Load Aware Self Optimisation Based Capacity Scaling Techniques for 5G Cloud Radio Access Networks

by

ZAINAB FAKHRI

Supervisor: ***Prof. Hamed Al-Raweshidy***

Department of Electronic and Computer Engineering

College of Engineering, Design and Physical Sciences

Brunel University London

United Kingdom

A doctoral thesis submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy (Ph.D.) in Communications Engineering

June 2020

Abstract

Expanding network capacity is the main pillar to realising the fifth-generation (5G) vision as well as the use cases of Enhanced Mobile Broadband (eMBB), Ultra-Reliable Low-Latency Communications (URLLC) and Massive Machine-Type Communications (mMTC). Capacity extension approaches involve bandwidth extension by using the Millimetre wave (mmW) band and enhancing the existing bandwidth efficiency through network densification as well as Multiple Input Multiple Output (MIMO). However, both approaches increase system complexity and cost. The Cloud Radio Access Network (C-RAN) architecture is a promising solution for the network densification cost issue in addition to introducing a high level of cooperation between the cells. Furthermore, a Self-Organising Network (SON) is another solution aimed at bringing intelligence and autonomous proactive adaptability, thereby reