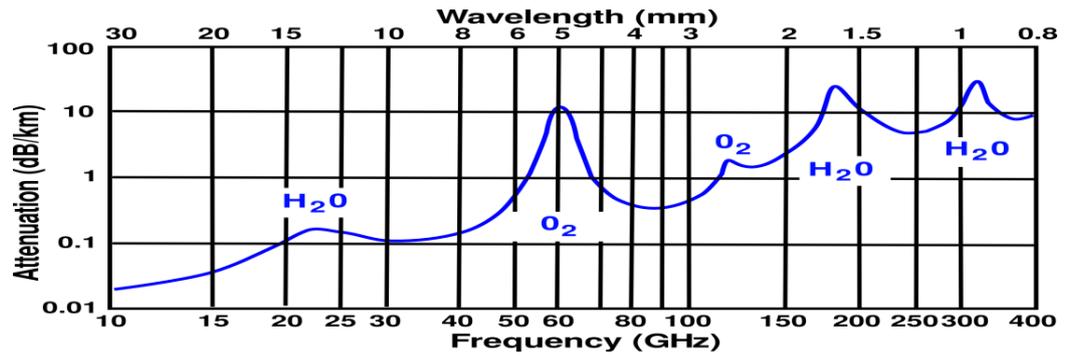
Figure 2.7: Atmospheric attenuation at different frequencies [Taken from [39]]

iv. Free Space Optics

Free Space Optics (FSO) uses laser light to communicate data through free space propagation [35]. FSO systems are able to provide point-to-point high bandwidth links without licensing requirements, and the installation and maintenance costs are relatively low. The transmission windows centred on the wavelengths of 850 nm and 1550 nm are suitable for FSO transmission because of their low attenuation, as well as the inexpensive transmitter and detector components. However, systems operating at a wavelength around the 1550 nm band have lower risk in terms of the eye safety which allows approximately 55 times more transmit power than those at 850 nm [35, 51].

The major disadvantages of the FSO are the vulnerability to dense weather conditions and beam dispersion [51, 52]. Although 99.999% availability is generally achievable for FSO links ranges less than 140 m as suggested in [52], it requires a considerable link margin. For the dense fog situation for example, at least a 31.5 dB link margin needs to be included in the link budget. This may not be a cost-effective solution as a higher transmit power is needed. Hybrid architectures of FSO and mm-wave have been proposed to increase availability because of their complementary propagation characteristics, but such benefits are only realised over long length links where rain and fog attenuation is significant.

In dense deployed small cell architectures, the rain attenuation at mm-wave is not a significant issue. Hence introducing FSO to the system may not able to bring

benefits in terms of availability and energy efficiency. However, in the case where a future mm-wave access network is deployed, where increased bandwidths and data rates are being proposed, a FSO backhaul would be one way of providing the high link capacities needed to support this increased capacity.

2.2.4. 5G Key Performance Indicators

Ultra High Definition streaming media and cloud computing have started to gain in popularity. In addition to high data rates, they have generic requirements for low latencies due to the conversational nature. While these applications have a broad range of requirements in terms of capacity, latency and information loss, the common challenge is to support a large amount of mobile devices/sensors over a wide area without affecting the performance of other services.

Based on the demands of the new applications and the ever growing number of mobile devices, industrial and research initiatives have identified a set of Key Performance Indicators (KPIs) [2, 12, 53]. Figure 2.8 illustrates a significant example on the expected enhancement of KPIs proposed for IMT-2020.

Some of the KPIs most relevant to this thesis are described below:

– **Area traffic capacity** is measured in $Mbps/km^2$ (or $Gbps/km^2$) and is the total amount of data traffic capacity of a wireless network in a given area.

– **Latency** in ms is the time it takes for a small data packet to be transmitted over the network from initial generation of data to its ultimate usable reception.

– **Mobility** in km/h is the maximum speed between a vehicle and a communicating node, at which the network is able to deliver the required QoS.

Some other performance evaluation metrics used in this thesis are thoroughly explained in Chapter 3.
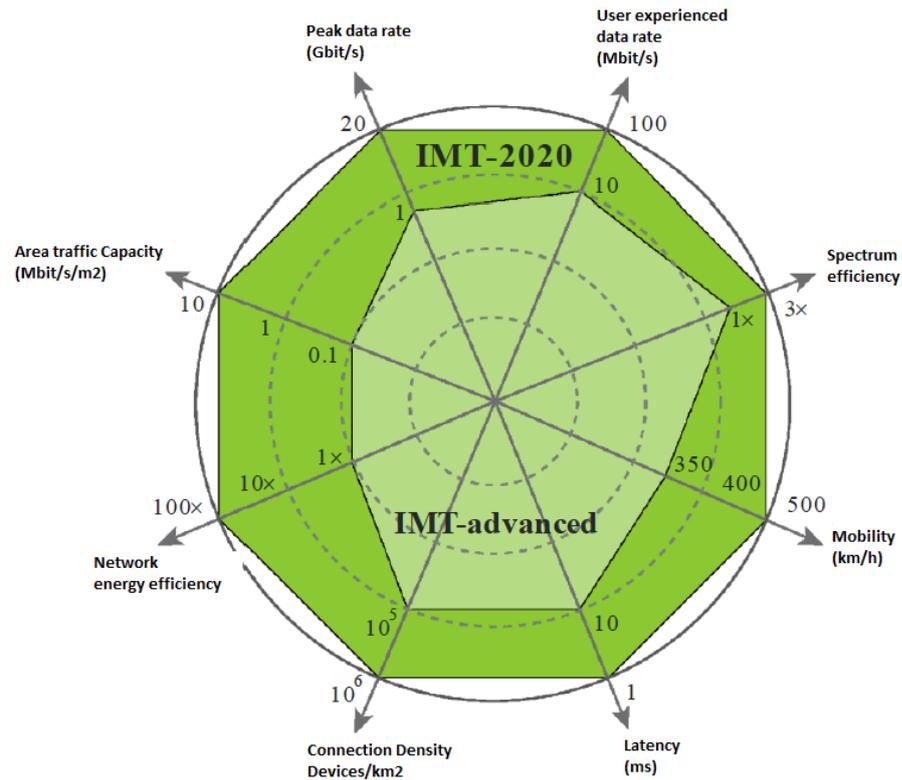
Figure 2.8: Enhancement of key parameters in 5G (directly reproduced from [2])

2.3. Wireless Ad-hoc Networks

Traditionally Wireless Ad-hoc Networks (WANETS) has a flexible topology but recent advances in fixed wireless networking where communicating nodes are turned ON and OFF while facilitating the mobile users [54, 32]. It results in a very dynamic network topology which changes with time. Thus, this mm-wave dual-hop network architecture explained earlier has increase in similarity with the WANET.

WANET is defined as a type of unmanaged multiple hop networks [32]. The topology layout is most often unpredictable and depends on the current conditions and requirements of network However, the presence of dynamic and adaptive routing protocols enables ad hoc networks to be formed quickly. The WANETs can be broadly classified into two types:

2.3.1. Fixed Ad-hoc Networks

Fixed Ad-hoc networks are also known as Wireless Mesh Networks (WMNETs) [54]. They rely on some pre-existing infrastructure. All centralized/decentralized nodes are interconnected to each other creating a mesh type structure. Many redundant paths are generated but by using sophisticated routing algorithms the intermediate nodes can choose appropriate paths to transfer their data packets to the correct destination. Different routing algorithms use network parameters like link quality, link congestion, relay burden and power consumption for selecting suitable routes over the entire network.

2.3.2. Mobile Ad-hoc Networks

Mobile Ad-hoc Networks (MANETS) are infrastructure-less networks of mobile devices with the capabilities of self-configuration. For MANETs the routing algorithms are quite complex and robust for tackling the high mobility issues. In MANETs, there is no centralized unit and all the operations are carried out in complete distributed manner. However, WANETS can be further categorized on the basis of applications.

2.4. Future Network Deployment Strategies

It is envisioned that the network architecture of 5G networks will be much more complex and self-Organizing Network capabilities (e.g., autonomous load balancing, interference minimization, spectrum allocation and power adaptation) for both the access and backhaul networks will be essential to the overall systems in order to reduce manual interventions and energy expenses. Also, a user will have simultaneous active connections to more than one BS or access point (AP) using the same or different Radio Access Technologies (RATs).

In [55, 56], the authors presented the results of a detailed survey on 5G cellular network architectures and some of the key emerging technologies that are helpful in improving the architecture and meeting the demands of users. In this detailed survey, the prime focus is on the 5G cellular network architecture, massive multiple input multiple output technology, and device-to-device communication. Also, a general probable 5G cellular network architecture is proposed, which shows that certain short range technologies, like mm-wave communication, ultra-small cell access nodes, D2D, network cloud and the visible light communication can be a part of 5G cellular network architecture.

Meanwhile, the emerging diverse traffic patterns in ultra-small cells, both spatial and temporal, together with the increasing demand in processing power, make the cloud technologies very attractive for enabling complex network management tasks. Backhaul networks need to implement new flexible management strategies as the rest of the system in order to improve the flexibility and resource utilisation of the network. In [9], the authors proposed the methodologies for wireless backhaul/fronthaul in ultra-dense wireless networks. They proposed millimetre wave backhaul deployment scenarios that serve outdoor users in an ultra-dense outdoor Radio Access Network (RAN) mounted on street lamps and traffic lights. The architecture is based on either Cloud-RAN (C-RAN) or Distributed-RAN (D-RAN) configuration. The analysis revealed that the unevenly distributed traffic fluctuations caused by different user types and mobility

requires that the location and number of backhaul links need to be carefully planned to overcome the barrier of bottlenecks on the backhaul links.

To realize ultra-dense network, a resilience, robust, gigahertz size bandwidth, small cells BSs are the mandatory requisite. Also, the mm-wave can easily be integrated with the massive MIMO for enhancement in link peformance. In [57], the authors addressed the feasibility of mm-wave massive MIMO based wireless backhaul solution for 5G. They also proposed a digital based phase shift network based hybrid scheme for mm-wave massive MIMO. Properties of mm-Wave massive MIMO are used to reduce the required cost and complexity of a transceiver. One key feature of the proposed scheme is that the macrocell BS can simultaneously support multiple small-cell BSs with multiple streams for each small-cell BS.

In [33], the authors focused on the throughput and energy efficiency for 5G wireless backhaul networks. They proposed two ultra-small cell scenarios using mm-wave communications to achieve gigabit transmission rates. In the centralized scheme, the small cells BSs are uniformly distributed in a star topology. These BSs transmit their backhaul traffic to the macrocell BS (MBS) by mm-wave short distance links, and then the aggregated backhaul traffic at the MBS is forwarded to the core network by fibre to the cell (FTTC) links. Meanwhile, in the distribution solution, there is no MBS to collect all backhaul traffic from small BSs, and all data traffic is relayed to a specified small BS using mm-wave links. These small BSs are also connected with the operator core network by fibre to the cell. From numerical results, it is indicated that the distributed solution is more flexible, robust and energy efficient.

In [6], the authors proposed a cognitive based backhaul deployment strategy for future ultra-dense wireless networks. The concept of cognitive Networks (CNs) is proposed in order to deal with the increasing network complexity. The network can learn from the changes taking place and use them to make future decisions, all while taking into account end-to-end goals. The authors also aimed to reduce the energy consumption in low traffic scenarios using energy efficient cognitive topology management strategy.

They proposed a reinforcement learning based backhaul link selection procedure which is aimed to concentrate distributed traffic on fewer backhaul beams within the network and puts the idle beams into sleep mode to improve energy efficiency. Significant reduction in energy consumption is achieved with marginal compromise in Quality-of-Service.

2.5. Reduction in Energy Consumption

The densification of the network driven by increasing traffic volumes can lead to increased CAPEX and OPEX. Therefore, energy efficiency has become an important research topic. New designs of 5G need to reduce the energy consumption by considering all layers, from physical components up to network level.

With the help of the proposed network architecture as shown in Section 2.5, the overall system throughput can be significantly increased when system traffic is at a high level. However, it is not an effective strategy in terms of energy efficiency to keep all nodes running when considering the behaviour of the users over a whole day. Figure 2.9 shows how the number of users varies throughout a typical day for accessing the Internet. The highest peak occurs around 22:20 and the lowest point is around 05:30 in the morning [58]. Therefore, we should devise some topology management techniques to switch off parts of the network in order to achieve higher energy savings. There are several papers that provide an overview of current energy saving techniques in a wide range of communication networks.

The potential mechanisms for minimizing the BS energy consumption are discussed in [59]. The authors explore different set of approaches for improving energy efficiency and reducing the burden on network operators. Different key aspects are considered to save energy in future wireless networks: Sleep mode techniques which enable switching off power consuming components in the BS, femto cell or relay deployments on cell edges and multiple antenna wireless systems.
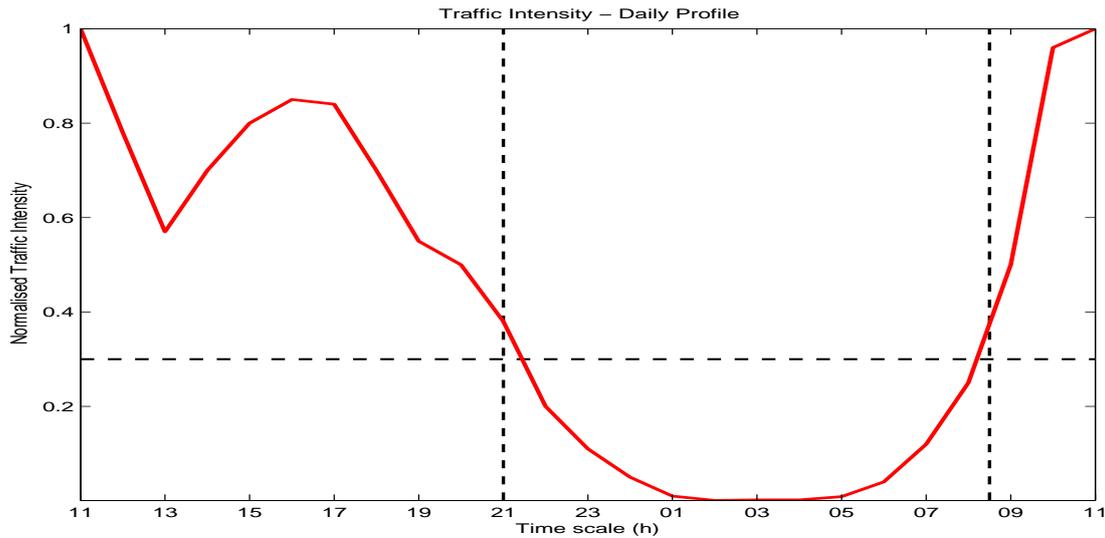
Figure 2.9: Daily Traffic Profile [Redrawn from [3]]

In [60], the authors discussed the dynamic BS planning (switching off redundant BSs) to reduce the energy consumption in cellular access networks. Since the BSs are designed to support peak time traffic, the utilization of BSs can be very inefficient during peak-time because the traffic profile is time varying. It is observed that the mean and variance of data traffic profile and the BS density are the dominant factors that govern the amount of energy saving that can be attained. They compared ideal and real traffic profiles to quantify the potential savings from dynamic BS switching in a realistic scenario.

In [17], the author discussed energy efficient resource allocation scheme which is based on Signal to Interference plus Noise Ratio (SINR) for the future mobile networks. It is observed that the lower SINR choices result in higher interference in the network and vice versa. Different SINR based choice restriction policies are evaluated to achieve energy savings.

In [61], the authors proposed an energy-aware topology management scheme which aims to reduce energy consumption by switching off as many underutilized BSs as possible in future cellular networks. The architecture under consideration is the BuNGee architecture which is selected for the temporary event city scenarios such as festivals, parades and city marathons. The main objective is to cope with the spatial traffic load fluctuations during temporary events. The scheme successfully tunes the ON/OFF state

of the BSs depending on the network traffic demands.

In [24], the authors presented a distributed topology management strategies for two-hop cellular network. The objective is to substantially reduce the energy consumption at low occupancy levels. The network uses few complex and expensive HBSs, combined with a large number of simple and cheap ABSs to provide the required throughput density. The proposed energy model is based on traffic loads and the behaviour of adjacent nodes. An ABS can be turned ON/OFF by the local information it receives rather than being controlled by its corresponding HBS. A balance between energy reduction and required Quality-of-Service is considered for the analysis. It is shown that the topology management schemes can achieve energy savings up to 35%-70% at low traffic loads compared with a deployment without the topology management scheme.

In [62], the authors focused on the impact of backhaul on the energy consumption of wireless access networks while considering different data traffic requirements. Different layouts with different technologies (i.e. copper, fibre and microwave) are analysed. It is suggested that the backhaul power expenditure has to be included in the energy efficiency analysis in order to attain truly green wireless access network architecture. In particular, it is shown that a hybrid backhaul solution combining a fibre to the Building (FTTB) option (i.e., to backhaul femto BSs) and microwave links (i.e., to backhaul macro BSs) is a promising candidate in setups where the wireless network is characterized by a high degree of femto BSs.

In [63], the authors presented channel assignment schemes to increase energy savings when used with green topology management schemes. These schemes activated and deactivated the BSs in accordance to local traffic demands to find a trade-off between energy savings and system throughput performance. These savings increase further by an additional average of 15% all occupancy levels.

- **Multi-hop Relay Links**

  From the Friis Free Space Equation, the power consumption of a BS increases with increase in distance between the transmitter and receiver [32, 64]. Thus, it is possible to use two or more short distance links (multi-hop relay) instead of one long distance hop to reduce the overall energy expense.

  Figure 2.10 gives a simple example of multi-hop network, where source node S would like to transmit data to the destination node D2. It has two options, one is from S-D2 and another is from S-D1-D2. Assuming that P1, P2 and P3 are the power consumed for link S-D1, D1-D2 and S-D2 separately.

  From Figure 2.10, the total energy consumption can be reduced when:

  $$P_1 + P_2 < P_3 \qquad (2.1)$$

- **Sleep Mode Algorithms**

  Sleep mechanisms are the key aspect of the topology management strategy. It is considered as one of the most common schedule management techniques. This refers to a low power mode for BSs in wireless networks because the BSs consume most of the energy in communication [65]. These modes generate significant energy reduction when the traffic load is at low level. This technique is always combined with other energy effective approaches to create a fusion solution.
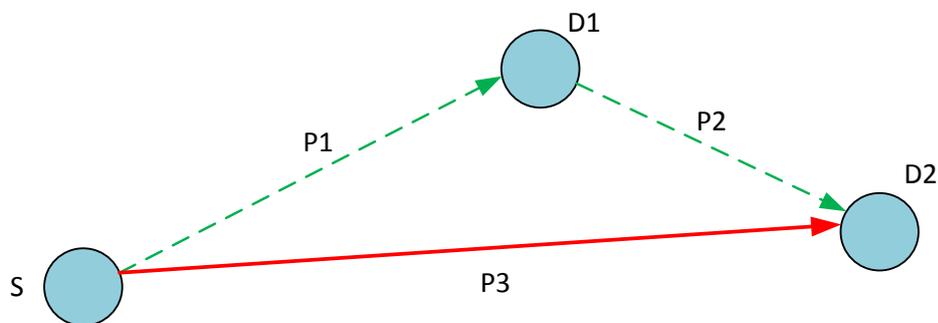
Figure 2.10: Simplified multi-hop relay network

In [66], a system selection algorithm is developed for cooperative 2G/3G networks in order to minimize the system power consumption whilst ensuring the required QoS is met. In addition, a network scale sleep mode is proposed in the system where the algorithm can achieve a large energy reduction.

In the work of [67], the authors aim to reduce the BS power consumption in the wireless cellular network using a BS sleep mode to serve consumers in periods of low traffic. The article focuses on the design of BS sleep and wake up transients. A practical case study is used to evaluate the performance of this technique. A similar work is carried out in [65]. The idea of green BS sleep mode design is proposed in both the time and spatial domains according to the LTE standards.

A distributed cell breathing technique is used in [68] for energy efficient cellular networks. When a BS attempts to go into its sleep mode, it tries to reallocate the traffic to its adjacent cells if they are not overloaded. The authors developed green cell breathing algorithms to avoid a centralized coordination and initialize the load thresholds.

The authors in [69] presented a green topology management scheme to facilitate future wireless network to adjust to traffic fluctuations in order to achieve energy savings. The objective is to first identify and then switch off the under performing BSs while taking into account the system performance. It is shown that 80% of energy consumption can be saved at low traffic loads.

An in-depth discussion of Sleep/Wake Up mechanisms is presented in [70]. The authors develop Sleep/Wake Up schemes for the BSs of a network comprising femto cells deployed within macro cells for the purpose of offloading part of its traffic. They use Markov Decision Processes (MDPs) to optimise the schemes based on information about traffic loads and user location in the cell.

In the aspect of BS radio efficiency, Power Amplifiers (PAs) account for significant portion of energy consumption of a BS in wireless networks. The PAs are very important for some mobile applications which battery lives are limited. A report from

[71] shows the radio frequency (RF) PAs consume around one third of the total input power consumption of a GSM BS and the efficiency of PA (output power versus input power) is less than 10%. Thus, there are two potential approaches to reduce energy consumption in this aspect: Improving the efficiency of PAs and switching off certain PAs at suitable time period.

In [72], the authors present a tuneable matching network that enhances both efficiency and linearity of power amplifiers. They demonstrated their works under 3GPP WCDMA modulated input. The results show that the PA with a dynamically controlled tuneable matching network can achieve up to 5% improvement compared with a PA employing a fixed matching network.

2.6. Fairness in Heterogeneous Networks

Mobile network operators are looking to facilitate and encourage subscribers to offload their data traffic from macro BSs to the alternative single/multi-hop small-cell networks and vice versa, essentially forming a basic heterogeneous network [30, 56]. In general, a HetNet consists of multiple tiers of networks of different cell sizes/footprints and/or of multiple radio access technologies. A macro BS overlaying a multi-hop, ultra-small cell network is a good example of a multi-tier heterogeneous network as illustrated in Figure 2.11.

In such layouts, signal degradation and user equipment restrictions will reduce the availability of some portion of the network to some of the users. In such situations, those limited users will face an inadequate network environment. A composite state with different levels of accessibility is expected to happen, where users with limited network choices and users with a more comprehensive set of connectivity options coexist in the same environment. The problem of limited availability has an effect on any cellular wireless system, be it mobile or fixed access, and should be taken into account in the cellular architecture design.
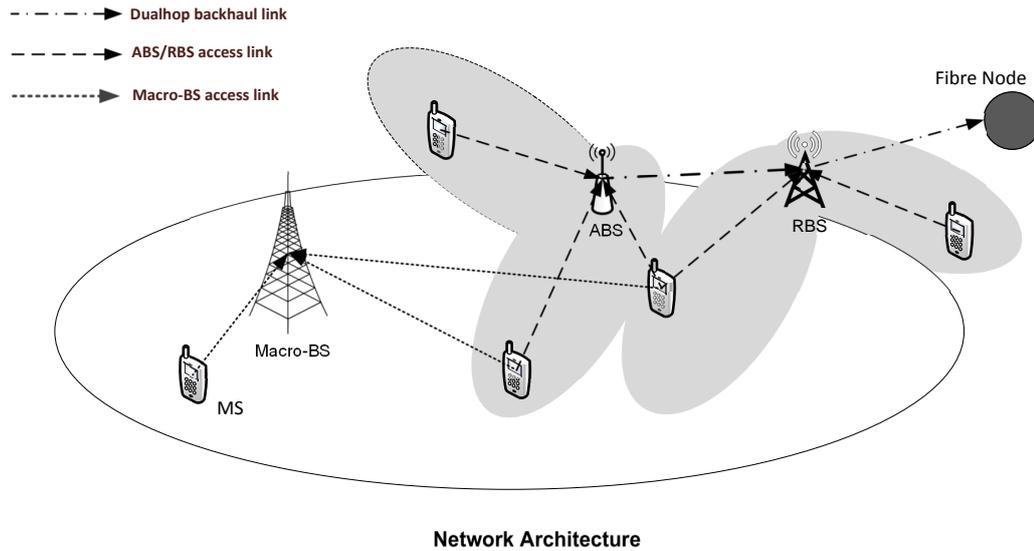
Figure 2.11: Multi-hop multi-tier high capacity heterogeneous network

In this work [18], the authors investigated the ways of enhancing the GoS in a coexistence scenario with different user types in a multiple High Altitude Platforms (HAPs) system with shared coverage area and radio spectrum. It is attained through the exploitation of HAP diversity. An analytical model based on a two-dimensional state-transition-rate diagram is developed to describe system behaviour of a coexistence scenario containing two user groups, which have full and limited HAP availability. On the basis of the analytical model, a restriction mechanism is implemented in order to achieve a fair balance of GoS for the two sets of users.

The algorithm restricts access to the channel resource for users with full HAP choice in order to give more chance of access to users with a more limited HAP selection. The analytical technique and controlling mechanism suggested in this paper are not only suitable for a multiple HAP system but are also potentially applicable to general communication systems. In Chapter 6, this restriction approach is applied to a coexistence scenario where users having greater freedom of choice can be directed on to other parts of the network, inaccessible by other users.

In [73], the authors presented an analytical and simulation models that obtain the minimum number of required communication channels for the end users and their alloca-

tions so as to increase the level of GOS per used of fixed wireless access communication systems. The algorithm is based on three different search methodologies by the end users, whose objective is to optimize the number of access nodes and their spatial allocation in the network. This approach also facilitates us to increase the level of fairness in a static wireless access network.

In [74], the authors proposed a new resource allocation scheme to ensure fairness in data rate offered by exploiting cell overlap regions. As the size of the cells on the ground increases, the cells start to overlap each other. The users in these overlap areas can therefore be assigned a channel from any of the overlapping cells. This technique works by blocking a proportion of users in the overlapping regions, even though there are free channels available. The channels saved from the overlapping region can then be used in non-overlapping sector. The scheme ensured uniform blocking levels within the coverage area as well as equal data rates across the user connections.

2.7. Conclusions

This chapter has provided the background information related to this thesis. The use cases and key requirements for 5G network deployment have been discussed. Ultra-Small cell networks were then introduced as a key enabler to 5G. The challenges in deploying ultra-small cell network including interference management, topology management for energy reduction and user association have also been discussed.

Wireless backhaul technologies using different frequency bands and topologies have been summarised. The proposed mm-wave dual-hop backhaul network architecture has been mentioned as a solution. Furthermore, the energy efficiency of cellular networks has been discussed. Enabling technologies for future wireless network architectures to improve the QoS and to reduce the energy consumption have been reviewed. Different user admission technologies for increasing system fairness have also been discussed.

Chapter 3.   System Modelling and Verification Methodologies

Contents

3.1. Introduction

I N this chapter, we present the system modelling techniques and the measurement methods used throughout this thesis. The system modelling techniques introduced in this chapter include simulations and analytical models. System modelling by professional simulation software is critically important in daily research. It is common practice to experiment with new ideas and thoughts on a system model before we make changes for real with the aid of professional softwares and powerful computers [24, 75]. Analytical models can be used to predict system behaviour. Especially, it is

always used at the design stage and is much affordable among the two system modelling techniques (simulation and analytical model).

The rest of this chapter is organised as follows: in Section 3.2, the system modelling techniques are described. An introduction to Markov Process is explained in Section 3.3. The system flow chart is illustrated in Section 3.4. In Section 3.5, the performance evaluation metrics are presented. Different system models necessary for system simulation are given in Section 3.6 to 3.8. Finally, conclusions are given in Section 3.9.

3.2. System Modelling Techniques

As computers have become indispensable for creative work in science in engineering, numerous simulation tools are developed to model wireless communication systems, such as C/C++, OPNET and MATLAB. Each of them offers rewards and shortcomings. The simulation environment that is used to build up this network model is MATLAB. It is mainly based on matrices, differential equations, arrays of data, plots and graphs [75, 76]. It also offers powerful built-in mathematical functions and toolboxes for a wide range of scientific areas such as Digital Signal Processing, Digital Image Processing, Linear Algebra, and Control System Design, which can efficiently diminish the time to develop programming codes [76].

Matlab provides effective ways to produce graphical results for performance evaluation. Furthermore, this work will consider a number of dynamic network behaviours, such as traffic, channel usage, network topology etc. Matlab can significantly reduce the time for code development. In recent years, Matlab is commonly used in both academe and industry. A system level simulator in Matlab provides transportable codes for some other researchers. Thus, Matlab is still considered as the most appreciate professional simulation tool in this work.

- **<u>Monte Carlo Simulation</u>**

A definition of Monte Carlo simulation was given in [77, 78] as representing the solution of a problem as a parameter of a hypothetical population, and using a random sequence of numbers to construct a sample of the population, from which statistical estimates of the parameter can be obtained. That means the more the trials are repeated, the more accurate simulation results we obtain. An event based strategy is applied here, which means the simulation is only carried out when a specific discrete event occurs, instead of the time-continuous simulation. The total simulation time can be significantly reduced because the programme only works when a new event happens.

The event here is defined as a thing that happens in the simulation, such as a user arriving, departing, blocking, dropping, and retransmission. The general process of the Monte Carlo simulation is illustrated in Figure 3.4. The simulator initializes parameters firstly, such as position information, traffic model, and topology management scheme. Next, the simulator passes through every event and takes measurement when the system is relatively stable. Further discrepancies in these measurements are eliminated by neglecting the start and end effects. If all events have been completed, a final simulation result will be obtained. The Monte Carlo simulation will be widely used in next few chapters, and its results will compare to the results from analytical models.

3.3. Verification Methodologies

Verification is used to analyse two or more independent methods by comparing their results for the purpose of system evaluation. Theoretical results are produced to validate the designed strategies. In this thesis, queuing theory is used to analyse and compare the behaviour of multi-hop, multi-tier HetNet scenario, to validate the resource and system fairness strategies.

3.3.1. Analytical Model - Markov Process

A definition of a Markov Process is given in [79, 80]:

> ***"A random process in which the probabilities of states in a series depend only on the properties of the immediately preceding state or the next proceeding state, independent of the path by which the preceding state was reached. It is distinguished from a Markov Chain in that the states of a Markov Process may be continuous as well as discrete."***

A Markov process helps us to generate a new sequence of random but related events mathematically, which looks similar to the original. It is useful for analysing dependent random events. In other words, the current event depends on what happened before (the previous event). For example, yesterday's weather has an influence on today's weather.

Assume that we have a set of states, $S = \{S_1, S_2, \cdots, S_n\}$, where n is the total number of states in a Markov chain. The process starts in one of these sates and moves successively from one state to another, or stays in its current state. Here, we define such a move as a transition. If the chain is in state S_i currently, and then it moves to state S_j with a probability denoted by P_{ij}, this probability does not depend on which states the chain was in before the current state [81]. The probability P_{ij} is called the transition probability from state S_i to state S_j. If the process remains in the state is was in, the probability is shown as P_{ii}.

If we use a matrix T to present all the probabilities transferring from one state to another, the matrix T can be shown as:

$$
T = \begin{Bmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1} & P_{n2} & \cdots & P_{nn} \end{Bmatrix}
$$

The matrix T is known as the transition matrix in a Markov chain.

Markov process is a useful mathematical method, and is widely used in different areas, such as statistics, queuing theory, information sciences, economics and etc. In wireless networks, researchers usually use a Markov process to generate analytical models or predict the system behaviour [82].

In this work, a two-dimensional state-transition diagram is developed to help set the parameter values to control the issuance of resources in coexistence scenarios as shown in Figure **??**.
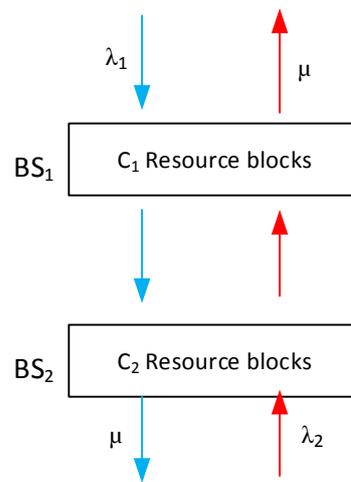
Figure 3.1: Two-dimensional Queuing Model

An example of a two dimensional queuing system is illustrated in Figure 3.1. The system has two base stations BS_1 and BS_2 with channel set c_1 and c_2 assigned respectively. The total number of states is equal to $(C_1 + 1)(C_2 + 1)$. By defining arrival rate of users in these BSs as λ_1, λ_2 and the departure rate as μ, the probability of the system having j_1 and j_2 channels occupied in each BS is given by the following equilibrium equations [82]:

$$(\lambda_1 + \lambda_2 + (j_1 + j_2)\mu) \cdot P(j_1, j_2) = (j_1 + 1)\mu \cdot P(j_1 + 1, j_2) + (j_2 + 1)\mu \cdot P(j_1, j_2 + 1)$$

$$+ \lambda_1 \cdot P(j_1 - 1, j_2) + \lambda_2 \cdot P(j_1, j_2 - 1)$$

$$(j_1 < C_1, \ j_2 < C_2) \qquad (3.1)$$

$$(\lambda_2 + (C_1 + j_2)\mu) \cdot P(C_1, j_2) = (j_2 + 1)\mu \cdot P(C_1, j_2 + 1) + \lambda_1 \cdot P(C_1 - 1, j_2)$$

$$+ \lambda_2 \cdot P(C_1, j_2 - 1)$$

$$(j_1 = C_1, \ j_2 < C_2) \qquad (3.2)$$

$$(\lambda_1 + (j_1 + C_2)\mu) \cdot P(j_1, C_2) = (j_1 + 1)\mu \cdot P(j_1 + 1, C_2) + \lambda_1 \cdot P(j_1 - 1, C_2)$$

$$+ \lambda_2 \cdot P(j_1, C_2 - 1)$$

$$(j_1 < C_1, \ j_2 = C_2) \qquad (3.3)$$

$$(C_1 + C_2)\mu \cdot P(C_1, C_2) = \lambda_1 \cdot P(C_1 - 1, C_2) + \lambda_2 \cdot P(C_1, C_2 - 1)$$

$$(j_1 = C_1, \ j_2 = C_2) \qquad (3.4)$$

The normalization equation is

$$\sum_{j_1=0}^{C_1} \sum_{j_2=0}^{C_2} P(j_1, j_2) = 1 \qquad (3.5)$$

It is assumed that the input flow into a state is equal to output flow from a state when the system is in equilibrium [82, 83]. The coefficient matrix which is generated using Equations 3.1-3.4 is diagonal dominant and positive, and all the entries off the main diagonal are negative or zero. Also, all the nonzero elements cluster along the main diagonal. Further details are given in Section 4.6.

The Markov Model and equilibrium equations vary for different systems, which will be discussed in detail for a more complex system in Chapter 6.

3.4. System Flow Chart

The simulation flow chart for this research work is categorized into the following four parts:

i. **Structure Layout**

The structure model is setup which consists of the most common parameters. They are usually constant for the whole simulation time. These parameters can be changed for different types of scenarios. The essential physical layer models like antenna orientations and gains can be built up and generated. As shown in Figure 2.9, the base stations and relay nodes are placed on pre-defined positions along the streets to form a tree like structure. The sub-function for the structure block is given in Figure 3.2.
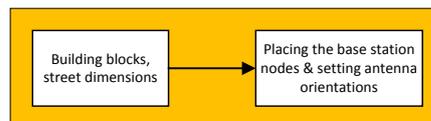
Figure 3.2: Network structure sub-function

ii. **Path loss and Power Calculation**

This section is responsible for all the necessary power related concerns. All the path losses, azimuth and elevation angles, antenna gains and received signal powers for the entire transmitting nodes are calculated. The proposed propagation models and antenna models explained in Section 3.6 and 3.7 are used in this sub-function for simulating different type of scenarios. The sub-function for the path loss and power calculation is given in Figure 3.3.

iii. **Channel Assignment**

The next step is the channel assignment process. A hop-by-hop link build up process is used in this simulation. It uses the propagation and antenna model
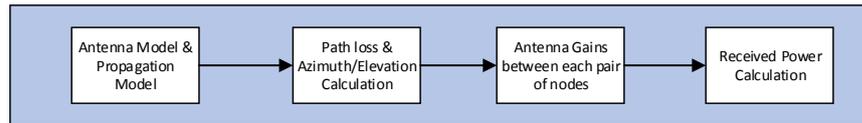
Figure 3.3: Path loss and power calculation sub-function

mentioned in second step to simulate the received signal and interference. However, there is no procedure to recheck the performance of previously constructed connections.

iv. Traffic Simulation

A large number of random time events are generated based on a Random Process to make it as realistic as possible [84]. The offered traffic here is made up of a sequence of events that can be made up of a large number of continuous packets. Once sufficient number of scenarios have been generated and then simulated, the overall performance of the system can be evaluated. A statistical approach is adopted to analyse the performance by plotting different QoS parameters. The traffic behaviour model used in this work is explained later in this chapter.

The all-inclusive simulation process including the two sub-functions is illustrated in the following flow chart. The Monte Carlo simulation will be widely used in next few chapters, and its results will compare to the results from analytical models.
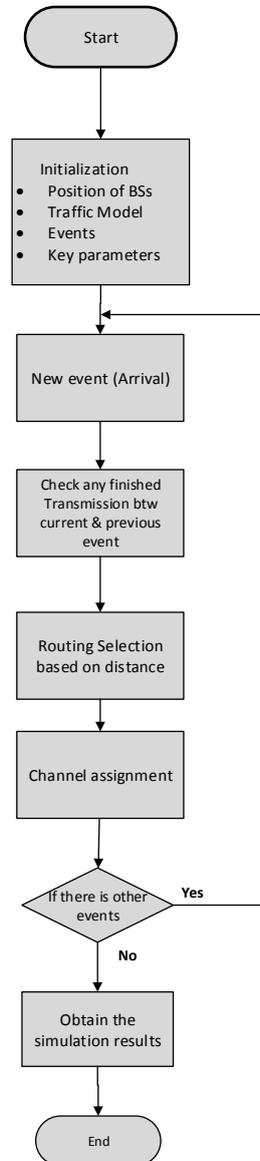
Figure 3.4: Flow chart for the Monte Carlo simulation process

3.5. Performance Evaluation Metrics

The following section describes some key parameters that are repetitively used in this research work. These parameters facilitate the comparison of the performance of the system under diverse scenarios using different algorithms. It determines whether a user can access the system or an existing user loses its current service.

i. Signal to Interference plus Noise Ratio

Signal to Interference plus Noise Ratio (SINR), which is also known as Carrier to interference plus Noise Ratio, is one of the fundamental parameters to measure the channel quality. It is defined as *"ratio of signal power to the combined noise and interference power"*[85]. The equation for calculating SINR is shown as:

$$SINR = \frac{P_S}{P_I + P_N} \tag{3.6}$$

where P_S represents received signal power, P_N stands for noise power and P_I represents the summation of signal interference from the entire interfering transmitters.

The received signal power on any link depends on the path loss, antenna gain and shadowing. It is given in logarithmic form [dB] as:

$$P_{S_{dB}} = P_{T_{dB}} + G_{T_{dB}} + G_{R_{dB}} - P_{L_{dB}} - P_{SH_{dB}} \tag{3.7}$$

P_T is the transmitter radiated power, G_T and G_R is the transmitter and receiver antenna gain respectively. P_L is the path loss factor and finally P_{SH} is the shadowing loss.

The SINR in decibels of a received signal at the receiver end is calculated as

follows:

$$SINR = P_{T_{dB}} + G_{T_{dB}} + G_{R_{dB}} - P_{L_{dB}} - P_{SH_{dB}} - P_{I_{dB}} - P_{N_{dB}} \qquad (3.8)$$

SINR will be used in this work to determine whether a new user or an existing user can receive reasonable service from the BS.

ii. Blocking Probability

The blocking probability (P_B) is the degree of the grade of service that a telecommunication system can provide. It is measured in percentage (%). This thesis focuses on the end-to-end performance as well as individual link behaviour. It is defined as, *"how often a packet/file request from a user is denied by the system"*[32]. The blocked transmissions remain in the system until they are successfully transmitted. The main reasons behind blocking are the unavailability of free channels or low SINR on free channels. Mathematically, it is denoted by the following equation:

$$P_B = \frac{N_B}{N_T} \qquad (3.9)$$

where N_B is the total number of activations being blocked and N_T is the total number of activations (including retransmissions) in the network.

iii. Average System Throughput

Throughput in a wireless network is defined as the average rate of successful data delivery. In system level research, throughput can be affected by both transmission and back off delay. It is measured in bits per sec. The average system throughput Thr_S and Throughput Density Thr_D of the entire network can be mathematically defined as:

$$Thr_S = \frac{N_{bit}(t)}{T_S} \qquad (3.10)$$

and

$$Thr_D = \frac{Thr_S}{A_S} \qquad (3.11)$$

where $N_{bit}(t)$ is the number of bits delivered within time t, which is contributed by the files delivered by all the users in the network, including those still in transmission. T_S and A_S represent the total simulation time and size of the coverage area respectively.

iv. Mean System Delay

The mean system delay D_S is also used to evaluate QoS performance of the network. The delay is calculated as the time between a users request to transmit and the completion of transmission. It is measured in seconds. Delay is particular important for real time applications (such video conferencing and live video streaming) that have require packets to be available at the destination within certain a given time period.

The average delay is evaluated by calculating the total delay for all successfully transmitted files D_{ST} and evaluating their average. Thus, the mean delay is obtained as follows:

$$D_S = \frac{D_{ST}}{F_{ST}} \qquad (3.12)$$

where F_{ST} is the total number of successfully transmitted files during the simulation.

v. Energy Consumption Rating

The energy efficiency performance of a wireless network can be estimated once the energy consumption is determined from a suitable energy or power model. The Energy Consumption Rating (ECR) has been defined as *"the energy consumed per information bit delivered"* and is measured in μJ per bit in this work. Based on the Green Radio Project definition [86], ECR is given by:

$$ECR = \frac{E_T}{SB_T} \tag{3.13}$$

where E_T is the total energy consumed in the network and SB_T is the total successfully transmitted information bit.

vi. <u>Resource Block Usage %</u>

It is the final parameter for exploring the system performance and is denoted by $RU\%$. It measures the percentage of each resource block utilized during the simulation process.

$$RU\% = \frac{RU_i}{\sum\limits_{k=1}^{n} RU_k} .100\% \tag{3.14}$$

where $RU\%_i$ represents the average percentage of the resource block i in use. The shape of $RU\%$ curve also confirms the correctness of algorithm used for the resource allocation process.

3.6. Radio Propagation Models

Signals transmitted over wireless communication channels usually experience weakening in signal strength due to opposing effects caused by natural phenomenon and man-made constructions. These channels are usually believed to be characterised by three main effects: path loss, shadowing and multipath fading. The path loss is the attenuation of the transmitted power and it is directly proportional to the distance between communicating nodes [42, 32]. Shadowing is the form of attenuation due to reflections and diffraction resulting from obstruction of the radio path by large entities while multipath fading is the variation in received signal as a result of transmitted signals arriving at the receiver through dissimilar paths with different attenuation and delay [42].

3.6.1. Backhaul Propagation Model

The main objective of this work is to increase the capacity density on the backhaul side by utilizing the path diversity in such densely populated network. Due to the presence of short range direct LOS links on the backhaul, the attenuation factors like small scale fading and multi-path effect are ignored. Measurements showed that mm-wave in LOS environments has almost identical path loss as free space [43].

Normally the received signal power decays with the square of the path length in free space [32]. In this model, line of sight connectivity is guaranteed between transmitting nodes, thus path loss is considered to define the signal attenuation. Equation 3.15 is the logarithmic form [dB] of Free Space Path Loss Model:

$$FSPL_{dB} = 20\log_{10}(d) + 20\log_{10}(f_c) + 20\log_{10}(\frac{4\pi}{c}) - G_{T_{dB}} - G_{R_{dB}} \qquad (3.15)$$

$G_{T_{dB}}$ and $G_{R_{dB}}$ is the transmitter and receiver antenna gain respectively. Where f_c is the carrier frequency in GHz and d is the separation distance in metres between any two communicating nodes. For Ultra-Small Cells network all the transceivers are 90 metres apart as shown in Figure 2.9. Thus, WINNER II B5c Line-of-Sight (LOS) Outdoor Propagation Model [87] can also be selected which is in compliance to this fixed dual-hop stationary BSs relay setup. Specifically, the path loss is calculated by:

$$PL_{dB} = 22.7\log_{10}(d) + 41.0 + 20\log_{10}(\frac{f_c}{5.0}) + \sigma \qquad (3.16)$$

σ is the shadow fading loss which is log-normal distributed with a 0 dB mean and a standard deviation of 3 dB. Many WINNER II channel models are described in [87]. These models are suitable for small BSs in metropolitan and rural areas.

3.6.2. Access Propagation Model

The path loss model for line of sight (LOS) communication between MS (Mobile Station) and BS is different from the non-line of sight (NLOS) model. The path loss for LOS scenarios is given by [87]:

$$PL_{LOS} = \begin{cases} 22.7 \log_{10}(d_1) + 41 + 20 \log_{10}(\frac{f_c}{5}) & 10 \, m < d_1 < d'_{BP} \\ 40 \log_{10}(d_1) + 9.45 - 17.3 \log_{10}(h'_{BS}) - 17.3 \log_{10}(h'_{MS}) \\ \qquad\qquad\qquad +2.7 \log_{10}(\frac{f_c}{5}) & d'_{BP} < d_1 < 5 \, km \end{cases}$$

$$(3.17)$$

where d_1 is the separation distance between MS and BS, f_c is the carrier frequency, d'_{BP} is the breakpoint distance, h'_{BS} and h'_{MS} are the effective BS and MS antenna heights.

$$d'_{BP} = 4.h'_{BS}.h'_{MS}.(\frac{f_c}{c}) \qquad c = 3 \cdot 10^8 \, m/s \qquad (3.18)$$

$$h'_{BS} = h_{BS} - 1.0 \, m \qquad (3.19)$$

$$h'_{MS} = h_{MS} - 1.0 \, m \qquad (3.20)$$

Meanwhile, the NLOS path loss is given by:

$$PL_{NLOS} = min(PL(d_1, d_2), PL(d_2, d_1)) \quad 10 \, m < d_1 < 5 \, km, w/2 < d_1 < 2 \, km$$

$$(3.21)$$

$$PL(d_k, d_l) = PL_{LOS}(d_k) + 20 - 12.5n_j + 10n_j \log_{10} d_l + 3 \log_{10}(\frac{f_c}{5}) \quad k, l \, \varepsilon \{1, 2\}$$

$$(3.22)$$

$$n_j = max(2.8 - 0.0024d_k, 1.84) \tag{3.23}$$

where $k, l\ \varepsilon\{1, 2\}$. w is the street width and when $0 < d_2 < w/2$, PL_{LOS} is applied. d_1 is the straight line distance of the BS to the centre of the perpendicular street where the NLOS MS lies while d_2 is the straight line distance of the MS to the centre of the street where the BS is located as shown in Figure 3.5.
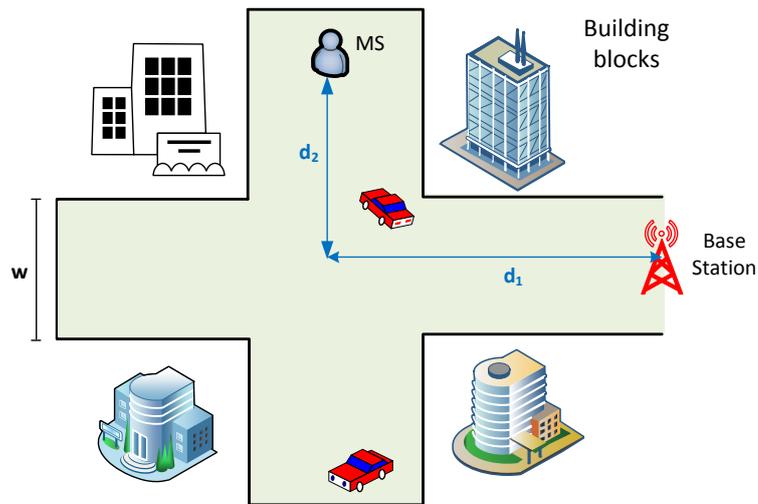
Figure 3.5: Illustration of access network Non-LOS scenario

3.7. Antenna Models

As discussed in Section 2.2.3, distance-dependent path loss, high attenuation by atmospheric gases and rain are not insurmountable challenges since the distance between the nodes considered here is typically a few tens of metres. The street canyons with high buildings in the network model also add benefits, such as interference reduction and higher frequency reuse.

Theoretically, it is also possible to use NLOS links by employing reflections from building walls and other objects. Here focus is on the multi-hop topology to introduce diversity into the system rather than the performance of individual links, so only LOS links are considered here. However, it is highly sensitive to node placement, antenna pattern and the local environment especially with the narrow beams. Electronically steerable antenna arrays are deployed at both the transmitting and receiving ends of a backhaul link [88]. Beam steering allows beam alignment during the initial stage and is resilient to the misalignments due to wind. It will also allow the antenna beams to point at other nodes dynamically, without angular constraints.

It is assumed that a maximum of 4 beams can be generated from the antenna module at the ABS/RBS end. The use of highly directional beams implies that both interference and multipath should be negligible, and the link is largely unaffected by its environment. Hence a simplified antenna model for the dual-hop backhaul network is used where only the main lobe with a fixed antenna gain is considered.

The network architecture presented in Section 2.3 indicates that the antenna model used for the access network is derived from a practical product, which is designed specifically for the BuNGee architecture [4, 37]. It can generate up to 24 beams pointing to their surrounding MSs. Each beam has a 15° azimuth beamwidth. The beam pattern of a BS antenna is shown in Figure 3.6. Only the uplink transmission is simulated in this work, from ABS to the HBS. The downlink transmission is not considered since a separate frequency band can be used, and it does not affect the key results investigated.
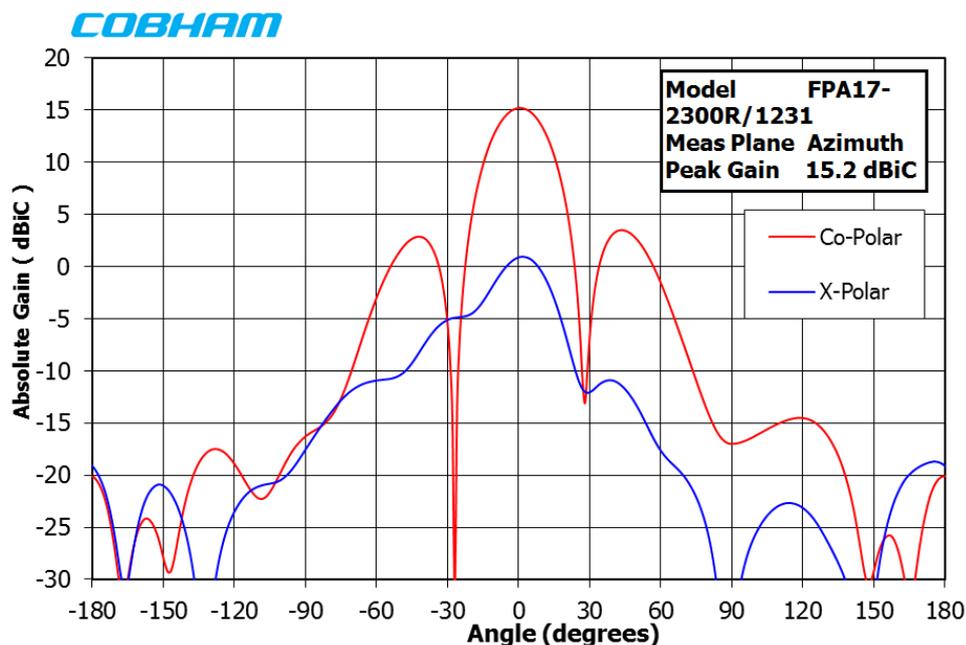
Figure 3.6: Antenna beam pattern for the BS antenna (directly reproduced from [4])

Another BS antenna radiation pattern is the 3-sector cell site pattern which is shown in Figure 3.7. It is suitable for a macro base station to provide access network coverage. Mathematically, the pattern is given by:

$$A(\theta) = -min\left[12\left(\frac{\theta}{\theta_{3dB}}\right)^2, A_m\right] \qquad where -180 \leq \theta \leq 180 \qquad (3.24)$$

θ_{3dB} is the $3\ dB$ beam-width which corresponds to 70°, and $A_m = 25\ dB$ is the maximum attenuation. The propagation model for this scenario is defined as:

$$PL = 128.1 + 37.6\log_{10}(d) \qquad (3.25)$$

where d is the BS-MS separation distance in kilometres.
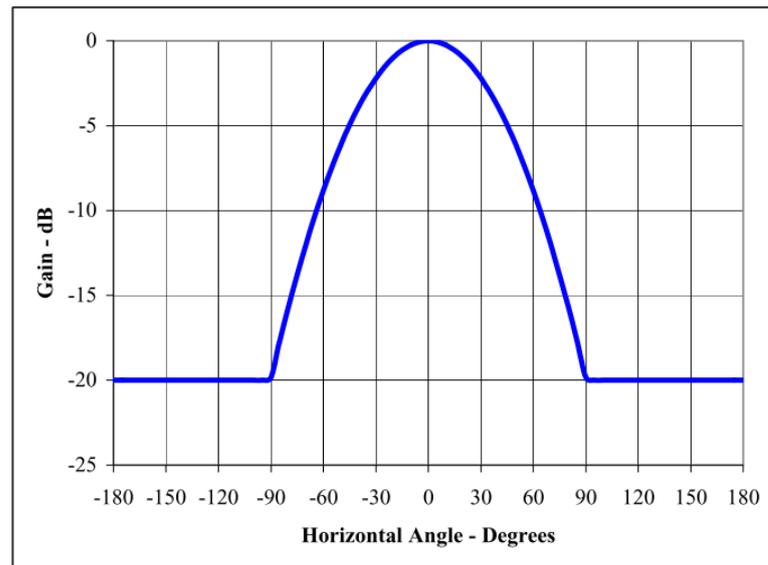
Figure 3.7: Antenna pattern for 3-sector cells (directly reproduced from [5])

3.8. Traffic Models

In traffic modelling, the start and end time of traffic flows are significantly important. They are usually characterised by the inter-arrival times between different user entities (which could be calls, packets, connections, or files) in the system and the length of time these entities exploit the resources in the wireless system. Data traffic can be modelled at the session, burst or packet levels. The session level models are usually characterized by inter- arrival and file size distribution. However, in this work same fixed file size is assumed for all the transmissions.

The data traffic model used in this work for traffic generation is Poisson Traffic Model. It is a traditional model widely accepted and used in illustrating the users arriving responses. It is a discrete probability distribution that describes the events with fixed average occurring time, which has the feature that the arrival rate of users within a period of time follows negative exponential distribution [83]. It is also known as OFF source, due to the idle (free) time slot between any two successive files with each file is denoting a single event as shown in Figure 3.8. The Poisson Traffic Model is ideally

suited for characterising the file arrivals into the system. It can serve well when the number of active users in the system is much lesser than the amount of overall users [82].
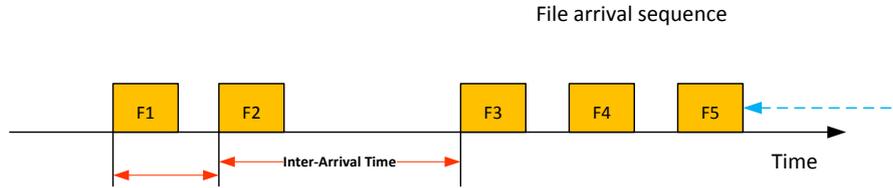
Figure 3.8: The Poisson Traffic Model

The Cumulative Distribution Function (CDF) of the exponential distribution is given in Equation 3.26 [83] with user file arrival rate is denoted by λ:

$$F(x) = 1 - e^{-(\lambda x)} \tag{3.26}$$

Mean Inter-Arrival Time $E(T_{IA})$ is equal to:

$$E(T_{IA}) = \frac{1}{\lambda} \tag{3.27}$$

The CDF in equation 3.26 can be used to produce the sample numbers following the distribution. The sequence of nth arrival times is calculated using equation 3.28 where N is a uniform random distribution [89].

$$T_{IA} = X = -\frac{\ln(1 - N)}{\lambda} \tag{3.28}$$

In a simple analysis involving a network consisting of a single base station, the accuracy of the simulation traffic model is verified. This has been done by evaluating the blocking probability of a system consisting of a single BS with a maximum of 10 resource blocks and 200 users are uniformly distributed in a circular service area. The blocking probability of the system is evaluated at the different offered loads and compared with the Erlang-B blocking probability model. It is observed that the blocking probability
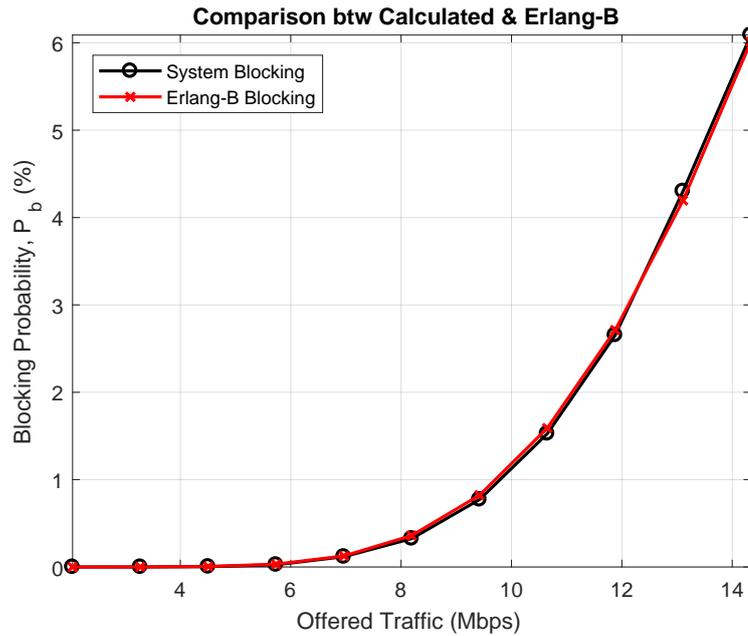
Figure 3.9: Comparison between simulation and Erlang-B Model

estimated by simulation matches the analytical model as shown in Figure 3.9.

3.9. Conclusions

This chapter described the approaches of system modelling tasks which may be applied in this thesis. We discussed the potential system modelling techniques firstly, including simulation and analytical model. MATLAB is selected as the main simulation tool in this work, and Monte Carlo approach is used to general statistically simulation results in order to obtain accurate results. An introduction to Markov processes is given which is chosen as an analytical tool for this thesis. A detail view of the system flow chart is illustrated to understand the system flow process. Next, the key measurements are defined and given to describe the system performance, such as SINR, blocking probability, system throughput, transmission delay and energy consumption rating. Different propagation models for outdoor scenarios, antenna models and the traffic model are also discussed in detail. Finally, the simulation and analytical model for a single BS are compared for the purpose of validation.

Chapter 4.   A Dual-hop Backhaul Network Architecture for 5G Ultra-Small Cells

Contents

4.1. Introduction

DUE to increasing complexity in network management and coordination among multiple network tiers, the network nodes will have the capability of self-organization (e.g., autonomous load balancing, interference minimization, spectrum allocation, power adaptation etc.) [29, 34]. Also, a MS will have simultaneous active connections to more than one BS using the same or different radio access technologies (RATs). Therefore, new network architectures need to be introduced in order to improve the flexibility and resouce utilization of the overall system which is the first key objective of this thesis hypothesis.

In this chapter, the mm-wave dual-hop backhaul network architecture which exploits the existing path diversity in wireless networks is introduced to deliver a QoS guarantee. The rest of the chapter is organised as follows: in Section 4.2, the concept of network densification is introduced to achieve high link capacity. The detailed backhaul network model is discussed in Section 4.3. The comparison between simplified dual-hop backhaul network and the Erlang-B model is illustrated in Section 4.4. The comparison with Erlang-B Model is presented in Section 4.5. The algorithm for solving the system of linear equations is mentioned in Section 4.6. It is further used extensively in Chapter 6 for solving 2-Dimensional Markov Model. In Section 4.7, results are illustrated from a simulation model. Finally, conclusions are given in Section 4.8.

4.2. Network Densification

The enhancement of wireless system capacity ever since the discovery of the radio right up to the present can be attributed to three main features (in decreasing order of impact): increase in the number of wireless communicating nodes, enhancement in link capacity and increased use of the frequency spectrum [28].

For a simple visualization of the key factors governing the performance of a cellular system, consider the following equation based on the capacity of an Additive White Gaussian Noise (AWGN) channel. The throughput capacity (R) of a user in a cellular system is upper-bounded by:

$$R = m \left(\frac{W}{n} \right) log_2 \left(1 + \frac{P_S}{P_I + P_N} \right) \qquad (4.1)$$

where W denotes the BS signal bandwidth, the integer parameter n (load factor) denotes the number of users sharing the given base station, the integer parameter m (spatial multiplexing factor) denotes the number of spatial streams between a base station and user device(s), and P_S denotes the desired signal power, while P_I and P_N denote the interference and noise power respectively, at the receiver [28].

Clearly, the signal bandwidth W can be increased by using more spectrum, which leads to a linear increase in data capacity. The load factor $n \ (\geq 1)$ can be decreased through cell splitting, which involves deploying a larger number of BSs, and ensuring that user traffic is distributed as evenly as possible among all the BSs. The spatial multiplexing factor m can be increased using a larger number of antennas at the BS and connecting devices.

The above ingredients for wireless capacity enhancement may be observed under a common umbrella of network densification. Network densification is a combination of spatial densification (which increases the ratio m/n) and spectral aggregation (which increases W). Spatial densification is realized by increasing the number of antennas per node (user device and base station), and increasing the density of BSs installed in the given geographic area, while ensuring nearly uniform distribution of outdoor users among all base stations which is one of the primary objective of this thesis. Finally, the spectral aggregation refers to bringing together different sections of spectrum allocation, spanning all the way from 500 MHz into the mm-wave bands (20 -300 GHz).
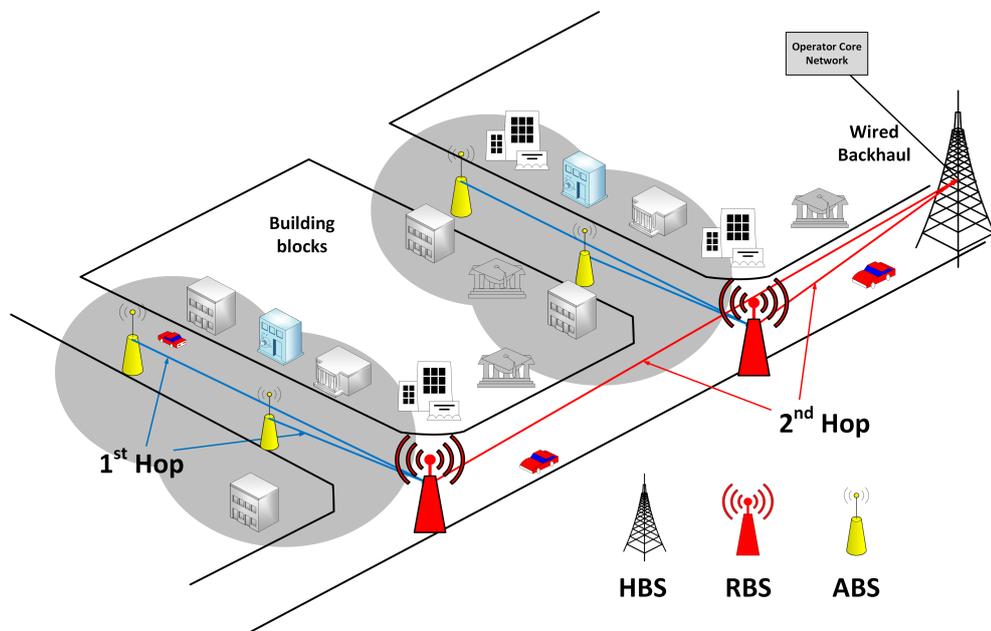
Figure 4.1: A simplified dual-hop backhaul network model

Figure 4.1 presents an in-depth view of the proposed network architecure which is pre-

dominantly based on dual-hop short distance links. It visualizes the intermediate relay points between backhaul communicating nodes. To make this network a high capacity density wireless system, the BSs are densely deployed and equipped with highly directional antennas which use short range mm-wave links. It is very much feasible for densely populated urban environments to achieve higher throughput capacity and better QoS [13].

4.3. The Dual-hop Wireless Backhaul Network

The network architecture which is proposed in this chapter is mapped on to an extended Manhattan grid in its simplest form with a dense deployment of BSs as shown in Figure 4.2. The distinctive trademark of this network architecture which makes it stand out over the traditional backhaul network is its two hop backhaul feature. In this layout, the RBS is introduced to relay the data traffic from ABS to the HBS with the number of hops to be equal to two anywhere in the network. As mentioned in Chapter 2, these RBSs are incapable of providing direct connectivity to the mobile users.

One of the techniques of network densification is to develop a dense network which comprises of a small geographical area as shown in Figure 4.2. The network consists of ABSs, RBSs and HBSs. These nodes are installed outdoors along the street at each street crossing (e.g. street lamps, traffic lights) separated by a short distance. The RBSs relay the data traffic between two communicating nodes.

Only one HBS is deployed on each corner of the block. These HBSs act as aggregation points and are connected to the Operators Core network via fibre as shown in Figure 4.2. This network architecture is a single zone representation of a larger cellular network. It can be scaled up by replicating and combining these single zones to make a large dual-hop backhaul network layout. The corner HBSs will then provide two further beams each to serve the neighbouring blocks.

This dual-hop feature ensures that the latency requirements of future 5G networks can
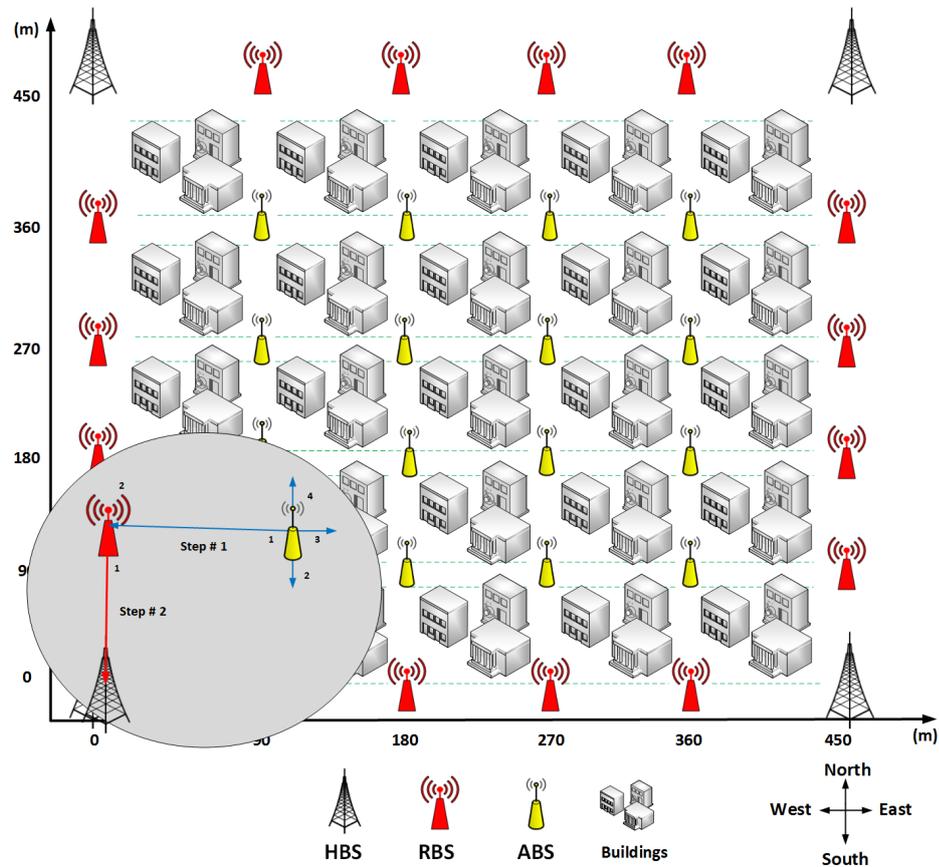
Figure 4.2: The dual-hop backhaul network layout

be met, while providing increased flexibility. These dual-hop connections are provided by highly directional antennas mounted on each ABS, RBS and HBS transceivers. The Half Power Beamwidth (HPBW) of these antennas is very narrow which significantly improves the link budget performance, while also mitigating the interference, thereby enhancing the throughput capacity.

Each ABS is capable of generating up to a maximum of four beams for transmission with main-lobe separation of 90° between two beams. At each RBS there is only one beam pointing towards four in line ABSs for reception and two more beams for transmission with main-lobe separation of 180° pointing towards the two HBSs deployed on the same street. However, each HBS is equipped with two beams with mainlobe separation of 90°. Each beam is capable of reception from the corresponding two RBSs deployed on the same street. The link between ABS to RBS is the first hop of the dual-hop connection as shown in Figure 4.1 while the link between RBS to HBS is the

second hop.

In this dual-hop wireless backhaul network, N ABSs are assumed. Each ABS has L directional narrow beams for backhaul connectivity. Each ABS has also L wide beams for the access network providing services to K MSs. The ABS transmits power for the access and backhaul network can be expressed as P_A^A and P_B^A respectively.

$$P_A^A = \begin{pmatrix} P_{A,1}^{A,1} & \cdots & P_{A,1}^{A,L} \\ \vdots & \ddots & \vdots \\ P_{A,N}^{A,1} & \cdots & P_{A,N}^{A,L} \end{pmatrix}$$

$$P_B^A = \begin{pmatrix} P_{B,1}^{A,1} & \cdots & P_{B,1}^{A,L} \\ \vdots & \ddots & \vdots \\ P_{B,N}^{A,1} & \cdots & P_{B,N}^{A,L} \end{pmatrix}$$

Where $P_{A,n}^{A,m}$ represents the access transmit power of beam m of ABS n whilst $P_{B,n}^{A,m}$ represents the backhaul transmit power of beam m of ABS n. For the N RBSs, which have 2 beams for backhaul. Its associated transmit powers can be expressed as:

$$P_B^R = \begin{pmatrix} P_{B,1}^{R,1} & P_{B,1}^{R,2} \\ \vdots & \vdots \\ P_{B,N}^{R,1} & P_{B,N}^{R,2} \end{pmatrix}$$

Further assumptions in this work are given below:

- Given that data is a prodominant transfer transmission method today, it is important to use a traffic model that can correctly capture file arrivals and departures from a source. One of the particular models that is in wide spread use is the 3GPP FTP Traffic Model 1 [90] with an fixed file size of 2 Mbytes. The file arrival rate is modelled as Poisson distribution with a fixed mean arrival rate of lambda (λ) as explained in Section 3.8.

- In order to see the impact of routing on these dual-hop backhaul links, it is assumed that RBS is only capable of relaying data traffic from the ABS towards HBS. The

impact of data traffic generated by the RBSs will be investigated in the future work.

- Only the uplink transmissions of the backhaul from the ABSs towards HBSs via RBS are considered with a fixed transmission rate and there is no interruption in the network.

- A perfect access network is assumed which means that blocking and delay are caused solely by the backhaul network. With this assumption in place, any negative effect that is caused by the access network can be neglected during the simulation process.

4.3.1. The Routing Process

The system model illustrated in Figure 4.2 is a modified version of Figure 2.9. A zoomed-in section is introduced in Figure 4.2 to explain the two steps of path selection algorithm. There are 16 ABSs, 16 RBSs and 4 HBSs in total serving a 450 x 450 m^2 area. The link between the ABS and RBS is the 1st hop while the link between the RBS and HBS is the second hop in this dual-hop backhaul network layout. It is worth mentioning here that the backhaul interference is significantly reduced due to the following reasons:

   i. Below roof-top deployment of ABSs/RBSs.

  ii. Installing highly directional mm-wave antennas with $90°$ of angular separation.

 iii. Street canyon effect from high-rise buildings.

- ## 1st Step

This step deals with the first link of the dual-hop connection. When a new event arrival occurs, the transmitting ABS has 4 RBSs to choose from. However, it selects the closest RBS deployed on the East-West Street using the first transmitting beams. Each beam is oriented in any one of the four directions (North, South, East, and West). If there is no free resource block with closest RBS (marked 1) then the ABS selects the next closest RBS on the North-South Street (marked 2) and so on as illustrated in Figure 4.3.
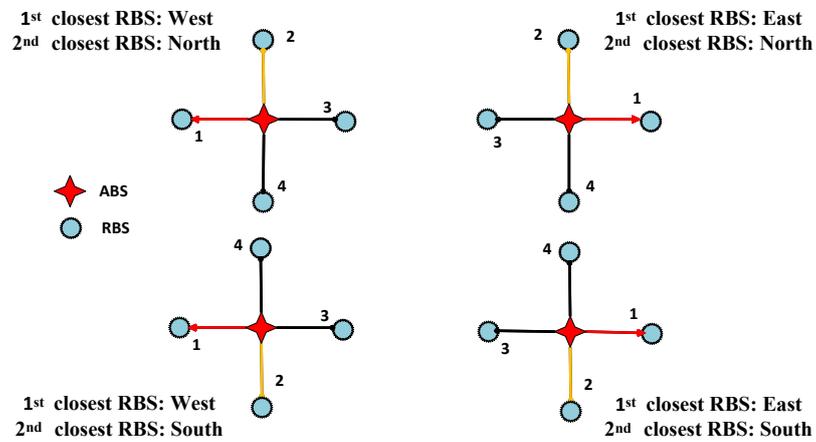
Figure 4.3: 1st Step of the routing process

- **2nd Step**

  This step is associated with the second link of the dual-hop connection. The RBS also selects a free resource block from any one of the two HBSs deployed on the same street where the transmitting RBS is deployed. This time the RBS selects a free resouce block from the closest HBS (marked 1) as shown in Figure 4.4. In case, if no free resource block is available from the closest one, the RBS moves to the second HBS on the opposite end of the street (marked 2) as shown in Figure 4.4. Once a free resource block is available on both hops of the connection, the end-to-end connection is established and the data transmission gets under way.
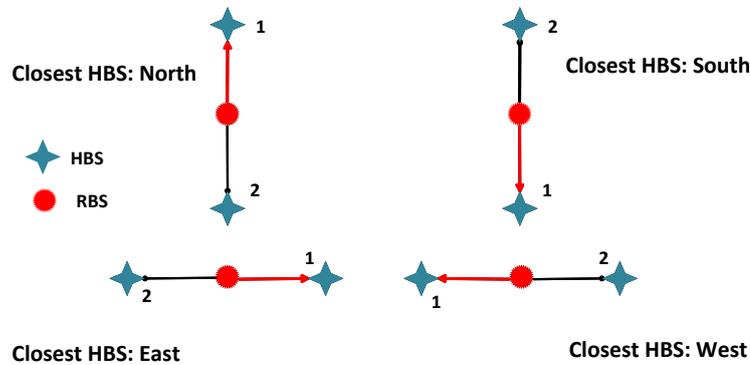
Figure 4.4: 2nd Step of the routing process

4.3.2. Resource Allocation Process

With the deployment of highly directional antennas using mm-wave and having short distance links, the link budget is significantly improved, thus allowing the same frequency bandwidth to be used on both the links. A Dynamic Channel Allocation (DCA) strategy is used for resource assignment where the resource blocks are assigned in sequence on First Available (FA) basis at both links separately. For simplicity, FA is choosen because it is not the main focus of this work. Furthermore, the resource assignment scheme is centralised where two separate RB state tables are maintained, one for all RBSs and other for all HBSs.

A file transmission is only initiated by the source ABS and transmitted to the destination

HBS via RBS. A Dual-hop end-to-end connection can only be established if there is at least one free resource block per hop [91], otherwise the transmission is blocked as shown in the flow diagram in Figure 4.5. This flow chart is just focused on the resource allocation process as compared to the one given in Chapter 3 which illustrated the whole simulation model. There is end-to-end signalling mechanism on the backhaul network which makes sure that the SINR is above the required threshold in order to establish the dual-hop connection. In case of blocking, the retransmission only starts after a random time interval, provided that a free good quality RB is available on both the links. Any link can be assigned multiple RBs, once there is more than one file transmission.
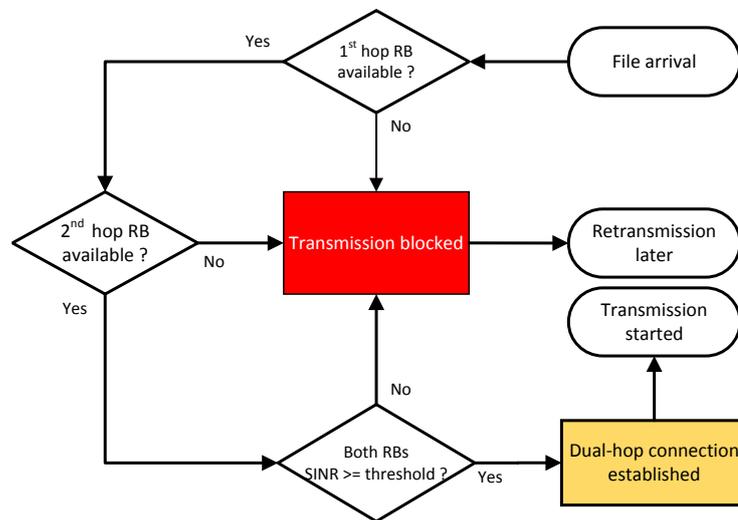
Figure 4.5: Illustrating the flow of resource allocation process

4.4. Erlang-B Comparison

A Dual-hop link model which is shown in Figure 4.1 is compared with the Erlang-B Model which is ideal for situations when there is no interference or frequency reuse in the system. The number of resource blocks are 20 for this comparison. The 3GPP File Traffic Model 1 [90] is designed as such that a single ABS can have any number of simultaneous file transmissions which can be approximated to infinite user population.

The bottleneck which occurs on the second hop only is analytically verified using the Erlang-B Model. As shown in Figure 4.6, there is a perfect match between Monte-Carlo simulation and the Erlang-B Model.
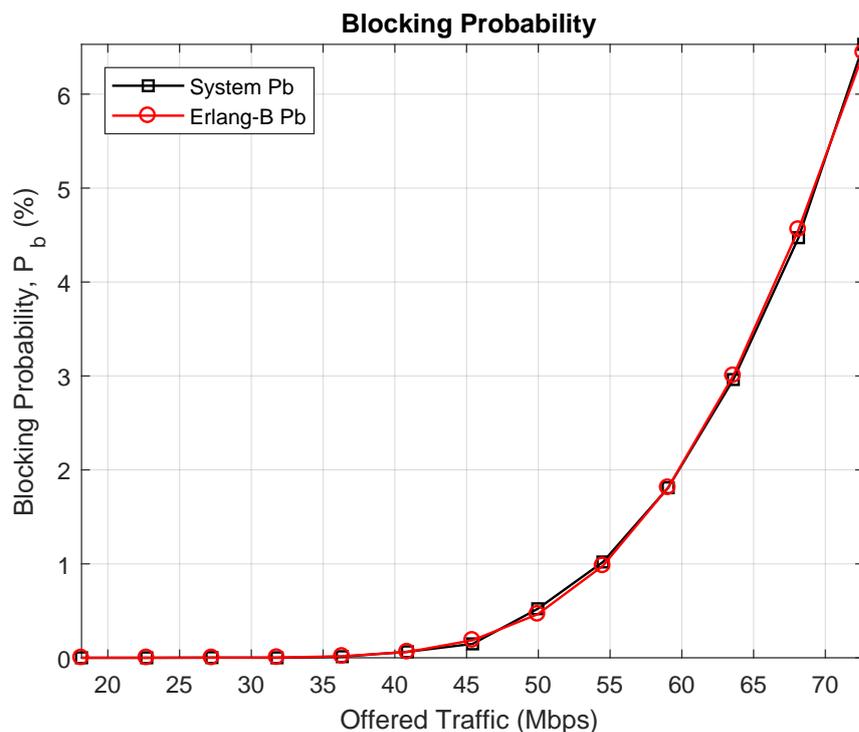
Figure 4.6: Comparison between simulated model & Erlang-B Model

As illustrated in Figure 4.7, the system throughput graph from simulation becomes restricted as the offered traffic increases beyond 50 Mbps. A line plotting the offered traffic, marked "comparison line" in the legend, is added to make clear where the throughput drops below the system offered traffic. This loss in system throughput is due to the bottleneck on the second hop. Also, the Erlang-B Model works on the princi-

ple of **Blocked Calls Cleared**, hence no retransmission or queuing is assumed for this analytical comparison.
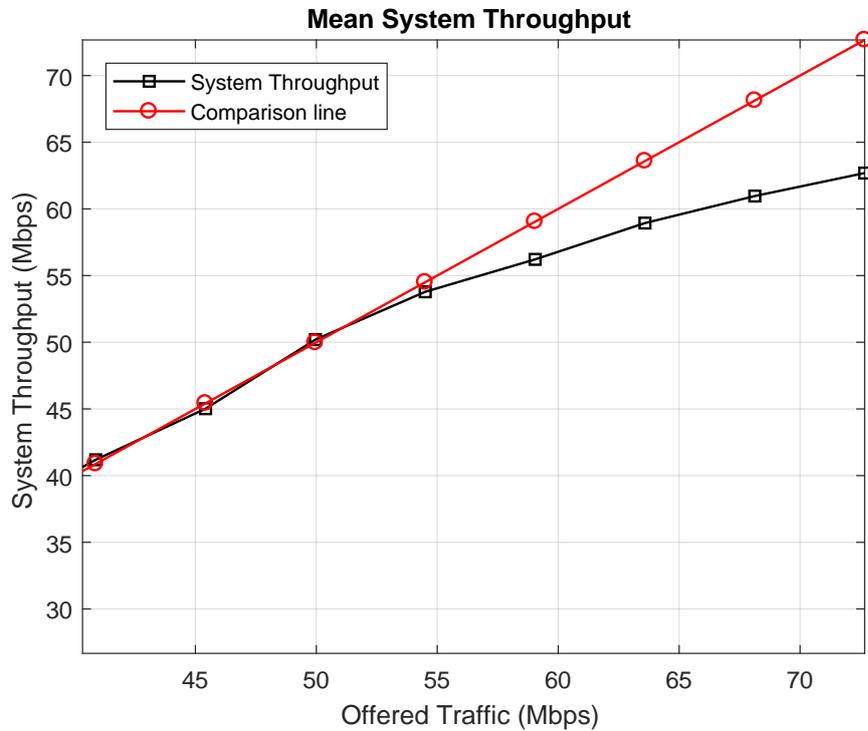
Figure 4.7: Illustrating system throughput saturation

4.5. Mathematical Analysis

The complexity of the analysis is increased in later chapters, therefore, here we introduce the key concepts that will be built on in the later part of this thesis. The behaviour of Erlang-B Model can be described using a 1-Dimensional Markov chain as shown in Figure 4.8. Each node in the diagram denotes a state which in turn denote the number of RBs being occupied or busy. The transition arrows between each state represent a single arrival or departure, such that the number of active RBs increases or decreases by one.

State n represents the state when there is no RB available on the second hop of the dual-hop connection. It is assumed that both the arrival and departure processes are negative exponential distributed with an infinite population. Thus, the arrivals (λ) and

departures (μ) when the system is in state k are given by:

$$\lambda_k = \begin{cases} \lambda & \text{for } k < n \\ 0 & \text{for } k \geq n \end{cases} \tag{4.2}$$

$$\mu_k = k\mu \qquad \text{for } k = 1, 2 \text{ .... } n \tag{4.3}$$

which represents a birth-death process [83], which means that at any given instant of time there will be only one arrival or departure. Furthermore, in statistical equilibrium, the transition rate into state j equals to the transition rate out of state j. The system will be in state j with state probability $P(j)$, i.e. the probability of observing the system in state j at a random point in time.
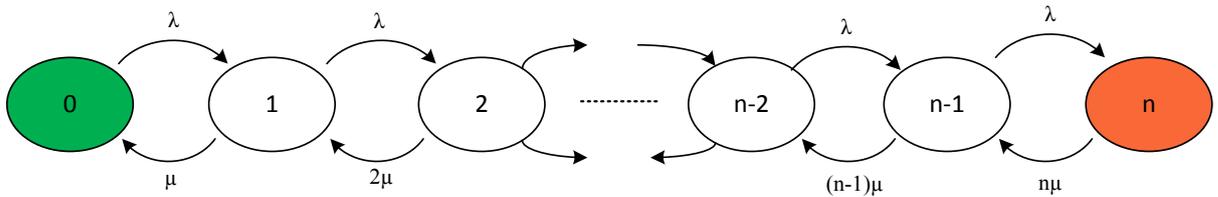
Figure 4.8: 1-Dimensional Markov chain for Erlang-B Model

For state (0), we have

$$\mu.P_1 = \lambda.P_0 \tag{4.4}$$

For state (j), $0 < j < n$,

$$\lambda.P_{j-1} + (j+1)\mu.P_{j+1} = (\lambda + j\mu).P_j \tag{4.5}$$

Finally, for state (n),

$$n\mu.P_n = \lambda.P_{n-1} \tag{4.6}$$

Regarding the boundary states, no state exist when $j < 0$,

$$P(-1) = 0 \tag{4.7}$$

Similarly, no state exist when $j > n$,

$$P(n+1) = 0 \tag{4.8}$$

As the system will always be in a state, the state probabilities must also satisfy the normalization equation:

$$\sum_{j=0}^{n} P(j) = 1 \tag{4.9}$$

Meanwhile, the blocking probability (P_b), when no free resource block is available equals to the probability of the system in state n.

$$P_b = P(n) \tag{4.10}$$

It is worth mentioning that a very basic Markov Model is analysed in this section. However in Chapter 6, a 2-Dimensional complex Markov analysis is performed for a coexistence scenario in a multi-tier heterogeneous network.

4.6. System of Linear Equations

The main purpose of the Markov analysis is to obtain the system probability at each state. There are multiple ways for solving a system of n linear equations with unknown X. One of which is an iterative technique known as **Gauss-Seidel Method** [92].

It is also known as the **Liebmann method** or the method of successive displacement [92], is an iterative method used to solve a linear system of equations. Convergence is only guaranteed if the matrix is either diagonally dominant, or symmetric and positive

definite.

$$\mathbf{AX} = \mathbf{B} \tag{4.11}$$

It is defined by the iteration

$$\mathbf{L}_*\mathbf{X}^{(k+1)} = \mathbf{B} - \mathbf{U}\mathbf{X}^{(k)}, \tag{4.12}$$

where $\mathbf{X}^{(k)}$ is the kth iteration of $\mathbf{X}$, $\mathbf{X}^{(k+1)}$ is the $k+1$ iteration of $\mathbf{X}$, and the matrix $\mathbf{A}$ is the coefficient matrix which is decomposed into a lower triangular component $\mathbf{L}_*$, and a strictly upper triangular component $\mathbf{U} : \mathbf{A} = \mathbf{L}_* + \mathbf{U}$.

$\mathbf{A}$, $\mathbf{X}$ and $\mathbf{B}$ in their components form:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Then the decomposition of A into its lower triangular component and its strictly upper triangular component is given by:

$$\mathbf{L}_* = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

The system of linear equations may be rewritten as:

$$(\mathbf{L}_* + \mathbf{U})\mathbf{X} = \mathbf{B} \tag{4.13}$$

$$\mathbf{L}_*\mathbf{X} = \mathbf{B} - \mathbf{U}\mathbf{X} \qquad (4.14)$$

The Gauss-Seidel method now solves the left hand side of this expression for $\mathbf{X}$, using previous value for $\mathbf{X}$ on the right hand side. Analytically, this may be written as:

$$\mathbf{X}^{(k+1)} = \mathbf{L}_*^{-1}(\mathbf{B} - \mathbf{U}\mathbf{X}^{(k)}). \qquad (4.15)$$

However, by taking advantage of the triangular form of $\mathbf{L}_*$, the elements of $\mathbf{X}^{(k+1)}$ can be computed sequentially using forward substitution [92]:

$$\mathbf{X}_i^{(k+1)} = \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}\mathbf{X}_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}\mathbf{X}_j^{(k)} \right), \qquad i = 1, 2, ..., n. \qquad (4.16)$$

The procedure is generally continued until the changes made by an iteration are below some tolerance threshold. The convergence property of this method is dependent on the coefficient matrix $\mathbf{A}$. Namely, the procedure is known to converge if matrix $\mathbf{A}$ is strictly diagonally dominant. Also, an in-depth view of a 2-Dimensional Markov Model is presented for a multi-tier heterogeneous network in Chapter 6 where this technique is thoroughly utilized.

4.7. Simulation Results

In earlier sections, a simplified version of the network with only one dual-hop backhaul setup is used, so that to make it possible to compare with the Erlang-B model. Here a full scale network is analysed through monte-carlo simulation alone and the results are derived. The urban street canyon scenario proposed in Section 4.2 is simulated to test the performance of the routing process explained in Section 4.3.1 for ultra-dense outdoor small cells.

The Free Space Path Loss Model given in Equation 3.15 is used in this model for cal-

Table 4.1: Simulation Parameters for dual-hop backhaul setup

Parameters	Values
Coverage area	450x450 m
Building size	75x75 m
Street width	15 m
Building height	6 m
SINR threshold	1.8 dB
Log-normal shadowing factor	3 dB
Carrier frequency	75 GHz
Antenna gain	38 dB
Number of HBS/RBS/ABS	4/16/16
Frequency Spectrum	500 MHz
Resource blocks	200
Inter-arrival time	Exponential Distribution
Noise floor	-114 dBm/MHz
Mean file size	2 MB
Transmission power	27 dBm
Iteration per offered traffic	200k

culating the path loss since it is in accordance with this type of dual-hop backhaul relay setup. The routing technique is performed for every iteration to establish an end-to-end connection. Other parameters used in this simulation are summarized in Table 4.1. The performance of this network is evaluated using discrete time event based simulation. The simulation runs for 200k iterations for each traffic load, and the measurements are taken when the system is relatively in stable state. Files arrive into the system with exponentially distributed inter-arrival times and leave the system once the file transmission is completed.

The QoS parameters of this proposed backhaul network are evaluated in terms of Blocking Probability and end-to-end Mean Transmission Delay. The interruption of the transmission link due to worsening link quality in terms of SINR is a minor factor compared to the overall system, due to the highly directional nature of the antennas and mm-wave frequency band which results in minimal interference.

The blocking occurs in this network on either link only due to the unavailability of a free RB. The adequate limit for the blocking probability is assumed to be 5%. All the performance parameters are plotted against six different routing combinations. For

this simulation the offered traffic, specifically the user demanded traffic is divided into three levels. This dual-hop routing approach could result in many possible combinations. Here we examine the impact when the number of possible routes is restricted. The routing levels available are specified using two numbers, the first integer indicates the maximum number of routes that a transmitting ABS has for the first link towards the intermediate RBS while the second integer is associated with the number of routes that the RBS has for the second link towards the destination HBS. The route options are limited on the basis of the shortest distance to all 4 possible HBSs. All the 6 combinations of routing are simulated for three different ranges of offered traffic and their corresponding blocking probability values are shown in Figure 4.9.

From Figure 4.9, it is evident that it is the second link where the greater routing choice to the HBS improves the performance. It also validates the assumption that the bottleneck effect is taking place on the second link at high occupancy levels, which increases the bottleneck. However, there is still slight improvement in the performance if the number of routes on the first link is increased.

For the baseline scenario, when there is no routing and path diversity (1-1) available on
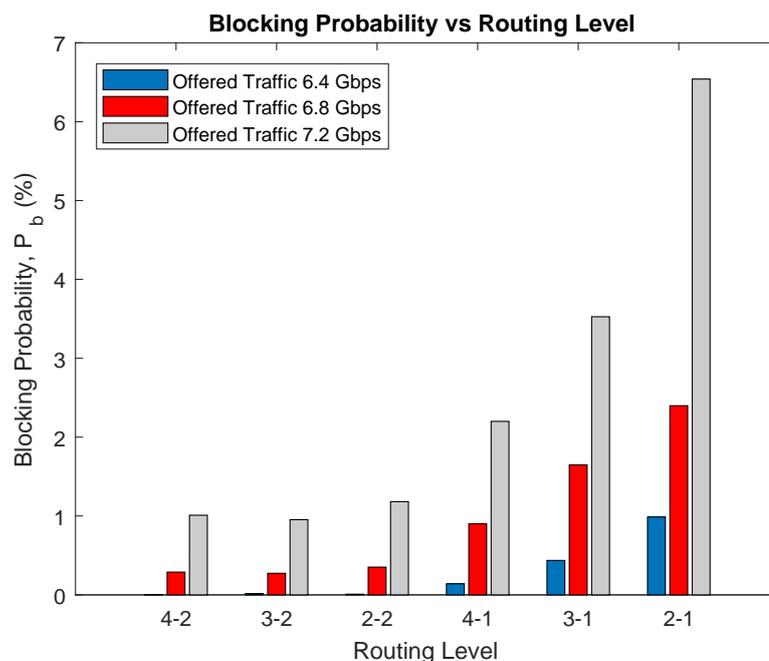
Figure 4.9: Routing impact on network blocking
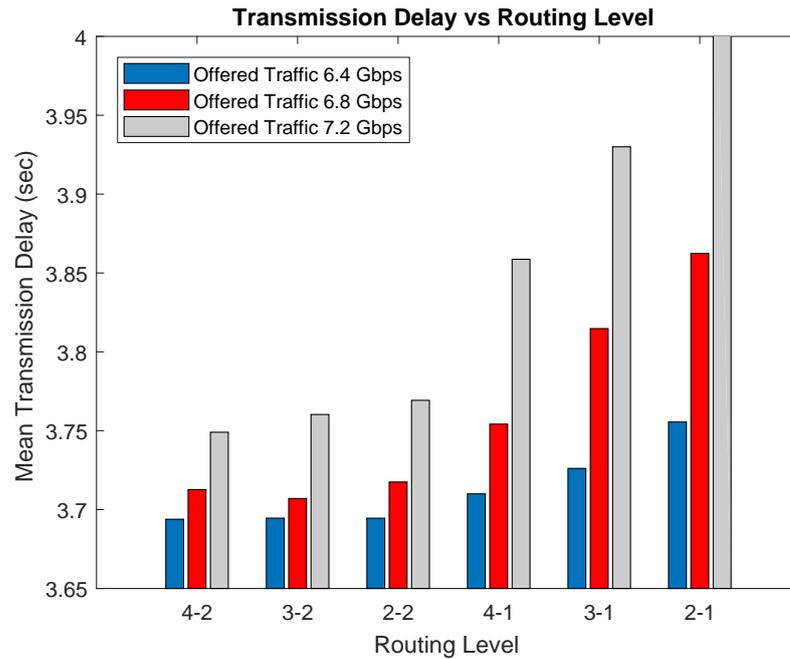
Figure 4.10: Routing impact on transmission delay

both the links, the network is only capable of providing the same QoS up to an offered traffic of 5 Gbps. It is this routing and path diversity which reduces the bottleneck on the second link and improves the blocking from 6.5% to 1.1% at high traffic loads.

Figure 4.10 on the other hand, illustrates the impact of this dual-hop network along with the routing process in terms of Mean Transmission Delay. In case of no blocking in the network, the mean delay is equal to the transmission time. Once again, at higher offered traffic the transmission delay significantly increases when there is only one routing option on the second link. Figure 4.10 also validates the fact that the bottleneck is taking place on the second hop as shown by blocking in Figure 4.9. It must be noted that the baseline routing level 1-1 is not shown in these simulation results because it is incapable of handling anything above 5 Gbps.

Figure 4.11 shows RBs Usage %. It illustrates that the starting RBs are used the most but as the RB index increases, RU% decreases which is in compliance with the Resource Allocation Scheme explained in Section 4.5. It also confirms that the routing process is independent of Resource Assignment. Those RBs which exceed 118 are used only for high offered traffic levels which give rise to the curve shape in the middle of the graph.
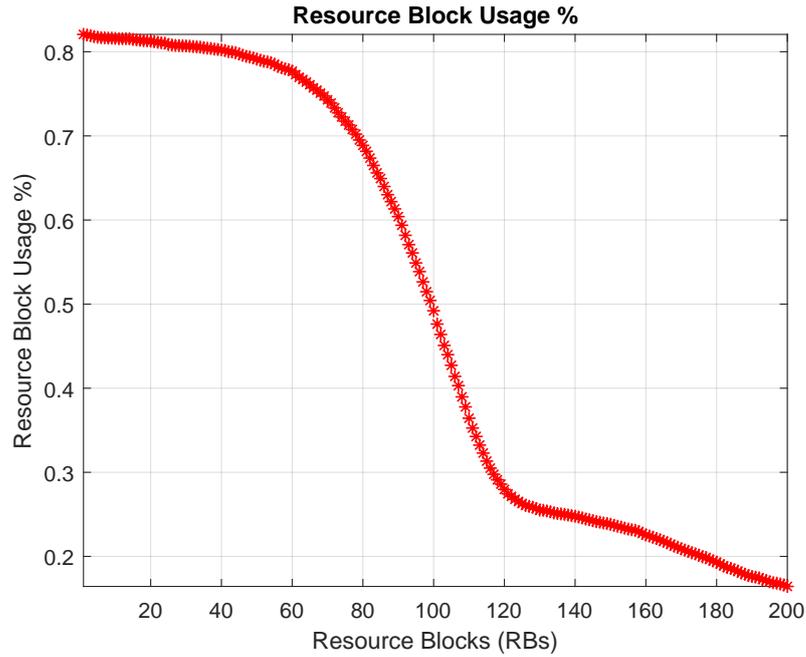
Figure 4.11: Illustrating the First-Available allocation process

Simulation results from Figure 4.9 to Figure 4.11 clearly demonstrate the level of expected improvement in GoS that can be achieved by introducing more routing and path alternatives into the dual-hop backhaul network. Much higher system throughput is achieved by having such close deployment of BSs. Also installing highly directional antennas using mm-wave band for such short distance links, the level of interference can be significantly mitigated.

4.8. Conclusions

This chapter proposes a dual-hop backhaul network architecture for ultra-small cells which put to use the available path diversity in the network. This provides significant flexibility in limiting the number of aggregation points where fibre needs to be present while ensuring that latency constraints can still be met. With this dual-hop approach it is shown that bottlenecks are restricted to the second hop links but routing flexibility mitigates this bottleneck impact to a larger extent. Furthermore, it is also analytically verified that with this type of dual-hop architecture, the hop closest to the aggregation point is experiencing the bottleneck. It is this unique dual-hop feature of the backhaul

network which makes it distinctive from the traditional single-hop backhaul networks. A detailed overview of the Gauss-Seidel Algorithm is explained for solving a square system of linear equations.

By utilizing the path diversity available in the network, the level of performance is improved which is the first primary objective of this thesis. Due to this path diversity, some of backhaul base stations can be turned off at low occupancy levels while still maintaining the same QoS thus minimizing the energy expenditure of the whole network.

Chapter 5.   Energy-Aware Topology Management for 5G Dual-hop Ultra-High Capacity Backhaul Networks

Contents

5.1. Introduction

As discussed in Chapter 1, new flexible topology management mechanisms need to be adopted in backhaul network in order to enhance the flexibility and resource management of the overall system. This project aims to enhance the capacity density while reducing the energy consumptions. To achieve such networks cost effectively, one option is to backhaul the data traffic via multihop links connecting a series of base stations to an aggregation point. However, a high density of BSs means high power expenditure, which may not be sustainable if existing design strategies are continued. Primarily, these ultra-dense networks are designed from the QoS point of view, where energy savings are not the prime objective.

The objective of this chapter is to exploit the available path diversity in a dual-hop backhaul network by using an energy-aware topology management strategy. The purpose is

to improve the throughput capacity plus, reducing the energy consumption especially at low traffic loads which is the second key objective of this thesis. As mentioned in Section 1.2, all the communicating nodes are equipped with highly directional antennas using mm-wave that can redirect the uplink data towards different BSs.

The rest of the chapter is organised as follows: The details about the topology management and its working mechanism is discussed in Section 5.2. To understand, the concept of path diversity available in the network, it is explained in Section 5.3. For improving the energy efficiency, the energy model is presented in 5.4. The simulation results for the whole network are illustrated in Section 5.5. Finally, conclusions regarding the topolgy management strategy are given in Section 5.6.

5.2. Topology-Management Strategy

The general backhaul network architecture is introduced in Chapter 4 where the pros and cons of the dual-hop mm-wave short distance links are discussed. In this chapter, the focus is on the reduction of energy consumption aspect of this densely deployed ultra-small cells network. The network architecture shown in Figure 4.2, the overall system throughput capacity can be significantly increased when the system traffic demand is at a high level. Due to this dense deployment, it is not an effective strategy in terms of energy consumption to keep all BSs in ON state during off peak periods. The highest peak occurs around 11:00 while the lowest point ranges from 01:00 to 05:00 in the morning as shown in Figure 2.11 [24]. The peak network traffic takes up only a small portion of the whole day plus the variation of data traffic both in time and spatial domain makes some portion of a network to be hot-spots and some to be not-spots, mainly due to the user accumulation and the bursty nature of voice and video intense applications. Therefore, appropriate rules are formulated especially in low to medium traffic scenarios to switch off some portion of the network to achieve energy savings.

This energy-aware topology management scheme dynamically regulates the structure

of a network to reduce the energy consumption without compromising the system QoS. It is accomplished by fine-tuning the different states (SLEEP / ACTIVE / ON) of the RBSs and HBSs. The BS remains in ACTIVE state until all the ON state BSs are 100% loaded. As shown by the state transition diagram in Figure 5.1, the topology management scheme consists of three main states and is known as the Sleep Mode Algorithm [93].

Primarily, when the offered traffic reaches to some predefined threshold and the network is in steady-state, those RBSs and HBSs which are still in ACTIVE state with no load are turned off. However, the ON state RBSs and HBSs cannot switch to SLEEP state unless they fulfil several conditions. Different values for the load threshold L_{Active} can be selected but in this work L_{Active} is set to 90% in order to allow a BS enough time to switch from SLEEP to ACTIVE state. However, L_{Sleep} is the load threshold defined for the number of users in service at which a BS is allowed to make a backward transition from an ACTIVE state to SLEEP state. In this case, L_{Sleep} is put to 50% for deactivating the BS. All the auxiliary conditions for switching from one state to another are explicitly mentioned in Figure 5.1.

- **Load on BS is still equal to zero**
- **BS is idle for further T$_{min}$**
- **No ON-State BS with load >= L$_{Sleep}$**

- **Load on BS is equal to zero**
- **BS is inactive for at least T$_{min}$**

SLEEP   ACTIVE   ON

Extra energy is required to make this transition   **No energy is required for this transition**

Any BS with load greater than L$_{Active}$ **All ON-State BSs are 100% loaded**
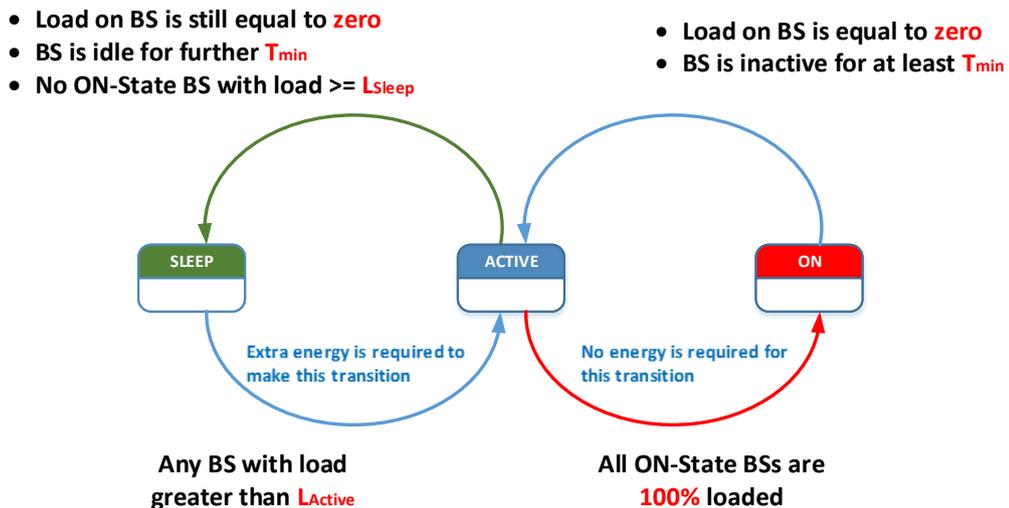
Figure 5.1: State Transition Diagram for TM scheme

For this scheme, T_{min} is calculated using Equation 5.1 from the actual offered traffic to keep RBSs and HBSs in the ON state for a period equal to it. It is used to sustain the

system by avoiding repetitive switching between any two states [69].

$$T_{min} = \frac{L_{Sleep}.N}{\lambda} \qquad (5.1)$$

where N is the total number of RBSs and HBSs used in the architecture, λ is the user mean arrival rate equal to $1/T$. The offered traffic is the user-requested traffic, whereas the system throughput is the traffic successfully carried by the network which is either less than or equal to the offered traffic. The T_{min} is used to forbid all the ON state HBSs from going into SLEEP mode straight away when L_{HBS} is equal to zero.

Moreover, if the above-mentioned conditions are met, the BS first enters into ACTIVE state for duration of T_{min}. In this state, the sleeping BS does not permit any new connection provided that there is a free resouce block available with an ON state HBS. Further assumptions related to this chapter are summarized in Section 4.3.

5.3. Path Selection Diversity

The major problem with network densification is the bottleneck which is taking place at the data aggregation points in the network. One way of solving this bottleneck is through introducing path diversity in a multi-hop fashion on the backhaul side of the network. The existing path diversity in the dual-hop network as shown in Figure 5.2 is utilized for reducing this bottleneck on second hop which in turns improve the throughput capacity. Due to this diversity, some portion of the network can be turned off under low traffic demands.

The graph $G(V; E)$ represents the directed graph for the path diversity of a single ABS node. It illustrates the available paths in the network for each ABS towards the aggregation point. $V(G)$ and $E(G)$ represent the set of vertices and directed edges respectively in this digraph. They can be expressed as the following:

$$V_i \in V, \qquad \forall_i \in [1, 13]$$
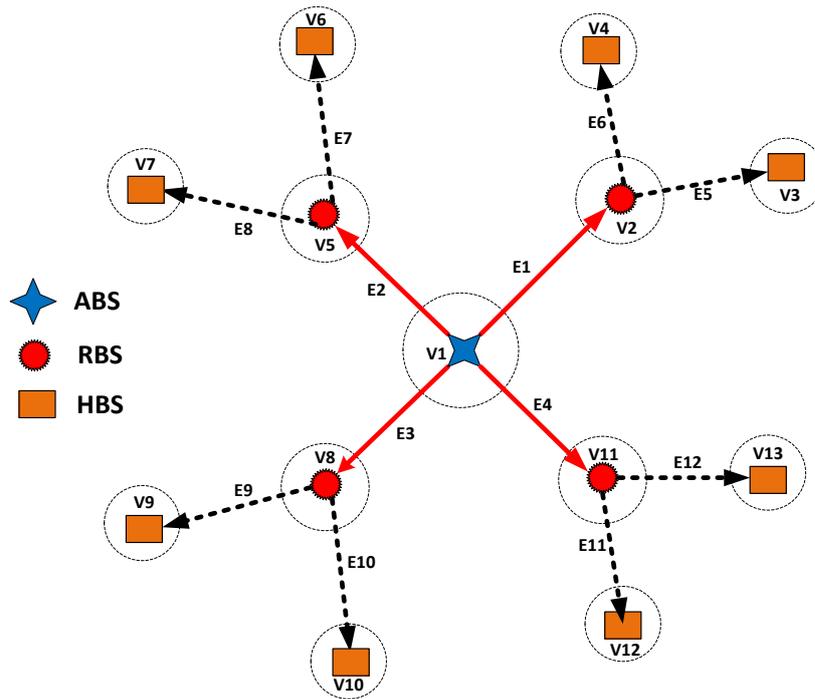
Figure 5.2: Graph illustrating the available path diversity in USCN

$$E_j \in E, \qquad \forall_j \in [1, 12]$$

An ABS is surrounded by four RBSs and each RBS is in direct Line-of-Sight of two HBSs placed on each corner of a block. It is verified that the path diversity graph G in this dual-hop setup is isomorphic and it is true for every ABS in the network. Each ABS has a maximum of 8 possible routes for transmitting its data to any aggregation point. However, the unidirectional arrows indicate the flow of uplink transmission towards the HBS. Six different combinations of routing are investigated in Chapter 4 for achieving better network GoS.

5.4. Energy Model

In this dual-hop backhaul network model, entities like ABS, RBS and HBS are the predominant source of energy consumption. The energy model proposed in [24] is used to estimate the total energy consumed in the network. The energy consumption of a BS is mainly divided into SLEEP, ACTIVE and ON states. It depends on the time spent in each of these three states as well as the number of times a BS makes a transition from SLEEP to the ACTIVE state. The overall BS energy consumption is determined according to the following equation:

$$E_{BS} = \sum_{i=1}^{n_{BS}} \left(T_{Sleep,i} P_{Sleep} + T_{Active,i} P_{Active} + T_{On,i} P_{On} + n_i E_{Wakeup} \right) \left(\frac{1}{1 - \mu_{sl} - \mu_c} \right)$$

(5.2)

where n_{BS} is the total number of BSs in the service area. $T_{Sleep,i} P_{Sleep}$, $T_{Active,i} P_{Active}$ and $T_{On,i} P_{On}$ are the energy consumed by the ith BS in SLEEP, ACTIVE and ON state respectively. E_{Wakeup} is the energy associated with the wake up process while n_i is the number of times the BS switches from SLEEP to ACTIVE state [24]. μ_{sl} is the loss in the power supply and battery backup and μ_c is the ratio of energy spent on cooling system compared with the total energy consumption of the BS. μ_c is only considered for HBS because it is much more powerful and complex than the RBS.

The energy consumed in the ON state alone is further divided into two parts; the energy consumed during the transmitting period $T_{Tx,i} P_{Tx}$ and the energy consumed in receiving mode $T_{Rx,i} P_{Rx}$. It is represented by the following equation:

$$E_{BS} = \sum_{i=1}^{n_{BS}} \left(T_{Tx,i} \frac{P_{Tx}}{\mu_{RF}} + T_{Rx,i} \frac{P_{Rx}}{\mu_{RF}} \right)$$

(5.3)

where μ_{RF} is the power amplifier efficiency. It is worth mentioning that the factor $T_{Tx,i} P_{Tx}$ for the HBS case in Equation 5.3 is equal to zero as only the uplink transmis-

sion is considered in this work. In Table 5.1 the values for the energy model parameters are summarized [24].

Table 5.1: Energy Model Parameters

Parameters	Values	
Power in Active mode	5 W	
Power in Sleep mode	250 mW	(5% of P_{Rx})
Max Transmit / Recieve power	5 W	
RF efficiency	20 %	
Supply loss efficiency	10%	
Cooling efficiency	50%	
Wakeup energy	50 J	

Furthermore, by not incorporating the ABSs in topology management, the access network high QoS is not compromised at all. Also, the BSs are not completely shut down to make sure that they can wake up in sufficient time to handle large variations of users accessing or departing the network.

5.5. Simulation Results

A Monte Carlo simulation is carried out to evaluate the performance of Path Diversity (PD) using the Topology Management (TM) strategy illustrated in Section 5.4. The backhaul path selection process is performed at each iteration. As described in Section 4.3, FA Resource Assignment Scheme is implemented for the dual-hop backhaul architecture which is widely used in the literature. We have not selected anything more sophisticated because it is not the main focus of the thesis.

As mentioned in Section 4.7, the Free Space Path Loss Model is used in this model. The simulated architecture consists of 16 ABSs, 16 RBSs and 4 HBSs distributed across the service area as shown in Figure 4.2. Other important simulation parameters are stated in Table 5.2.

To assess the effectiveness of routing-based topology management scheme and find the trade-off between energy consumption and system QoS, 4 different versions of this

Table 5.2: Simulation Parameters

Parameters	Values
Coverage area	450x450 m
Building size	75x75 m
Street width	15 m
Building height	6 m
Max antenna gain	17 dBi
SINR threshold	1.8 dB
Log-normal shadowing factor	3 dB
Carrier frequency	75 GHz
Antenna gain	38 dB
Number of HBS/RBS/ABS	4/16/16
Frequency Spectrum	500 MHz
Resource blocks	200
Inter-arrival time	Exponential Distribution
Noise floor	-114 dBm/MHz
Mean file size	2 MB
Iteration per offered traffic	200k

scheme are investigated. It is compared with a baseline network having no dynamic topology management capabilities and no path diversity in the network. At very low occupancy levels, at least 4 RBSs and 1 HBS are needed to be in ON state for providing the backhaul connectivity to all 16 ABSs. This is the minimum requirement for the topology management to work accurately otherwise QoS will significantly degrade.

The first metric is the Power Consumption (PC) which is illustrated in Figure 5.3 and Figure 5.4 for all HBSs and RBSs respectively. The PC for the HBS with no topology management is almost half the PC for RBS because the HBSs are only working in receiving mode. As shown in Figure 5.3, the topology management scheme with the path diversity reduces the PC by a factor of four for traffic loads less than 4 Gbps and is effective up to 6 Gbps.

With no path diversity, the same scheme reduces the power consumption by half and is only effective up to its saturation point at 4.2 Gbps. However, the rapid rise in PC for high traffic scenarios is mainly due to bottleneck on the second hop. At this point, the blocking starts showing its influence on the power consumption. Therefore, it is this path diversity which is utilized by the network topology strategy for power reduction.
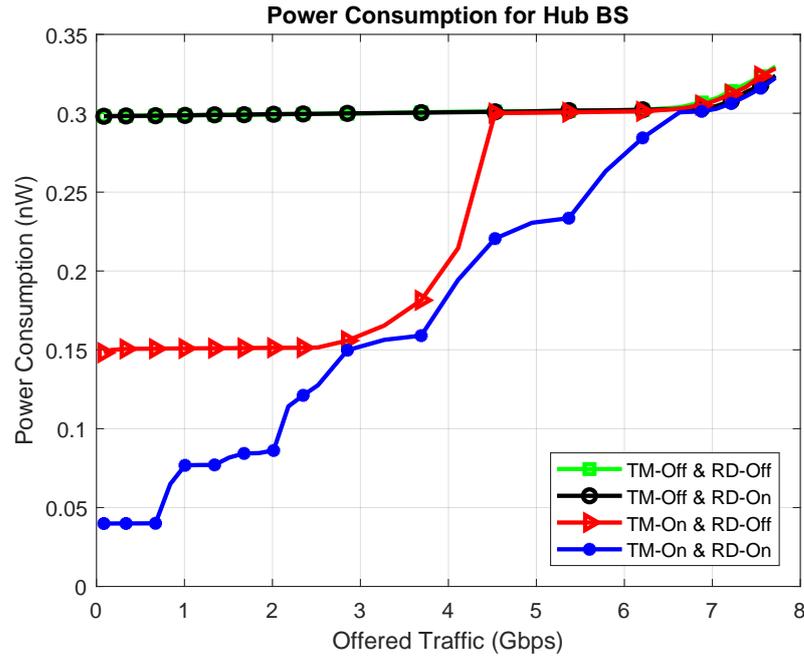
Figure 5.3: Impact of Topology Management on HBS power consumption

Figure 5.5 demonstrates the whole network Energy Consumption Rating (ECR) [58]. Unlike PC, it is measured on a logarithmic scale in (μJ per bit) [86]. These ECRs include the energy consumption of all ABSs which are not equipped with the proposed dynamic network capabilities. It validates earlier results that notable QoS and energy savings can be achieved with a scheme equipped with path diversity and topology management, especially in low to medium traffic demands. It flexibly adjusts the topology according to the traffic demand having an average of 75% and 35% energy reductions at low (0.5-2.5 Gbps) and medium (2.8-5.8 Gbps) traffic loads respectively. Beyond 7 Gbps, the energy savings reach zero as the system is most likely to operate at its full capacity. It is worth mentioning that both these rating parameters are based on the successful number of bits sent.

Blocking Probability and System Throughput are the preferred QoS parameters for measuring the performance of this dual-hop backhaul network. The topology management scheme is analysed for energy reduction with blocking probability less than 5% [24]. From Figure 5.6, it is evident that significant improvement in GoS can be achieved with path diversity because multiple path alternatives reduce the bottleneck from second hop. Schemes with the path diversity perform better in term of QoS than
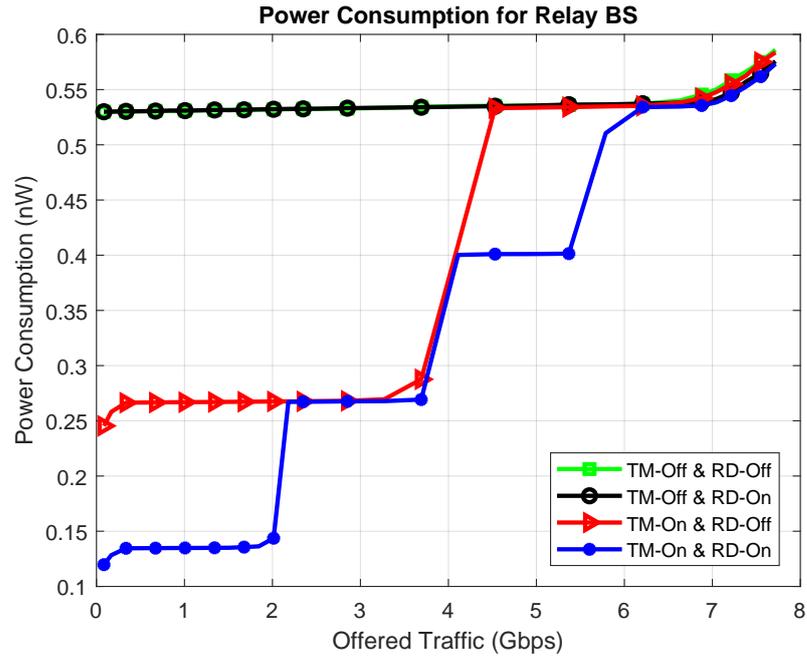
Figure 5.4: Impact of TM on RBS PC while increasing the offered traffic
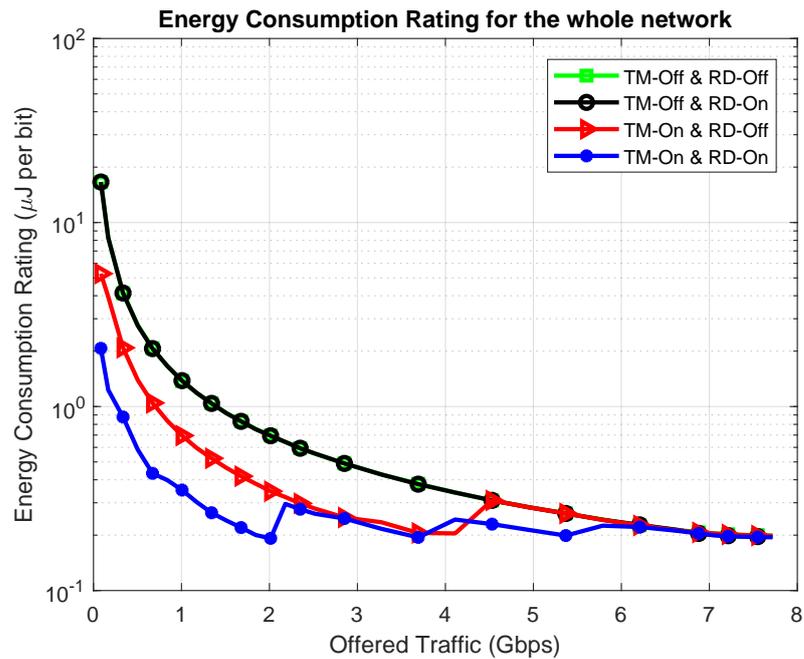
Figure 5.5: Impact of TM on Energy Consumption while increasing the offered traffic

the scheme with no routing capabilities. At high offered traffic, the bottleneck is diminished by shifting the newly arriving user to the least loaded section of the network. Also, the slight discrepancy in the absence of routing diversity is due to the traffic burstiness in simulation.

Blocking Probability (Pb) in the network
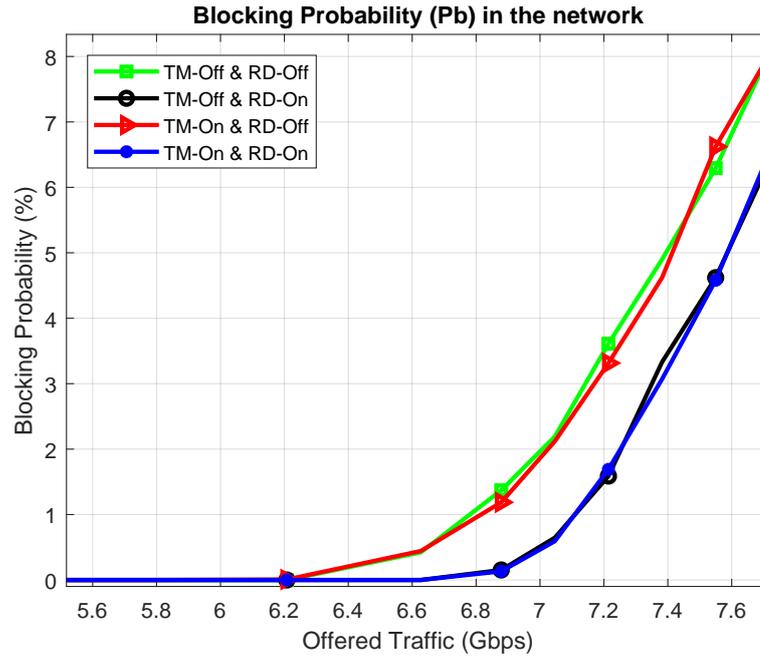
Figure 5.6: No degradation in QoS (Blocking) with TM scheme in place

At 7 Gbps, the system crosses saturation point and the throughput is restricted from further increasing due to channel blockage as illustrated in Figure 5.7. However, at low traffic loads, when the power savings are dramatically increased by a factor of four, the QoS parameters are barely showing any deterioration. Thus, confirming the significance of routing diversity in such ultra-dense environment which can be utilized for reducing the power expense. This path diversity enables all ABSs in the network to obtain connectivity to the far most HBS whilst the topology management exploits it to lessen the power expense by switching as many BSs to dormant mode.

Figure 5.8 illustrates the effect of topology management on Resource Block Usage. It enables the network to use the entire spectrum more efficiently and evenly. The low range RBs are having high $RB_\%$, thus validating the correctness of resource assignment scheme used. However, in the absence of topology management, high range resources $(RB_i \geq 160)$ are having $RB_\%$ almost equal to zero. On the one hand, path diversity directly improves network QoS performance but on the other hand, it is exploited by the topology management for energy reduction.
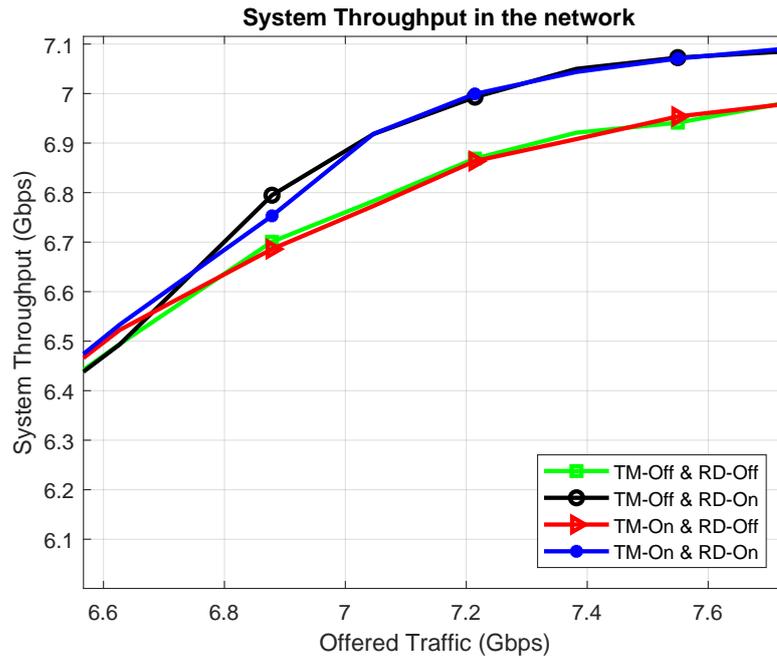
Figure 5.7: Impact of increasing Offered Traffic on System Throughput
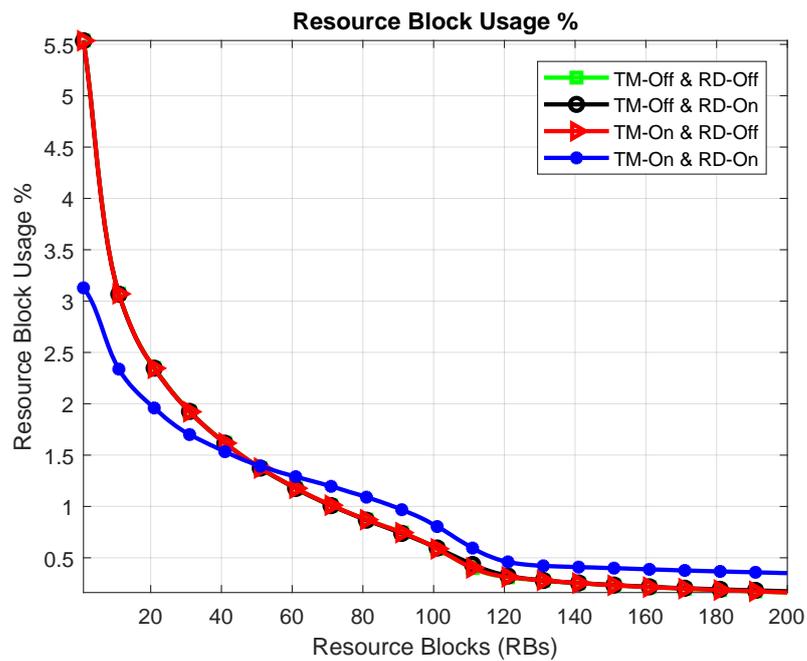
Figure 5.8: Impact of TM on Resource Blocks

5.6. Conclusions

This chapter presented an efficient energy based topology management strategy for future generation dual-hop backhaul network which enhanced network performance in

terms of power savings while still preserving exactly the same high QoS demand. Thus, validating the second key objective of this thesis which is to enhance the overall energy efficiency. The topology management strategy achieves best for data traffic levels less than 7 Gbps. It efficiently exploits the existing path diversity between the BSs in such ultra-dense scenarios in order to lessen the energy consumption from 35% to 75% in low to medium traffic loads. It is illustrated that the QoS is significantly improved by increasing the number of path alternatives which to a greater extent mitigates the bottleneck. It is also shown that if a BS is not turned back on in time, the system suffers from severe blocking. Therefore, adequate time is allocated for the initial transition stage.

Chapter 6. Grade of Service Enhancement in a mm-Wave Multi-hop, Multi-tier Heterogeneous Network

Contents

6.1. Introduction

I N this chapter, the dual-hop backhaul network architecture introduced in Chapter 4 is combined with a macro base station to make a heterogeneous network (Het-Net). It is applied in order to facilitate and encourage consumers to offload their data traffic from macro BSs to the alternative single/multi-hop small-cell networks and vice versa to achieve complete fairness. This GoS fairness in the overall network is the third key objective of this thesis.

In general, a HetNet consists of multiple tiers of networks of different cell sizes/footprints and/or of multiple radio access technologies. A macro base station overlaying a multi-hop, ultra-small cell network is a good example of a multi-tier heterogeneous network [30]. From the perspective of a consumer, HetNets need to provide ubiquitous coverage, secure and high data rates, high throughput capacity, always-on, and always-connected-to-best-network user experience [22].

The design of a HetNet system should also consider the probability that the network cannot establish a connection, i.e. the probability that no resource is available to users. However, various key factors including different obstacles, signal degradation and user equipment restrictions will reduce the availability of some portion of the network to some of the users. In such situations, those limited users will face an inadequate network environment.

A composite state with different levels of accessibility is expected to happen, where users with limited network choices and users with a more comprehensive set of connectivity options coexist in the same environment. This is primarily because different users often have dissimilar terrestrial locations, elevation/azimuth angles and/or antenna equipment choices [18]. When different types of users coexist in the same coverage area sharing a mutual resource pool, the resource utilization level of the two groups will affect each other. The presence of different types of users can worsen the performance of the disparate users and make the management of frequency spectrum challenging. Therefore, a mechanism is required for user admission control to guarantee system fairness.

The rest of this chapter is organised as follows: in Section 6.2, the next generation heterogeneous network architecture is introduced. The concept of users in a coexistence scenario having different levels of network accessibility is discussed in Section 6.3. Followed by the restriction mechanism technique in Section 6.4. Two different case studies are presented in 6.5 for analytical comparison. In Section 6.6, large scale simulation results are illustrated. Finally, conclusions are given in Section 6.7.

6.2. Multi-hop Multi-tier Heterogeneous 5G Network

A mm-wave, multi-hop, multi-tier heterogeneous architecture is proposed as an ultra-dense network solution for future 5G cellular networks as shown in Figure 6.1. This HetNet is composed of a multi-hop, Ultra-Small Cell Network with an overlay of a Macro Base Station (MBS). The MBS provides conventional single-hop access connectivity to users on the ground in coexistence with the USCN in the same coverage area.
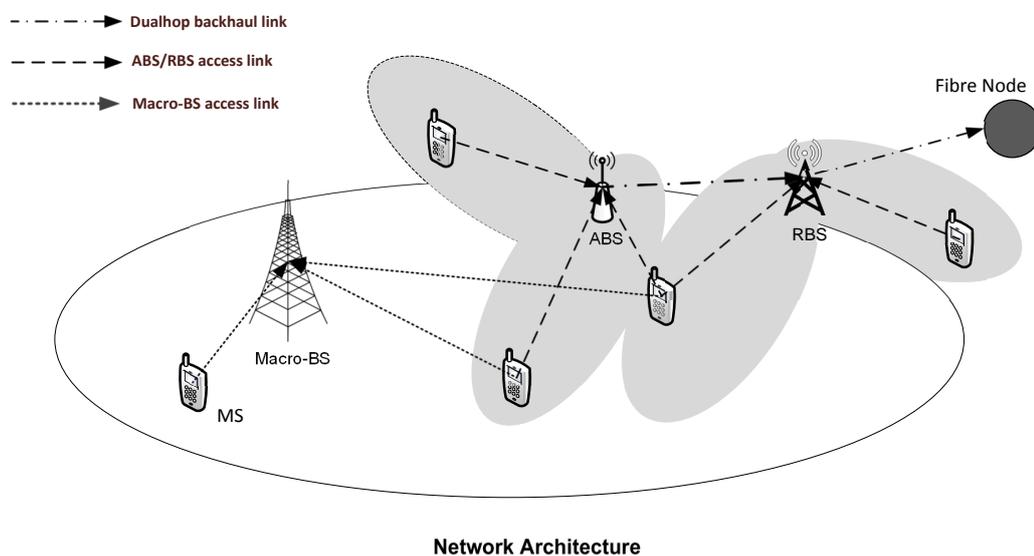
Figure 6.1: Multi-tier heterogeneous network for ultra-dense scenarios

Most research work in HetNet domain is primarily related to single-hop femto-cells, pico-cells and relay-cells overlaid on the edges of a macro-cell for cell range expansion [29]. In this work, it is extended one step further to multi-hop, ultra-small cells which are spread throughout the macrocell coverage area as shown in Figure 6.1. These multi-hop links consist of single-hop access links and dual-hop backhaul links which give the flexibility to respond to sudden changes in the network as well as to minimize the energy consumption of the entire network. These short range multi-hop links enable data to be directed from the users towards different aggregation points when some portion of the network is unavailable at low occupancy levels [3].

The multi-hop network can be further classified into a low frequency access network and a mm-wave single/dual-hop backhaul network. The main objective is to deliver high throughput capacity density with the least latency in the service area with minimum cost. The majority of key elements of this network architecture are explained in detail in Section 2.3.

6.2.1. Single-hop Access Network

Chapter 4 and 5 are only related to the backhaul section of the USCN. In this chapter, an access network is introduced along with the backhaul portion to investigate a complete end-to-end link scenario. The MSs are facilitated by this single-hop access network. As mentioned earlier, only the outdoor MSs are considered, since in 5G it is expected that the MSs that are located indoors will be served by indoor infrastucture [94]. The SINR for ABS/RBS n (signal transmitted from MS k using RB w) is:

$$\gamma_{n,w}^k = \frac{P_{A,k}^M \, g_w^{A,k,n}}{\displaystyle\sum_{i=1,i\neq k}^{P} P_{A,i}^M \, g_w^{A,i,n} + \sigma^2} \tag{6.1}$$

where $P_{A,k}^M$ and $g_w^{A,k,n}$ is the MS transmit power and gain of access link respectively from MS k to ABS/RBS n. The factor $\displaystyle\sum_{i=1,i\neq k}^{P} P_{A,i}^M \, g_w^{A,i,n}$ is the signal interference from all other active MSs ($i \neq k$) in the network to ABS/RBS n using the same resource RB w. For the access network, with the assumption of an omnidirectional antenna at the MS end, the link gain is obtained by:

$$g_w^j = \frac{G_j(\theta_j)}{PL(d_j)} \tag{6.2}$$

The channel propagation is modelled using the WINNER II B1 propagation model [5] for the low frequency ultra-small cell access network as both the BSs and MSs are deployed outdoors. The path loss for access scenarios is explained in Section 3.6.2.

During the uplink transmission, the effective signal strength at the receiver is obtained by accounting for the gains of MS and BS antennas, shadowing, path loss on the channel and interference from other users using the same RBs.

6.2.2. Dual-hop Backhaul Network

This dual-hop backhaul network of USCN is exactly the same as mentioned in earlier chapters. The links are mainly composed of single-hop and dual-hop backhaul connections. The data from/to multiple users is routed with low latency either via ABS/RBS to the Fibre Node through single/dual-hop mm-wave backhaul links. All the details related to this portion of the wireless network are thoroughly mentioned in Chapter 4. The available path diversity in this dual-hop backhual network is illustrated in Figure 5.2.

The path loss for the dual-hop section is explained in Section 3.6.1. The use of short range highly directional beams with large angular separation between the nodes in such densely deployed scenario implies that both interference and multipath should be negligible, and the backhaul links are largely unaffected by its environment. The street canyons with high buildings also add benefits, such as interference reduction and higher frequency reuse. Although, extensive research work is carried out recently on phased array mm-wave approaches which exploit reflections but it is not considered in this thesis.

Meanwhile, the MBS provides an overlay to all the ultra-small cells in the same coverage area as shown in Figure 6.1. It is the second layer of this multi-tier heterogeneous network which is equipped with a massive MIMO system. Due to this system, it is possible to have an equal number of resource blocks for both the MBS access network and the Ultra-Small Cells network. The radiation pattern and propagation model for the MBS are given in Section 3.7.

6.3. Coexistence Scenario in Multi-hop, Multi-tier Heterogeneous Networks

A coexistence scenario is analysed for a different domain in [18], which dealt with the users having restricted antenna directionality. In this chapter, a range of elements including various obstacles, signal attenuation and MS terminal restrictions will lessen the accessibility of one tier of a network to some of its users. Due to this inadequate network environment, we extend the applicability of restriction to different scenarios in a mm-wave, multi-hop, multi-tier HetNet domain. This restriction approach is developed further and analysed in a multi-tier scenario for users having distinct levels of network accessibility.

A coexistence scenario is likely to happen where users with a full access choice and users with a restricted choice coexist in the same coverage area as shown in Figure 6.2. This is due to the fact that different MSs frequently have different geographical locations, elevation angles and antenna equipment choices [18]. The existence of different types of users can actually reduce the performance of the overall system and make the spectrum utilization process significantly inefficient. The two user types which are investigated in this study are given below.

- **Privileged Users**

  The privileged group of users can potentially access the MBS as well as the USCN as shown in Figure 6.2. This group of users are not suffering from radio link outage and can access both the tiers of the network, therefore they are considered as privileged. In realistic scenarios, there are no obstacles to block this set of users from connecting to both the tiers due to their suitable geographical location and smart antenna systems. They have access to the resources available on both the tiers as shown in the restriction mechanism Figure 6.3.
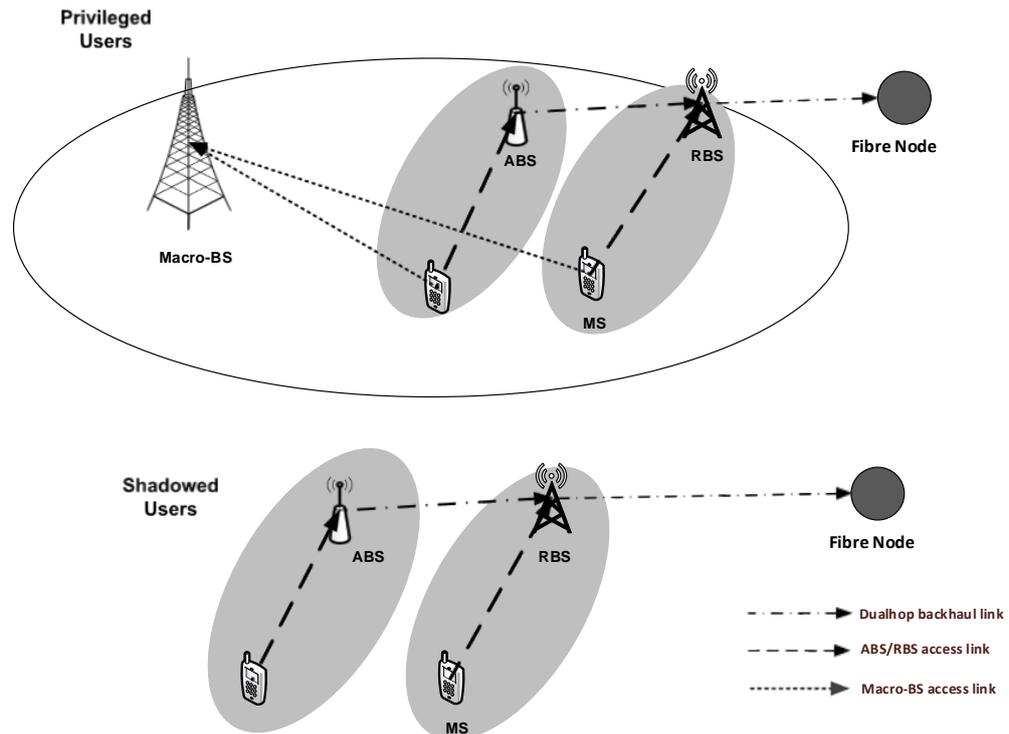
Figure 6.2: A coexistence scenario with different types of user groups

- **Shadowed Users**

Shadowed users represent that set of users which are suffering from radio link outage, caused by earthly obstacles, significant signal degradation or instead they represent users equipped with a simple fixed antenna system. Therefore, this user group has limited resource options as shown in Figure 6.3.

For GoS control in such scenarios, the MBS will require knowledge of load levels on single/dual-hop backhaul links. This can be achieved by a Fibre Node communicating this information to the MBS via a dedicated control interface. The Fibre Node and MBS then decide the access restrictions depending on whether the incoming user is a member of privileged or shadowed user group. In this work, through the restriction process, the GoS can be controlled for both types of users.

6.4. Restriction Mechanism

A restriction mechanism is investigated in [18], where the antenna directionality is restricted to some users. In this thesis, the restriction technique is applied to mm-wave, multi-hop, multi-tier HetNet scenarios where the users mainly suffer from radio link outage, caused by earthly obstacles, significant signal degradation or instead they represent users equipped with a simple fixed antenna systems. The main objective is to develop an access control approach for heterogeneous networks that will help to balance the traffic loads between different tiers having disparate groups of users.

In order to eliminate the performance disparity, the restriction mechanism stops the privileged users from gaining access to the USCN under specific circumstances. Meanwhile, the MBS is the only alternate option that they have to get a connection. In other words, we temporarily change the privileged users into a new type of restricted users by limiting their choice availability. Also, the shadowed users only have access to the USCN resources as shown in Figure 6.3. In a controlled and flexible way, this approach prevents privileged users from accessing the USCN when their load levels is above a certain threshold limit as shown in Figure 6.3. The most basic approach is the constant restriction function which is given below:

$$ r(j) = C_c \qquad 0 < C_c < 1 \tag{6.3} $$

where C_c is the coefficient of restriction function. It reserves the remainder of the available resource blocks in the ultra-small cell network for shadowed users which have more limited choice flexibility in this wireless network. However, once the load level of the USCN falls below the threshold set by Equation 6.3, the temporary restriction which is applied earlier on privileged users is removed. It is this restriction mechanism that maximizes the system performance by limiting the choice availability for its users. Some more advanced restriction functions are discussed in [18] which is beyond the focus of this research work.
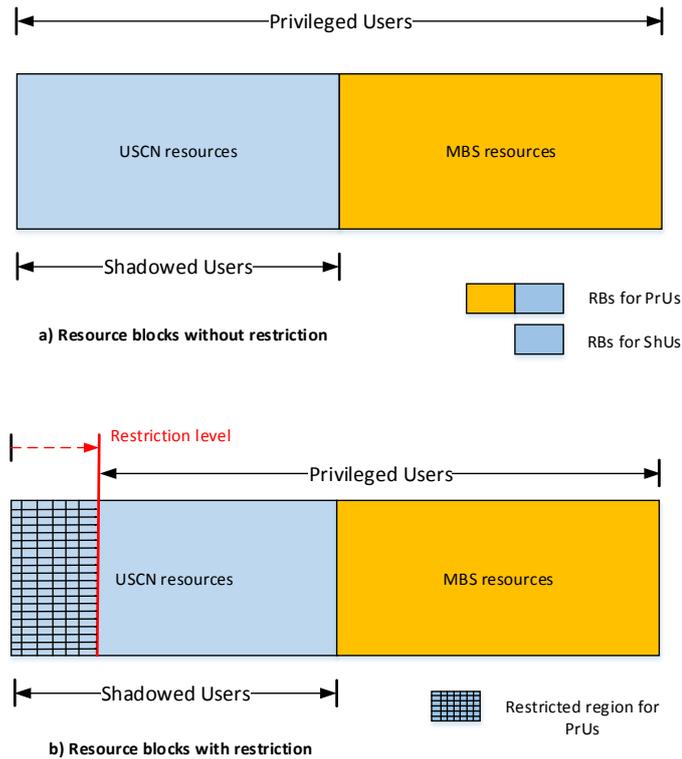
Figure 6.3: Restriction mechanism used to compensate for the inferior GoS performance of shadowed users

6.5. Two Dimensional Markov Analysis

In this section, the network performance of the coexistence scenario is investigated. Considering individual users accessing the system, poisson arrival and departure processes for both Privileged and Shadowed users are assumed. Figure 6.4 depicts a state transition rate diagram to illustrate the behaviour of n-resource blocks in the two tier heterogeneous network.

Each node in the diagram denotes a state. The first digit in the node represents the number of resource blocks occupied on the MBS while the second digit in the node stands for number of resources occupied on the last hop of the USCN. This final hop is responsible for the bottleneck in the USCN due to data accumulation. Also, it is assumed for this analytical study that there is no retransmission in case of blocking and no signal interference. The arriving process for both types of users is constant and

exponential distributed. Notations used in Figure 6.4 are explained below:
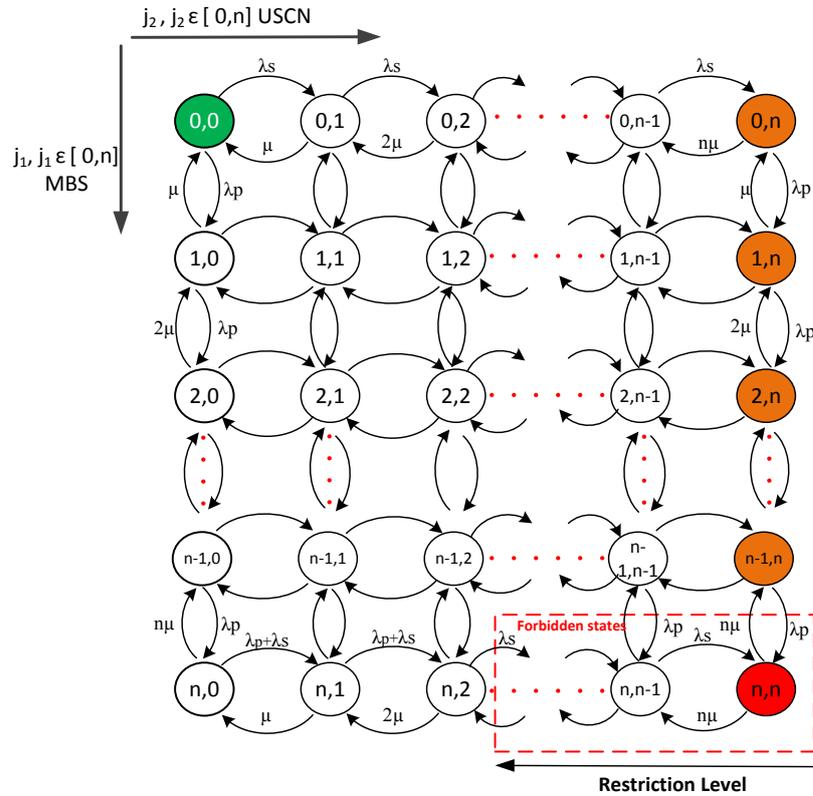
Figure 6.4: State-Transition Markov Model with Restriction Mechanism

- λ_S Arrival rate of shadowed users to access the network.

- λ_P Arrival rate of privileged users to access the network.

- n Number of resource blocks on each tier of a network.

- μ Departure rate per channel is constant and the service time $(1/\mu)$ is exponentially distributed.

- **Case Study # 1**

This case study describes the best case scenario where both group of users are detached as much as possible. They only interact with each other once the MBS is fully occupied as shown by the bottom edge in Figure 6.4. The resource allocation mechanism is a birth-death process [82, 83]. Transitions in a vertical direction represent the arrival and departure process on the MBS, while transitions in the horizontal direction

represent the arrival and departure process on the USCN. In the vertical direction, the arrival rate on the MBS is equal to λ_P.

In the horizontal direction, the arrival rate on the USCN is λ_S, when resources on the MBS are not fully occupied ($j1 < n$). When resources on the MBS are fully occupied, i.e. $j1 = n$, the arrival rate in the horizontal direction becomes $\lambda_P + \lambda_S$. This is because when privileged users initially arriving at the MBS cannot find any resources available on the MBS, only then will they access the USCN in search of free resource blocks. The total system arrival rate is split into equal halves for the two user groups ($i.e. \lambda_P + \lambda_S = \lambda_T$). The departure rate in any direction is equal to $k\mu$, where k is the number of busy resource blocks of that state.

The restriction mechanism explained in Section 6.4 is modelled in such a way to improve the inferior GoS of shadowed users. In Figure 6.4, it is illustrated that due to restriction, those states are restricted for the privileged users along the horizontal direction ($j1 = n$). However, these states are accessible in a vertical direction by the privileged users. The restriction level extends from 0% all the way up to 100%. It is important to control and balance the resource distribution to achieve a fair allocation pattern in the coexistence setup.

The restriction process also equalizes the performance of both user groups. In a controlled and flexible way, it blocks some privileged users to reserve more bandwidth for shadowed users which have the more limited choice flexibility. It is this compensation effect that allows the network to achieve a balanced blocking probability as shown in Figure 6.5. It is worth mentioning that the GoS performance of each user group is more important than any other criterion like the collective GoS of both the user groups. In this thesis, the poor GoS of the shadowed users is due to their restricted access. The aim is to improve the GoS performance of shadowed users at minimum expense in terms of overall system capacity and complexity.

- **Equilibrium Analysis**

Law of Conservation of Flow at statistical equilibrium states that the rate of flow into state (j_1, j_2) is equal to the rate of flow out of state (j_1, j_2). At any random point of time, the system can be in any state (j_1, j_2) with a state probability $P(j_1, j_2)$ [82]. The states in Figure 6.4 can be split into four components in the corners, six components on the edge and one component in the centre with total of eleven different equilibrium equation formats respectively. However, it is impossible to have a negative number of occupied resources in system, thus the condition $P(-1, j_2) = P(j_1, -1) = 0$, is applied to simplify the equilibrium expression. In this way, the number of equation formats can be reduced to just six, which are presented as follows:

For state (j_1, j_2), $0 \leq j_1, j_2 < n$, we have

$$(\lambda_P + \lambda_S + j_1\mu + j_2\mu).P(j_1, j_2) = (j_1+1)\mu.P(j_1+1, j_2) + (j_2+1)\mu.P(j_1, j_2+1)$$
$$+ \lambda_P.P(j_1-1, j_2) + \lambda_S.P(j_1, j_2-1) \quad (6.4)$$

For state (j_1, n), $0 \leq j_1 < n$,

$$(\lambda_P + j_1\mu + n\mu).P(j_1, n) = (j_1+1)\mu.P(j_1+1, n) + \lambda_P.P(j_1-1, n) + \lambda_S.P(j_1, n-1)$$
$$(6.5)$$

For state (n, j_2), $0 \leq j_2 < R_L$,

$$(\lambda_P + \lambda_S + n\mu + j_2\mu).P(n, j_2) = (j_2 + 1)\mu.P(n, j_2 + 1) + \lambda_P.P(n - 1, j_2)$$
$$+ (\lambda_P + \lambda_S).P(n, j_2 - 1) \quad (6.6)$$

For state (n, R_L),

$$(\lambda_S + n\mu + R_L\mu).P(n, R_L) = (R_L + 1)\mu.P(n, R_L + 1) + \lambda_P.P(n - 1, R_L)$$
$$+ (\lambda_P + \lambda_S).P(n, R_L - 1) \quad (6.7)$$

For state (n, j_2), $R_L < j_2 < n$,

$$(\lambda_S + n\mu + j_2\mu).P(n, j_2) = (j_2+1)\mu.P(n, j_2+1) + \lambda_P.P(n-1, j_2) + \lambda_S.P(n, j_2-1)$$

$$(6.8)$$

Finally, for state (n, n),

$$(n\mu + n\mu).P(n, n) = \lambda_P.P(n-1, n) + \lambda_S.P(n, n-1) \qquad (6.9)$$

In the $(n+1)^2$ equations above, one of them is redundant, which means it can be derived from other $(n+1)^2 - 1$ equations. As the system always will be in a state, the state probabilities must also satisfy the normalization equation [82].

$$\sum_{j_1=0}^{n} \sum_{j_2=0}^{n} .P(j_1, j_2) = 1 \qquad (6.10)$$

The $(n+1)^2$ equations and the normalization equation can be expressed in a matrix format

$$\mathbf{AP} = \mathbf{B} \qquad (6.11)$$

where $\mathbf{A}$ is the $(n+1)^2$ x $(n+1)^2$ coefficient matrix, $\mathbf{P}$ is the $(n+1)^2$ x 1 state probability vector and $\mathbf{B}$ is the $(n+1)^2$ x 1 constant vector. By solving the matrix equation, we can obtain the state probability vector $\mathbf{P}$ and effectively all the $(n+1)^2$ state probabilities $P(j_1, j_2), 0 \leq j_1, j_2 \leq n)$

$$\mathbf{P} = \mathbf{A}^{-1}\mathbf{B} \qquad (6.12)$$

It is too complex to derive an expression of $P(j_1, j_2)$ in a closed form for the restricted scenario presented in this chapter. Therefore, the Gauss-Seidel Method explained in Section 4.6 is used to solve this system of linear equations.

- **Blocking Probability**

The state probabilities $P(j_1, j_2), \ (0 \leq j_1, j_2 \leq n)$ incorporating the restriction function are calculated. The blocking probability of Shadowed User (PB_{SH}) is equivalent to the sum of state probabilities on the right edge in Figure 6.4.

$$PB_{SH} = \sum_{j_1=0}^{n} .P(j_1, n) \tag{6.13}$$

The blocking probability of Privileged Users (PB_{PR}) should take into account not only the state probability $P(n, n)$ but also the blocking probability caused by the restriction mechanism.

$$PB_{PR} = \sum_{j_2=R_L+1}^{n} .P(n, j_2) \tag{6.14}$$

The restriction function also provides a certain degree of controllability over the whole network performance. For example, the probability of both the user groups can be equalized, i.e. For a complete fair network.

$$PB_{PR} = PB_{SH} \tag{6.15}$$

- **Analysis of Analytical and Simulation Results**

In this section the analytical model and the Monte-Carlo simulation results of the coexistence scenario for 24 resource blocks $(n = 24)$ are compared. The arrival rates for both the user groups are identical $(\lambda_P = \lambda_S = \lambda_T/2)$. 0% restriction means that the privileged users have complete access to the USCN as well as MBS while at 100% restriction level, the privileged users can only access the MBS. At 100% restriction, the whole USCN resources are reserved for the shadowed users.

As expected, with an increase of restriction level, the blocking probability for the shadowed users decreases while the blocking for the privileged users increases. From
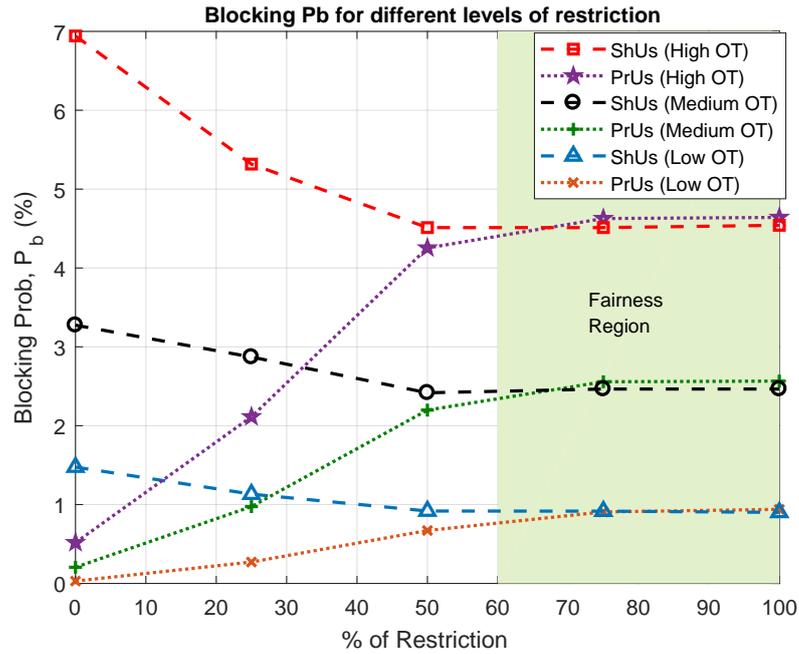
Figure 6.5: Performance evaluation of restriction for certain levels of offered traffic

60% restriction onwards, the improvement in performance is negligible due to system saturation. The restriction mechanism is able to equalize the performance of both the user groups as indicated by the fairness region in Figure 6.5. These results also indicate the inherent discrimination in this system when both these users types coexist. In the case of no restriction, the shadowed users have much poorer performance compared to privileged users due to their confined number of choices.

Figure 6.6 and Figure 6.7 illustrates the effect of increasing offered traffic on the blocking probability performance of both the user groups. At 180 Mbps, the PB_{SH} is reduced from 6.9% to 5.4% and 4.5% with restriction levels of 25% and 75% respectively. From these results it is very clear that the restriction mechanism is mostly effective at high traffic loads. Therefore, it is most beneficial to postpone the restriction process until high offered traffic.

The restriction is intended to provide free resource blocks for the shadowed users primarily in times of high traffic. For privileged users, the same restriction levels increases PB_{PR} from 0.4% to 2% and 4.5%. Overall, it indicates that the analytical model is a good representation of the system.
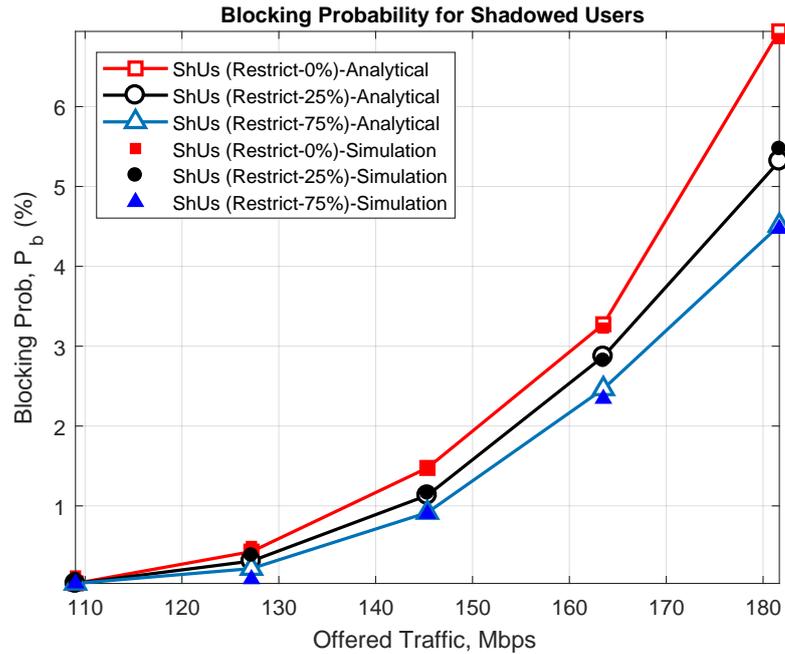
Figure 6.6: Impact of increasing offered traffic on shadowed users
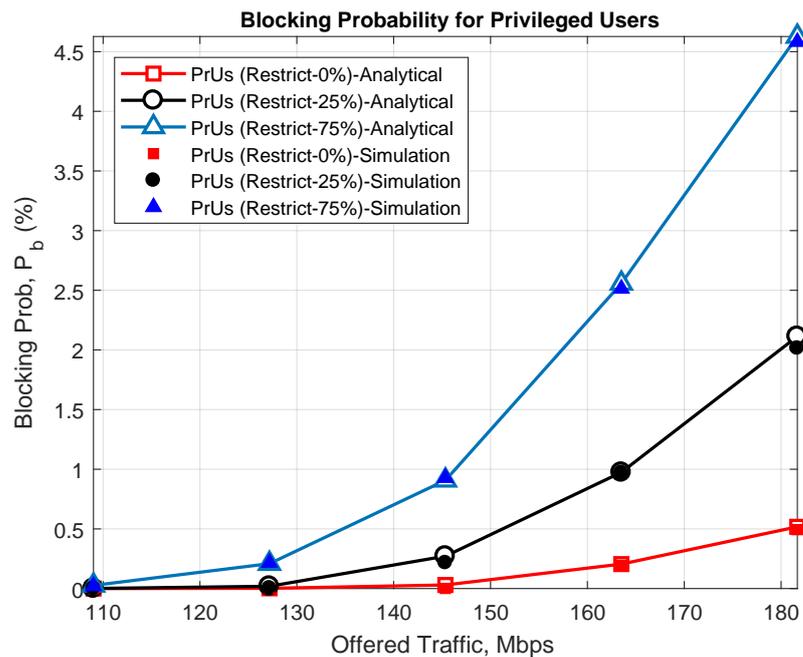
Figure 6.7: Privileged users blocking for different levels of Offered Traffic

- **Case Study # 2**

In the previous study, the case is analysed where the privileged users only access the USCN when there are no resources available on the MBS. However, this case study is more general and flexible where the performance of the restriction mechanism is

analysed when the privileged users are allowed to have equal access to both types of networks from the very beginning.

This approach is likely to be less advantageous due to greater interaction between both types of users. The aim is to understand the degradation in system performance which results from this higher degree of interaction.

Two more notations for this case study are:

- λ_{P1} Arrival rate of of privileged users to access MBS.

- λ_{P2} Arrival rate of privileged users to access USCN.

The total arrival rate for such scenario is written as, $\lambda_T = \lambda_{P1} + \lambda_{P2} + \lambda_S$ while for privileged users, the two arrival rates ($\lambda_{P1} = \lambda_{P2}$) are equal to $\lambda_T/4$. The restriction is applied in a similar way as explained in Section 6.4 but the rate of transitions are different along both vertical and horizontal directions as indicated in Figure 6.8.
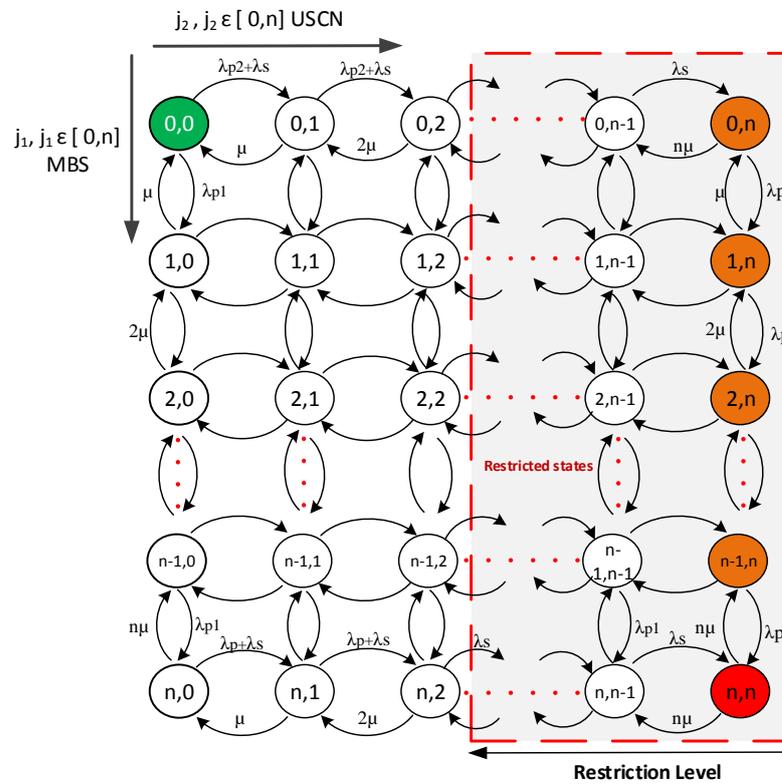
Figure 6.8: State-Transition Diagram in case of allowing a % of PrUs to access USCN

Initially, when the load level on the USCN is below the restriction limit, the horizontal transition is the summation of λ_{P2} and λ_S but once that restriction limit is crossed, the horizontal transition is only caused by the λ_S. These are restricted states in the horizontal direction for the privileged users but they are accessible vertically when those users want to access the MBS as indicated by vertical transition λ_{P1}. However, the remainder of the forward and backward transitions are the same as explained earlier in Case Study # 1.

Once, the resources on the MBS are fully occupied i.e. $(j_1 = n)$, the states which are restricted by the restriction process become forbidden for the privileged users as shown by the bottom edge in Figure 6.8. With such a configuration, there are not enough resources available for the shadowed users on the USCN due to severe competition between both the user groups and as a result the performance is severely degraded.

The set of equilibrium equations for this case study are given below, when the privileged users are allowed to access the USCN from the very start.

For state (j_1, j_2), $0 \le j_1 < n, 0 \le j_2 < R_L$, we have

$$(\lambda_{P1}+\lambda_{P2}+\lambda_S+j_1\mu+j_2\mu).P(j_1, j_2) = (j_1+1)\mu.P(j_1+1, j_2)+(j_2+1)\mu.P(j_1, j_2+1)$$
$$+ \lambda_{P1}.P(j_1 - 1, j_2) + (\lambda_{P2} + \lambda_S).P(j_1, j_2 - 1)$$
$$[P(-1, j_2) = P(j_1, -1) = 0]$$

$$(6.16)$$

For state (j_1, R_L), $0 \le j_1 < n$,

$$(\lambda_P+\lambda_S+j_1\mu+R_L\mu).P(j_1, R_L) = (j_1+1)\mu.P(j_1+1, R_L)+(R_L+1)\mu.P(j_1, R_L+1)$$
$$+ \lambda_P.P(j_1 - 1, R_L) + (\lambda_{P2} + \lambda_S).P(j_1, R_L - 1)$$
$$[P(-1, R_L) = 0]$$

$$(6.17)$$

For state (j_1, j_2), $0 \le j_1 < n$, $R_L < j_2 < n$,

$$(\lambda_P + \lambda_S + j_1\mu + j_2\mu).P(j_1, j_2) = (j_1+1)\mu.P(j_1+1, j_2) + (j_2+1)\mu.P(j_1, j_2+1)$$

$$+ \lambda_P.P(j_1 - 1, j_2) + \lambda_S.P(j_1, j_2 - 1)$$

$[P(-1, j_2) = P(j_1, -1) = 0]$

$$(6.18)$$

For state (j_1, n), $0 \le j_1 < n$,

$$(\lambda_P + j_1\mu + n\mu).P(j_1, n) = (j_1+1)\mu.P(j_1+1, n) + \lambda_P.P(j_1-1, n) + \lambda_S.P(j_1, n-1)$$

$[P(-1, n) = 0]$

$$(6.19)$$

For state (n, j_2), $0 \le j_2 < R_L$,

$$(\lambda_P + \lambda_S + n\mu + j_2\mu).P(n, j_2) = (j_2 + 1)\mu.P(n, j_2 + 1) + \lambda_{P1}.P(n - 1, j_2)$$

$$+ (\lambda_P + \lambda_S).P(n, j_2 - 1)$$

$[P(n, -1) = 0]$

$$(6.20)$$

For state (n, R_L),

$$(\lambda_S + n\mu + R_L\mu).P(n, R_L) = (R_L + 1)\mu.P(n, R_L + 1) + \lambda_P.P(n - 1, R_L)$$

$$+ (\lambda_P + \lambda_S).P(n, R_L - 1) \quad (6.21)$$

For state (n, j_2), $R_L < j_2 < n$,

$$(\lambda_S + n\mu + j_2\mu).P(n, j_2) = (j_2+1)\mu.P(n, j_2+1) + \lambda_P.P(n-1, j_2) + \lambda_S.P(n, j_2-1)$$

$$(6.22)$$

Finally, for state (n, n),

$$(n\mu + n\mu).P(n, n) = \lambda_P.P(n-1, n) + \lambda_S.P(n, n-1)$$
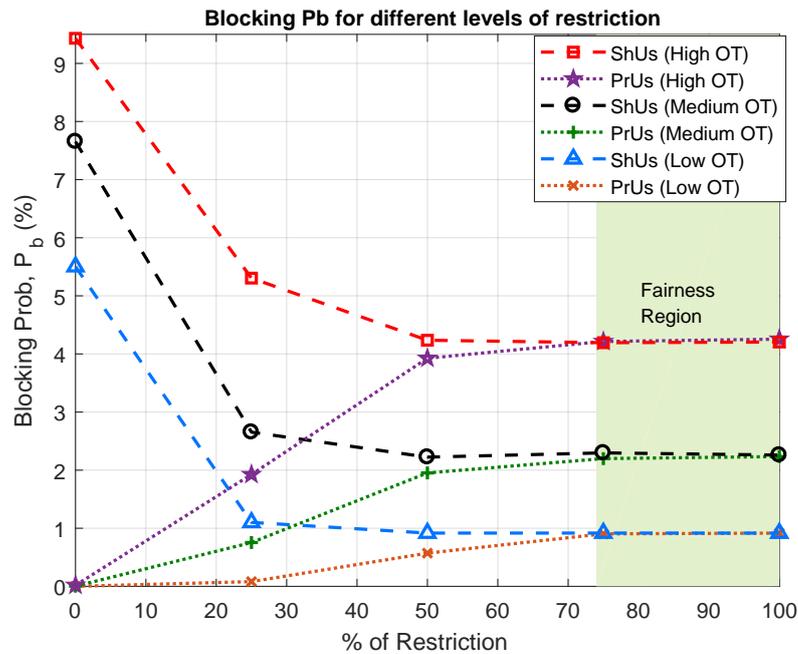
Figure 6.9: Performance evaluation of restriction for certain levels of offered traffic

In this case, the resources available for the privileged users are significantly under utilized. Figure 6.9 clearly indicates that PB_{SH} is very poor even at low traffic loads but with the application of restriction, the performance of these users is enhanced. Without restriction, PB_{PR} is 0% due to the excessive availability of resources to privileged users. Also, above 60% restriction, the improvement in performance is negligible due to system saturation. From Figure 6.9 to 6.11, it is clear that this kind of access for the privileged users makes the entire system underperform and inappropriate.

Figure 6.10 also illustrates that PB_{SH} is reduced from 9% to 4.2% with the restriction process. On the contrary, PB_{PR} is increased from 0% to 3.5% with 50% restriction. Again, it equalizes the blocking probability performance of both the user groups but the fairness region in this scenario is reduced. It is this compensation effect that

allows the system to achieve a balanced GoS.
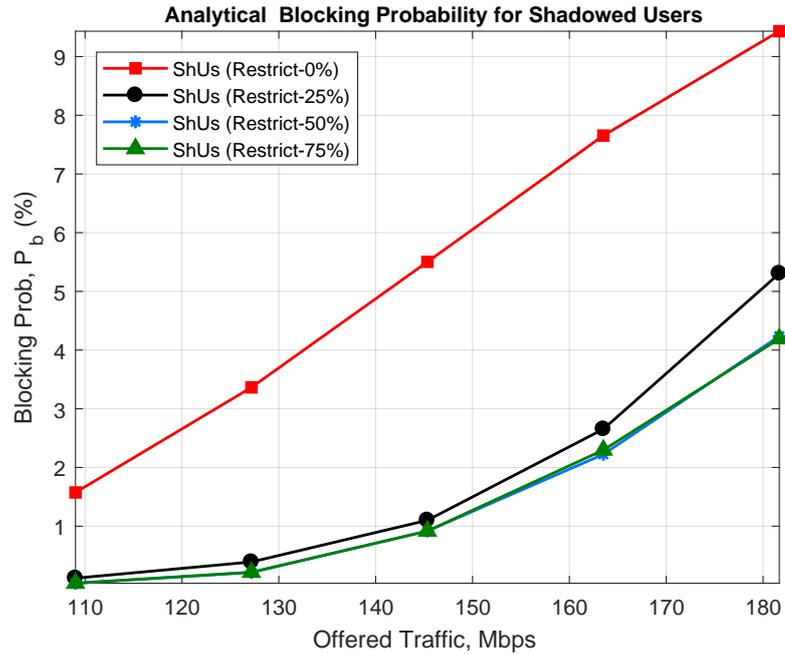
Figure 6.10: Impact of increasing offered traffic on shadowed users
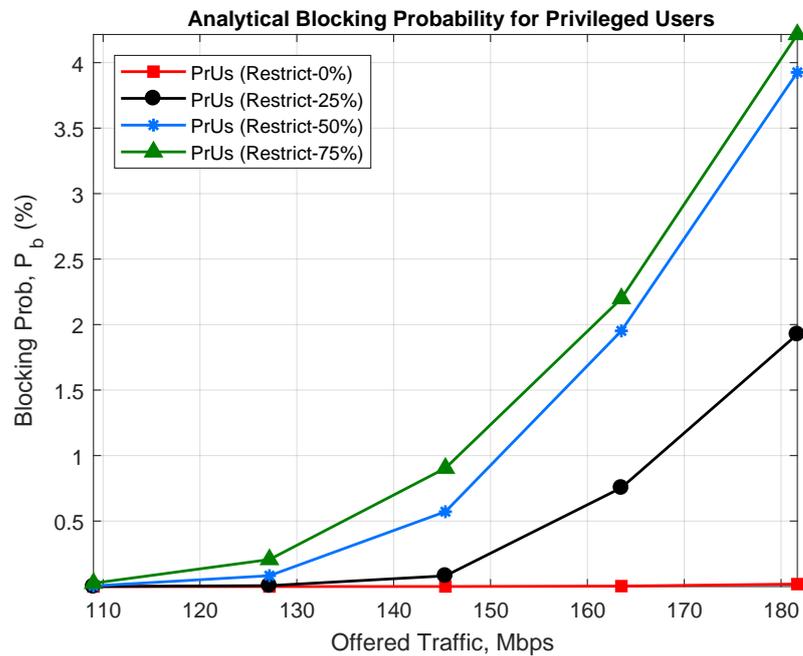
Figure 6.11: Privileged users blocking for different levels of Offered Traffic

Also, it is clear that the restriction process is most effective at high traffic loads. Thus, it is beneficial to suspend the restriction until high traffic levels because at low occupancy levels, it is likely to cause unnecessary blocking. Therefore, the users

with a high degree of choice are kept on those segments of a network which are inaccessible by the users with a more limited resource options.

6.6. Large Scale Simulation Results

This section presents results which are obtained when a number of constraints, like signal interference and retransmission are included in the system. It is carried out to evaluate the performance of restriction mechanism explained in Section 6.4 on a large scale network with different levels of occupancy. The BS antenna profiles, gains and transmit powers are defined in [5, 95]. Important simulation parameters are listed in Table 6.1.

Table 6.1: Simulation Parameters

Parameters	Values
Coverage area	450x450 m
Building size	75x75 m
Street width	15 m
Building height	6 m
Transmit power of RBS/ABS	35/35 dBm
Transmit power of MS for USCN/MBS	10/23 dBm
Antenna height of FB/RBS/ABS/MS/MBS	10/4/4/1.5/15 m
Antenna Beams for Rx FN/RBS/ABS	2/3/4
Antenna Beams for Tx RBS/ABS	2/4
Antenna elements in MBS	8
SINR threshold for USCN/MBS	1.8/-3 dB
Log-normal shadowing factor	3 dB
Carrier frequency for USCN/BN/MBS	3.5/75/3.5 GHz
Number of FN/RBS/ABS/MS	4/16/16/2000
Resource blocks for USCN/BN/MBS	20/20/20
Inter-arrival time	Exponential Distribution
Mean file size	2 MB
Iteration per offered traffic	200k

The main assumptions in this work are summarized earlier in Section 4.3. The metric used in this simulation to assess the full scale network performance is the blocking probability. It is worth mentioning that there are three main factors which give rise to blocking in the system i) Unavailability of resources. ii) Due to restriction process. iii) Due to poor quality of a channel.

Figure 6.12 illustrates the effect of increasing the restriction level on the performance of full scale network. In the absence of restriction, the blocking of shadowed users is severely high whilst for the privileged users, it is at very low level.

As indicated in Figure 6.12, the system achieves fairness around 50% of restriction for all levels of occupancy. It is clear that the restriction process is most adequate at high traffic loads. Thus, it is appropriate to postpone the restriction until high occupancy because at low traffic loads it is likely to cause unnecessary blocking in the system. These simulation results also validate the analytical results mentioned in Section 6.5.
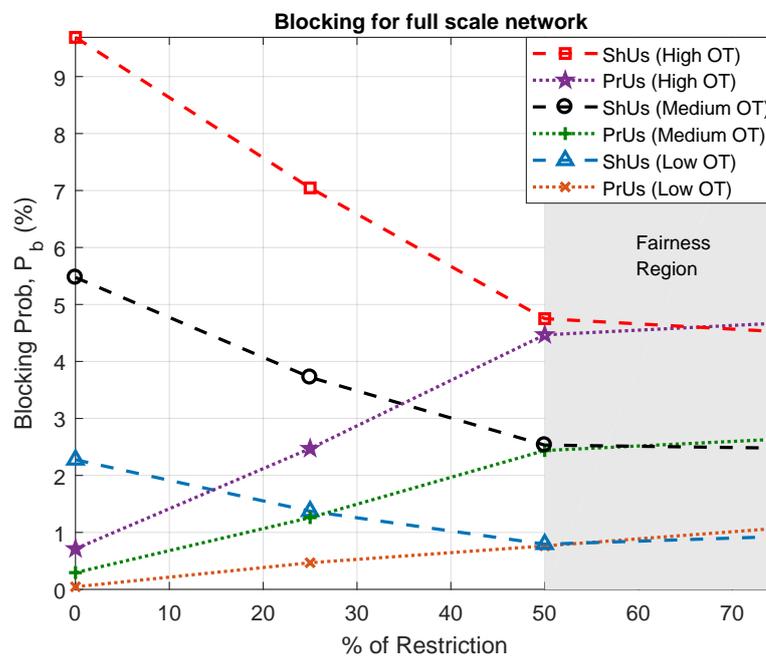
Figure 6.12: Performance evaluation of restriction for all levels of offered traffic

From Figure 6.13, it is clear that in case of restriction, the rise in blocking for the privileged users is comparatively higher from the analytical model explained in Case Study # 1. It is due to the presence of signal interference and retransmission in the system. These simulation results also validate that the basic premise of applying the restriction factor is still applicable on a large scale. Furthermore, it can be predicted from the analytical model what the restriction level needs to be for a complete fair system.
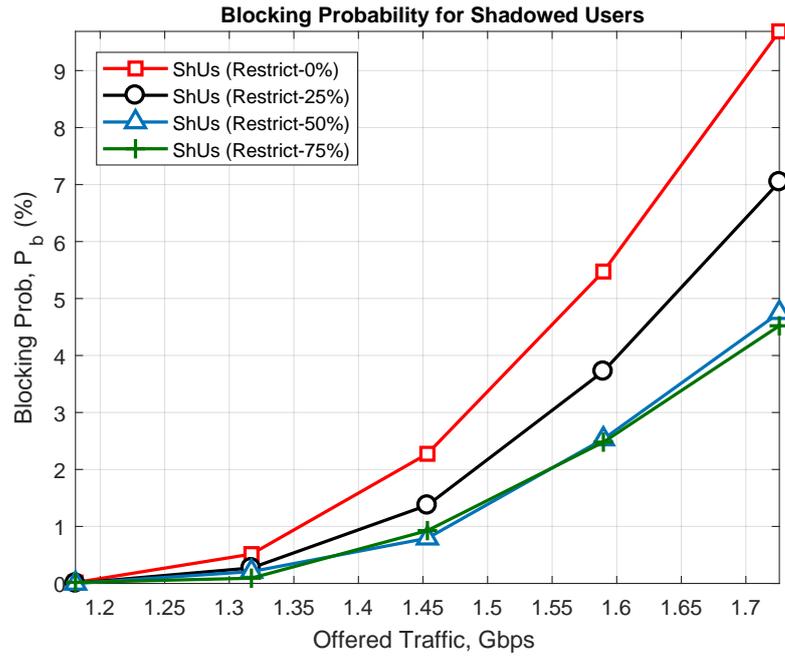
Figure 6.13: Impact of increasing offered traffic on Shadowed Users
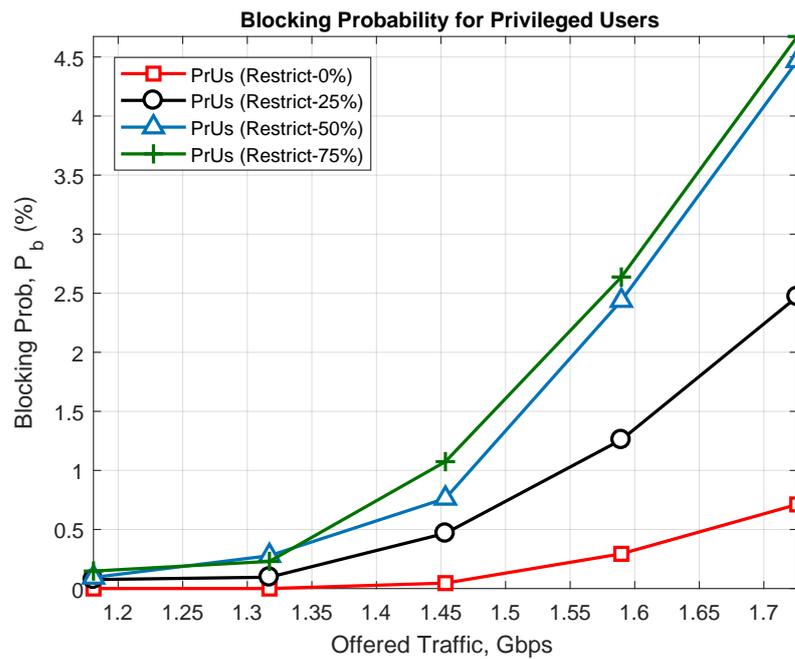
Figure 6.14: Impact of increasing offered traffic on Privileged Users

6.7. Conclusions

In this chapter, a mm-wave, multi-hop, multi-tier heterogeneous network is investigated for the improvement in GoS and to achieve overall network fairness which is a third

key objective of this thesis. A restriction mechanism has been applied to this multi-tier network which sacrifices some diversity of mobile users for the sake of others. The users are divided into two user groups. This classification of users is based on certain network parameters like geographical sites, azimuth/elevation angles and antenna equipment choices.

Different levels of restriction mechanism are examined and a comparison is drawn using certain performance metrics. From the analytical model, it is concluded that the blocking of shadowed users are significantly reduced from 7% to 4.5% at high traffic loads. It is achieved by sacrificing some choice flexibility of full access users for more limited type of users. The restriction mechanism reserves some portion of ultra small cells network for users with restricted choice flexibility, once the load level crosses that load threshold. It is this compensation effect that allows the network to achieve a balanced and fair GoS performance.

From both the case studies, it is concluded that such a restriction technique will be useful in real systems, where users having greater freedom of choice can be directed on to other parts of the network, inaccessible by other users. It is also shown how the multi-hop, ultra-small cells network operates well in the presence of macro base station and provides the flexibility to adapt to the dynamic changes in the network.

Chapter 7.   Conclusions and Future Work

Contents

7.1. Conclusions

The work presented in this thesis has concentrated on designing and evaluating a novel wireless dual-hop backhaul architecture using mm-wave that serves a densely populated area. The main objective of the thesis hypothesis was to understand whether the backhaul could be made more resilent and robust whilst improving resource utilization, energy efficiency and overall system fairness. This study have shown that in general this is possible. This has been looked at in a variety of ways and proven both through simulation and analytically.

Chapter 1 provided a general introduction to whole work. In Chapter 2, background information related to the area of next generation cellular network architectures, radio resource management, energy reduction and system fairness of wireless network have been presented.

The modelling, simulation and analysis methodologies used in this thesis have been presented in Chapter 3. The network has been modelled with appropriate topology, antenna, propagation and traffic models. Matlab is selected to carry out Monte Carlo simulation in this thesis. The complete simulator has been built upon the architecture, physical layer, traffic, spectrum and topology management modules. Results have been evaluated in a long term averaged manner to obtain steady state performance. Furthermore, Markov Modelling has been extensively discussed as an effective tool to analyse

system capacity and QoS.

A dual-hop short range mm-wave backhaul network architecture has been proposed in Chapter 4. It served densely populated ultra-small cells on street-level fixtures. Mm-wave backhaul offers abundant spectrum and therefore multi-gigabit data rates can significantly improve system performance. The proposed architecture also enabled exploiting the existing path diversity within the network, and thus delivered efficient resource utilization and high energy reductions. Due to this path diversity, some of backhaul base stations can be turned off at low occupancy levels while still maintaining the same QoS thus minimizing the energy expenditure of the whole network. It has been shown that the hop closest to the aggregation point experienced a bottleneck because too much traffic was directed towards it. Furthermore, a detailed overview has been provided for solving state probabilities.

Next, the application of energy based network management for 5G deployments has been studied in Chapter 5. This chapter introduced an Energy-Aware based Topology Management strategy for future cellular networks which fine-tunes the status of the network nodes depending on the traffic demands. A time based parameter has been used to forbid all the ON state nodes from going into SLEEP mode straight away when the load level is minimum. The scheme enhances the operation of the network in terms of energy efficiency as well as QoS by increasing the number of path alternatives between the communicating nodes.

Finally, an analytical model to control GoS performance for two different types of users with different levels of accessibility has been investigated in Chapter 6. A multi-tier heterogeneous network has been investigated for the improvement in Grade-of-Service and to achieve overall system fairness. As a way of improving system fairness, a restriction mechanism has been applied to the multi-tier network which sacrificed some diversity of mobile users for the sake of others and it is shown to be effective. The users have been divided into two user groups. This classification of users has been based on certain network parameters. Different case studies have been investigated and it

has been concluded that such a restriction technique is useful in situations where users having greater freedom of choice could be directed on to the parts of the networks, inaccessible by other users.

7.2. Novel Contributions

This thesis has proposed novel backhaul architecture options and evaluated various characteristics of designing backhaul solutions that serves a dense 5G cellular network. These highlight the need to design flexible and robust backhaul networks which can deliver tailored services for different deployment scenarios that support a vast range of applications with different performance requirements. The details of the original contributions are given in this section.

- **Dual-hop Backhaul Connectivity**

  To increase the network densification, other people have looked at scenarios where the nodes were mounted onto the street infrastructure [9, 28]. In this thesis, it has been extended further to increase the number of hops to two for improving the path diversity and the flexibility of the backhaul network. This work is published in [13].

  A novel dual-hop backhaul network architecture has been proposed where the nodes are mounted onto the street lamps and traffic lights to provide cost-effective and robust millimetre wave backhaul network for outdoor dense ultra-small cells. This approach can dramatically reduce the expenditure and deployment time compared to a more conventional approach whereby the nodes are typically mounted on high rise buildings.

- **System Fairness in a Multi-hop, Multi-tier Dense Network**

  A multi-tier dense network has been investigated for the improvement in GoS and to achieve overall system fairness. Based on the restriction mechanism, which was originally developed in [18] for High Altitude Platforms. It has been applied here

for the first time to multi-hop, multi-tier dense network. Based on the mathematical model, the restriction technique reserved some portion of the network for users with limited choice options to allow balanced and fair GoS performance. This work will be submitted to Transactions on Emerging Telecommunications Technologies.

- **Exploiting Path Diversity for Backhaul Topology Management**

The proposed topology management strategy exploited path diversity in dense urban scenario to provide a robust and flexible dual-hop backhaul solution. Other people have investigated a single-hop backhaul network, but in this thesis the TM strategy has been applied to a dual-hop setup. It has been shown to improve the network performance in terms of power savings while still preserving exactly the same high QoS demand. Due to this path diversity, some of backhaul nodes can be turned off at low occupancy levels thus minimizing the energy expenditure of the whole network. This work has published in [3]. It is expected that the technique can also be applied more generally to other types of dual-hop mm-wave networks, where more complex MIMO based antenna systems are used.

- **Two Different Network Architectures Working Simultaneously**

It has also been illustrated that how the multi-hop, ultra-small cells network operated well in the presence of macro base station and provides the flexibility to adapt to the dynamic changes in the network. The novel contribution in this case is that how the available flexibility in one tier of the network can be sacrificed for the sake of users associated to the second tier of the network alone. This work will be submitted to Transactions on Emerging Telecommunications Technologies.

7.3. Future Work

This section presents recommendations for future work, mainly built on the work of this thesis. The potential related applications and the extension of the ideas proposed in the thesis are discussed.

- **More User Groups in a HetNet Scenario**

  Only two user groups have been investigated in Chapter 6 for examining the system fairness. One user group has association with the dense ultra-small cells network while the second user group has access to both the tiers of the network. However, we can analyse a third user as well which has only access to the MBS part of the network. Intially, the restriction mechanism has been applied only in one dimensional but we can apply the restrction along both the dimensions. Necessary adjustments have been made in Figure 7.1 to illustrate the flow of system between any two transition states.

- **Splitting the Data Traffic of Privileged Users Unequally**

  In Case Study # 2 of Chapter 6, the data traffic generated by the privileged users has been divided equally between both tiers of HetNet. However, we can study the effect of splitting the same data traffic into any two unequal portions.

- **Different Functions for the Restriction Mechanism**

  In Chapter 6, a constant restriction function has been used to implement the restriction process for achieving network fairnesss. However, different types of restriction functions can be studied that progressively contrains access to channels at high occupancy level [18].

- **Millimetre-wave Phased Array Antennas**

  In recent years, extensive research campaigns are carried out on combination of massive MIMO and mm-wave phased array antennas[47, 48, 49]. These phased array antennas can exploit a more complex mm-wave channels with multipath, e.g. that
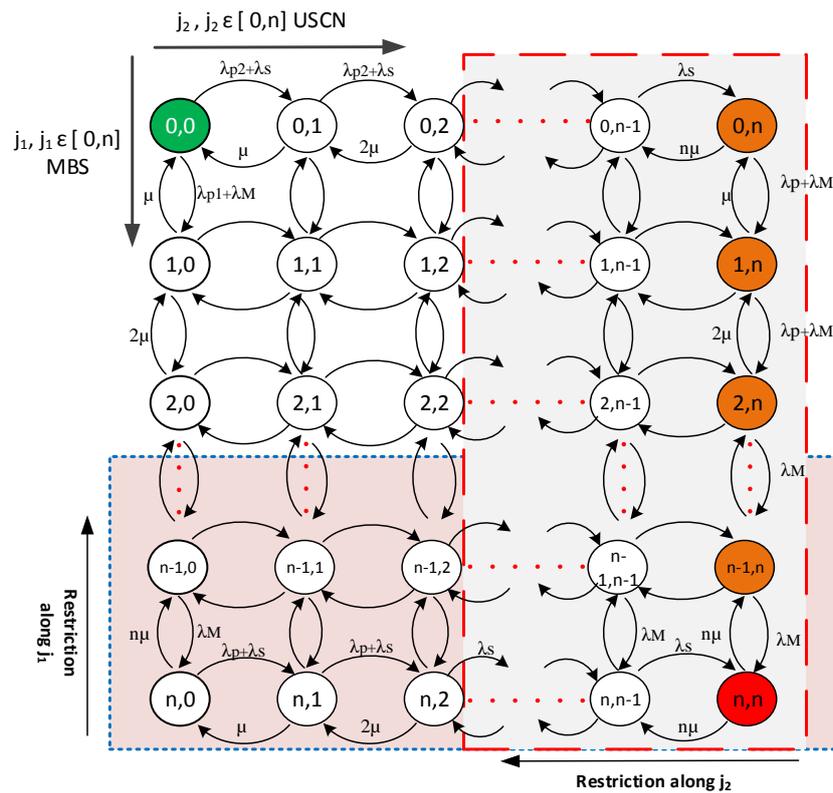
Figure 7.1: State-Transition Diagram when the restriction process is applied along both the dimensions

caused by reflections. In addition, to increased capacity and spectral efficiency, such antenna systems also have compact dimensions. Signal processing issues including channel characterization, estimation and new modulation techniques need to be investigated. Moreover, innovative antenna architectures will be required to handle the wideband frequency channels.

- **Load based Path Selection Process**

The path selection process in Chapter 4 has been performed on the basis of shortest distance link. However, different criteria e.g, load based can also be used for the path selection algorithm. This will make the load level more evenly distributed across the entire network.

- **Artificial Intelligence for Backhaul Topology Management**

The complexity of future mobile networks presents new challenges for network management. Although centralised controlled network can offer high resource utilisation. It is also worth considering distributed solutions for flexibility and scalability reasons. For a highly dynamic network environment, machine learning techniques e.g. Reinforcement Learning can be used to improve performance. The historical information of the mobility patterns can be used to predict the user movement and traffic load conditions for routing the backhaul network.

- **Energy Efficient mm-wave Backhauling**

Energy-efficient topology management strategies have been investigated on a large scale in wireless access networks but relatively little attention has been paid to the overall backhaul energy consumption management options. In Chapter 5, a backhaul TM strategy has been proposed for dual-hop backhaul architecture in order to minimise the backhaul network energy consumption. There, the BS energy consumption model for Beyond Next Generation Mobile Broadband system was used [24, 4]. However, for mm-wave backhaul architectures, the energy efficiency aspect has not been studied. The research on energy consumption models for mm-wave BSs is still in its early age, especially with the brisk ongoing progress in the development of radio hardware. Hence it would be interesting to evaluate the advantages of switching part of the backhaul network OFF or switching them to adaptive sleep mode.

Glossary

3GPP 3rd Generation Partnership Project

ABS Access Base Station

BBU Baseband Unit

BP   Blocking Probability

CoMP Coordinated Multipoint

C-RAN Cloud-RAN

CAPEX Capital Expenditure

D2D Device-to-Device

ECR Energy Consumption Rating

FN   Fibre Node

FDD Frequency Division Duplex

FDMA Frequency Division Multiple Access

FTP File Transfer Protocol

GoS Grade-of-Service

HetNet Heterogeneous Network

HBS Hub Base Station

ITU-R International Telecommunication Union-Radiocommunication Sector

IMT International Mobile Telecommunications

KPI Key Performance Indicator

LOS Line-of-Sight

LTE Long Term Evolution

LTE-A Long Term Evolution-Advanced

M2M Machine-to-Machine

MIMO Multiple-Input-Multiple-Ouput

MAC Medium Access Control

NLOS Non-Line-of-Sight

OT   Offered Traffic

OPEX Operational Expenditure

PD   Path Diversity

QoS Quality-of-Service

RBS Relay Base Station

RAN Radio Access Network

RAT Radio Access Technology

RRH Remote Radio Head

RSSI Recieved Signal Strength Indicator

RSRP Reference Signal Recieve Power

RB   Resource Block

RRM Radio Resource Management

SINR Signal-to-Interference plus Noise Ratio

USCN Ultra-Small Cell Network

UG   User Group

UHD Ultra-High Definition

List of References

[1] *Small Cell Forum: White Paper on Backhaul Technologies for Small Cells*. 2015.

[2] *IMT Vision - Framework and overall objectives of the future development of IMT for 2020 and beyond*. ITU-R M.2083-0, Geneva, Sep. 2016.

[3] A. Ahmed and D. Grace. Energy-Aware Topology Management for 5G Dual-hop Ultra-High Capacity Backhaul Networks Exploiting Path Diversity. In *2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN)*, pages 1020–1025, July 2016.

[4] University of York. "Bungee Deliverable D4.1.1:Interim Simulation. *Beyond Next Generation Mobile Networks Project Tech*, V1.0, February, 2011.

[5] ETSI TR. *Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Frequency (RF) System Scenarios(3GPP TR 36.942 version 11.0.0 Release 11)*. 2012.

[6] J. Lun and D. Grace. Cognitive Green Backhaul Deployments for Future 5G Networks. In *2014 1st International Workshop on Cognitive Cellular Systems (CCS)*, pages 1–5, Sept 2014.

[7] C. Dehos, J. Gonzalez, A. Domenico, D. Ktenas, and L. Dussopt. Millimeter-wave Access and Backhauling: The Solution To the Exponential Data Traffic Increase in 5G Mobile Communications Systems? volume 52, pages 88–95, September 2014.

[8] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta. Massive mimo for next generation wireless systems. *IEEE Communications Magazine*, 52(2):186–195, February 2014.

[9] J. Lun, D. Grace, A. Burr, Y. Han, K. Leppanen, and T. Cai. Millimetre Wave Backhaul/fronthaul Deployments for Ultra-Dense Outdoor Small Cells. In *2016 IEEE Wireless Communications and Networking Conference*, pages 1–6, April 2016.

[10] A. I. Sulyman, A. T. Nassar, M. K. Samimi, G. R. Maccartney, T. S. Rappaport, and A. Alsanie. Radio Propagation Path Loss Models for 5G Cellular Networks in the 28 GHz and 38 GHz Millimeter-Wave Bands. *IEEE Communications Magazine*, 52(9):78–86, September 2014.

[11] G. R. MacCartney and T. S. Rappaport. 73 GHz Millimeter Wave Propagation Measurements for Outdoor Urban Mobile and Backhaul Communications in New York City. In *2014 IEEE International Conference on Communications (ICC)*, pages 4862–4867, June 2014.

[12] B. Bangerter, S. Talwar, R. Arefi, and K. Stewart. Networks and Devices for the 5G Era. volume 52, pages 90–96, February 2014.

[13] A. Ahmed and D. Grace. A Dual-hop Backhaul Network Architecture for 5G Ultra-Small Cells using Millimetre-Wave. In *2015 IEEE International Conference on Ubiquitous Wireless Broadband (ICUWB)*, pages 1–6, Oct 2015.

[14] J. Qiao, L. X. Cai, X. S. Shen, and J. W. Mark. Enabling Multi-hop Concurrent Transmissions in 60 GHz Wireless Personal Area Networks. *IEEE Transactions on Wireless Communications*, 10(11):3824–3833, November 2011.

[15] Ajay R. Mishra. *Fundamentals of Cellular Network Planning and Optimisation.* A John Wiley and Sons Ltd, 2011.

[16] E. Oh and B. Krishnamachari. Energy Savings through Dynamic Base Station Switching in Cellular Wireless Access Networks. In *2010 IEEE Global Telecommunications Conference GLOBECOM 2010*, pages 1–5, Dec 2010.

[17] A. Fisusi, D. Grace, and P. Mitchell. Interference Aware, Energy Efficient Resource Allocation for Beyond Next Generation Mobile Networks. In *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 2197–2202, Sept 2013.

[18] Y. Liu, D. Grace, and P. D. Mitchell. Exploiting Platform Diversity for GoS Improvement for Users with Different High Altitude Platform Availability. *IEEE Transactions on Wireless Communications*, 8(1):196–203, Jan 2009.

[19] D. Grace K. Katzis, D. A. J. Pearce. Fairness in Channel Allocation in a High Altitude Platform Communication System exploiting Cellular Overlap. In *Wireless Personal Mobile Conference*, 2004.

[20] Gordon L. Stuber. *Principles of Mobile Communication.* Kluwer Academic Publishers, 3rd edition edition, 2009.

[21] M. F. Hossain, K. S. Munasinghe, and A. Jamalipour. An Eco-Inspired Energy Efficient Access Network Architecture for Next Generation Cellular Systems. In *2011 IEEE Wireless Communications and Networking Conference*, pages 992–997, March 2011.

[22] M. Agiwal, A. Roy, and N. Saxena. Next Generation 5G Wireless Networks: A Comprehensive Survey. *IEEE Communications Surveys Tutorials*, 18(3):1617–1655, thirdquarter 2016.

[23] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi. A Survey on 3GPP Heterogeneous Networks. *IEEE Wireless Communications*, 18(3):10–21, June 2011.

[24] Y. Han, D. Grace, and P. Mitchell. Energy Efficient Topology Management for Beyond Next Generation Mobile Broadband Systems. pages 331–335, Aug 2012.

[25] A. Fisusi, D. Grace, and P. Mitchell. Interference Aware, Energy Efficient Resource Allocation for Beyond Next Generation Mobile Networks. In *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 2197–2202, Sept 2013.

[26] *Skyfiber Forum: White Paper on Breaking The Backhaul Bottleneck.* 2016.

[27] *Ericsson: White Paper on Wireless Backhaul in Future Heterogeneous Networks.* 2017.

[28] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. T. Sukhavasi, C. Patel, and S. Geirhofer. Network Densification: The Dominant Theme for Wireless Evolution into 5G. *IEEE Communications Magazine*, 52(2):82–89, February 2014.

[29] 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*, 30(3):497–508, April 2012.

[30] Yi Qian Rose Q. Hu. *Heterogeneous Cellular Networks*. A John Wiley and Sons Ltd, 2nd edition edition, 2015.

[31] G. Shen, J. Liu, D. Wang, J. Wang, and S. Jin. Multi-hop Relay for Next-Generation Wireless Access Networks. *Bell Labs Technical Journal*, 13(4):175–193, Winter 2009.

[32] A. S. Tanenbaum and David J. Hetherall. *Computer Networks*. Prentice Hall Publishers, 5th edition, 2003.

[33] X. Ge, H. Cheng, M. Guizani, and T. Han. 5G Wireless Backhaul Networks: Challenges and Research Advances. *IEEE Network*, 28(6):6–11, Nov 2014.

[34] O. Tipmongkolsilp, S. Zaghloul, and A. Jukan. The Evolution of Cellular Backhaul Technologies: Current Issues and Future Trends. *IEEE Communications Surveys Tutorials*, 13(1):97–113, First 2011.

[35] M. A. Khalighi and M. Uysal. Survey on Free Space Optical Communication: A Communication Theory Perspective. *IEEE Communications Surveys Tutorials*, 16(4):2231–2258, Fourthquarter 2014.

[36] U. Siddique, H. Tabassum, E. Hossain, and D. I. Kim. Wireless Backhauling of 5G Small Cells: Challenges and Solution Approaches. *IEEE Wireless Communications*, 22(5):22–31, October 2015.

[37] N. Haklai M. Goldhamer, Z. e. Gross and O. Marinchenco. "bungee deliverable d1. 2: Baseline bungee architecture,. *Beyond Next Generation Mobile Networks Project Tech*, V1.0, November , 2010.

[38] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta. Massive MIMO for Next Generation Wireless Systems. *IEEE Communications Magazine*, 52(2):186–195, February 2014.

[39] B. Li, D. Zhu, and P. Liang. Small Cell In-Band Wireless Backhaul in massive MIMO Systems: A Cooperation of Next-Generation Techniques. *IEEE Transactions on Wireless Communications*, 14(12):7057–7069, Dec 2015.

[40] *Ericsson: White Paper for NLOS Microwave Backhaul Solutions*. 2014.

[41] K. Dyer. Ericsson makes case for Non Line of Sight backhaul scenario for small cells Network. `http://the-mobile-network.com/2013/03/ericsson-makes-case-for-nlos-backhaul-for-small-cells/`, 2013. [Online; accessed 24-August-2016].

[42] G. R. MacCartney, Junhong Zhang, Shuai Nie, and T. S. Rappaport. Path Loss Models for 5G Millimeter Wave Propagation Channels in Urban Microcells. pages 3948–3953, Dec 2013.

[43] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun. Wideband Millimeter-Wave Propagation Measurements and Channel Models for Future Wireless Communication System Design. *IEEE Transactions on Communications*, 63(9):3029–3056, Sept 2015.

[44] Yong Niu, Yong Li, Depeng Jin, Li Su, and Athanasios V. Vasilakos. A Survey of Millimeter Wave Communications (mm-wave) for 5G: Opportunities and Challenges. *Wireless Networks*, 21(8):2657–2676, Nov 2015.

[45] Z. Pi and F. Khan. An Introduction to Millimeter-Wave Mobile Broadband Systems. *IEEE Communications Magazine*, 49(6):101–107, June 2011.

[46] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh. Millimeter Wave Beamforming for Wireless Backhaul and Access in Small Cell Networks. *IEEE Transactions on Communications*, 61(10):4391–4403, October 2013.

[47] M. Stanley, Y. Huang, H. Wang, H. Zhou, A. Alieldin, and S. Joseph. A Novel mm-Wave Phased Array Antenna with 360 Coverage for 5G Smartphone Applications. In *2017 10th UK-Europe-China Workshop on Millimetre Waves and Terahertz Technologies (UCMMT)*, pages 1–3, Sept 2017.

[48] X. Gao, L. Dai, and A. M. Sayeed. Low RF-Complexity Technologies to Enable Millimeter-Wave MIMO with Large Antenna Array for 5G Wireless Communications. *IEEE Communications Magazine*, 56(4):211–217, APRIL 2018.

[49] B. Sadhu, Y. Tousi, J. Hallin, S. Sahl, S. Reynolds, . Renstrm, K. Sjgren, O. Haapalahti, N. Mazor, B. Bokinge, G. Weibull, H. Bengtsson, A. Carlinger, E. Westesson, J. E. Thillberg, L. Rexberg, M. Yeck, X. Gu, D. Friedman, and A. Valdes-Garcia. 7.2 A 28GHz 32-Element Phased-Array Transceiver IC with Concurrent Dual Polarized Beams and 1.4 Degree Beam-Steering Resolution for 5G Communication. In *2017 IEEE International Solid-State Circuits Conference (ISSCC)*, pages 128–129, Feb 2017.

[50] A. I. Sulyman, A. T. Nassar, M. K. Samimi, G. R. Maccartney, T. S. Rappaport, and A. Alsanie. Radio Propagation Path Loss Models for 5G Cellular Networks in the 28 GHz and 38 GHz Millimeter-Wave Bands. *IEEE Communications Magazine*, 52(9):78–86, September 2014.

[51] Wikipedia. Free-Space Optical Communication. https://en.wikipedia.org/wiki/Free-space_optical_communication. [Online; accessed 06-July-2018].

[52] Eric J. Korevaar Isaac I. Kim. Availability of Free-Space Optics FSO and Hybrid FSO/RF Systems, 2001.

[53] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. A. Uusitalo, B. Timus, and M. Fallgren. Scenarios for 5G Mobile and Wireless Commu-

nications: The Vision of the METIS Project. *IEEE Communications Magazine*, 52(5):26–35, May 2014.

[54] Y. C. Hsu and Y. D. Lin. Multihop Cellular: A Novel Architecture for Wireless Data Communications. *Journal of Communications and Networks*, 4(1):30–39, March 2002.

[55] A. Gupta and R. K. Jha. A Survey of 5G Network: Architecture and Emerging Technologies. *IEEE Access*, 3:1206–1232, 2015.

[56] M. Agiwal, A. Roy, and N. Saxena. Next Generation 5G Wireless Networks: A Comprehensive Survey. *IEEE Communications Surveys Tutorials*, 18(3):1617–1655, thirdquarter 2016.

[57] Z. Gao, L. Dai, D. Mi, Z. Wang, M. A. Imran, and M. Z. Shakir. Mmwave massive-mimo-based wireless backhaul for the 5g ultra-dense network. *IEEE Wireless Communications*, 22(5):13–21, October 2015.

[58] J. Wu, Y. Zhang, M. Zukerman, and E. K. N. Yung. Energy-Efficient Base-Stations Sleep-Mode Techniques in Green Cellular Networks: A Survey. *IEEE Communications Surveys Tutorials*, 17(2):803–826, Secondquarter 2015.

[59] S. Mclaughlin, P. M. Grant, J. S. Thompson, H. Haas, D. I. Laurenson, C. Khirallah, Y. Hou, and R. Wang. Techniques for Improving Cellular Radio Base Station Energy Efficiency. *IEEE Wireless Communications*, 18(5):10–17, October 2011.

[60] E. Oh and B. Krishnamachari. Energy Savings through Dynamic Base Station Switching in Cellular Wireless Access Networks. In *2010 IEEE Global Telecommunications Conference GLOBECOM 2010*, pages 1–5, Dec 2010.

[61] S. Rehan and D. Grace. Energy-Aware Topology Management for High Capacity Density Temporary Event Networks. In *2013 International Conference on Advanced Technologies for Communications (ATC 2013)*, pages 307–311, Oct 2013.

[62] S. Tombaz, P. Monti, F. Farias, M. Fiorani, L. Wosinska, and J. Zander. Is Backhaul Becoming a Bottleneck for Green Wireless Access Networks? In *2014 IEEE International Conference on Communications (ICC)*, pages 4029–4035, June 2014.

[63] S. Rehan and D. Grace. Combined Green Resource and Topology Management for Beyond Next Generation Mobile Broadband Systems. In *2013 International Conference on Computing, Networking and Communications (ICNC)*, pages 242–246, Jan 2013.

[64] Behrouz A. Forouzan. *Data Communcations and Networking*. McGraw-Hill Higher Education, 4th edition, 2007.

[65] R. Wang, J. S. Thompson, H. Haas, and P. M. Grant. Sleep Mmode Design for Green Base Stations. *IET Communications*, 5(18):2606–2616, Dec 2011.

[66] L. Saker, S. E. Elayoubi, and H. O. Scheck. System Selection and Sleep Mode for Energy Saving in Cooperative 2G/3G Networks. In *2009 IEEE 70th Vehicular Technology Conference Fall*, pages 1–5, Sept 2009.

[67] A. Conte, A. Feki, L. Chiaraviglio, D. Ciullo, M. Meo, and M. A. Marsan. Cell Wilting and Blossoming for Energy Efficiency. *IEEE Wireless Communications*, 18(5):50–57, October 2011.

[68] L. A. Suarez, L. Nuaymi, and J. M. Bonnin. Energy Performance of a Distributed BS Based Green Cell Breathing Algorithm. In *2012 International Symposium on Wireless Communication Systems (ISWCS)*, pages 341–345, Aug 2012.

[69] S. Rehan and D. Grace. Macro-Cell Overlaid Green Topology Management for Beyond Next Generation Mobile Broadband Systems. In *2013 Future Network Mobile Summit*, pages 1–9, July 2013.

[70] L. Saker, S. E. Elayoubi, R. Combes, and T. Chahed. Optimal Control of Wake Up Mechanisms of Femtocells in Heterogeneous Networks. *IEEE Journal on Selected Areas in Communications*, 30(3):664–672, April 2012.

[71] MathWorks. Improving the Efficiency of RF Power Amplifiers with Digital Predistortion. `https://uk.mathworks.com/company/newsletters/articles/improving-the-efficiency-of-rf-power-amplifiers.html`. [Online; accessed 07-July-2018].

[72] J. S. Fu and A. Mortazawi. Improving Power Amplifier Efficiency and Linearity Using a Dynamically Controlled Tunable Matching Network. *IEEE Transactions on Microwave Theory and Techniques*, 56(12):3239–3244, Dec 2008.

[73] N. Blaunstein, R. Giladi, A. Freedman, and M. Levin. Unified Approach of GoS Optimization for Fixed Wireless Access. *IEEE Transactions on Vehicular Technology*, 51(1):200–208, Jan 2002.

[74] D. A. J. Pearce K. Katzis and D. Grace. Fairness in Channel Allocation in a High Altitude Platform Communication System Exploting Cellular Overlap. In *Wireless Personal Mobile Conference*, pages 17–25, July 2004.

[75] R. Prasad H. Harada. *Simulation and Software Radio for Mobile Communications*. Prentice Hall Publishers, 2nd edition, 1998.

[76] Y. Chen D. Xue. *System Simulation Techniques with MATLAB and Simulink*. The Wiley Publishers, 1st edition edition, 2013.

[77] Wikipedia. The Monte Carlo Method. `https://en.wikipedia.org/wiki/Monte_Carlo_method`. [Online; accessed 5-April-2018].

[78] Monte Carlo Simulation. Probabilistic Engineering Design. `http://web.mst.edu/~dux/repository/me360/ch8.pdf`. [Online; accessed 25-May-2018].

[79] H. Abu-Ghazaleh and A. S. Alfa. Application of Mobility Prediction in Wireless Networks Using Markov Renewal Theory. *IEEE Transactions on Vehicular Technology*, 59(2):788–802, Feb 2010.

[80] U. Narayan Bhat. *An Introduction to Queueing Theory*. Birkhauser Boston, 1st edition, 2008.

[81] C. M. Grinstead and J. L. Snell. *Introduction to Probability. Providence, RI: American Mathematical Society (RI)*. 2000.

[82] Robert B Cooper. *Introduction to Queueing Theory*. North Holland New York Oxford, 2nd edition, 1981.

[83] Leonard Kleinrock. *Introduction to Queueing Systems*. John Wiley and Sons, 1st edition, 1975.

[84] D. A. J. Pearce. *Communications ECAD course notes*. University of York, 2001.

[85] T. S. Rappaport. *Wireless Communications: Principles and Practice*. Upper Saddle River, NJ. : Prentice Hall, 2002.

[86] C. Han, T. Harrold, S. Armour, I. Krikidis, S. Videv, P. M. Grant, H. Haas, J. S. Thompson, I. Ku, C. X. Wang, T. A. Le, M. R. Nakhai, J. Zhang, and L. Hanzo. Green Radio: Radio Techniques to Enable Energy-Efficient Wireless Networks. *IEEE Communications Magazine*, 49(6):46–54, June 2011.

[87] Pekka Kysti and J. Meinil. *WINNER II Channel Models, IST-4-027756 WINNER II D1.1.2 V1.2*. 2007.

[88] A. Mozharovskiy, A. Artemenko, A. Sevastyanov, V. Ssorin, and R. Maslennikov. Beam-Steerable Integrated Lens Antenna with Waveguide Feeding System for 71 7176/8186 GHz Point-to-Point Applications. In *2016 10th European Conference on Antennas and Propagation (EuCAP)*, pages 1–5, April 2016.

[89] Wikipedia. Exponential Distribution. `https://en.wikipedia.org/wiki/Exponential_distribution`. [Online; accessed 12-January-2018].

[90] 3rd Generation Partnership Project. *Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects)*. 2013.

[91] Q. Zhao and D. Grace. Application of Cognition Based Resource Allocation Strategies on a Multi-hop Backhaul Network. In *2012 IEEE International Conference on Communication Systems (ICCS)*, pages 423–427, Nov 2012.

[92] Wikipedia. GaussSeidel Method. `https://en.wikipedia.org/wiki/Gauss%E2%80%93Seidel_method`. [Online; accessed 25-March-2018].

[93] L. M. Correia, D. Zeller, O. Blume, D. Ferling, Y. Jading, I. Gdor, G. Auer, and L. V. Der Perre. Challenges and Enabling Technologies for Energy Aware Mobile Radio Networks. *IEEE Communications Magazine*, 48(11):66–72, November 2010.

[94] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang. What Will 5G Be? *IEEE Journal on Selected Areas in Communications*, 32(6):1065–1082, June 2014.

[95] Abimbola Fisusi, David Grace, and Paul Mitchell. Energy saving in a 5g separation architecture under different power model assumptions. *Computer Communications*, 105:89 – 104, 2017.

[96] J. Lun and D. Grace. Software Defined Network for Multi-Tenancy Resource Sharing in Backhaul Networks. In *2015 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pages 1–5, March 2015.

[97] H. Mei, P. Jiang, and J. Bigham. Dynamic Frequency Allocation and Network Reconfiguration on Relay Based Cellular Network. In *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 2984–2989, April 2012.

[98] Y. Li, Q. Miao, M. Peng, and W. Wang. Frequency Allocation in Two-Hop Cellular Relaying Networks. In *2007 IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1–5, Sept 2007.

[99] N. Morozs, D. Grace, and T. Clarke. Case-Based Reinforcement Learning for Cognitive Spectrum Assignment in Cellular Networks with Dynamic Topologies. In *2013 Military Communications and Information Systems Conference*, pages 1–6, Oct 2013.

[100] D. Cassioli, L. A. Annoni, and S. Piersanti. Characterization of Path Loss and Delay Spread of 60-GHz UWB Channels vs. Frequency. In *2013 IEEE International Conference on Communications (ICC)*, pages 5153–5157, June 2013.

[101] Tareq M. Shami, David Grace, and Alister Burr. User Association in Cell-less 5G Networks Exploiting Particle Swarm Optimisation, 08 2017.

[102] T. Issariyakul, E. Hossain, and A. S. Alfa. Markov-Based Analysis of End-to-End Batch Transmission in a Multi-hop Wireless Network. In *IEEE International Conference on Communications, 2005. ICC 2005. 2005*, volume 5, pages 3505–3509 Vol. 5, May 2005.

[103] R. Taori and A. Sridharan. Point-to-Multipoint In-band mm-wave Backhaul for 5G Networks. *IEEE Communications Magazine*, 53(1):195–201, January 2015.

[104] S. Cho, E. W. Jang, and J. M. Cioffi. Handover in Multihop Cellular Networks. *IEEE Communications Magazine*, 47(7):64–73, July 2009.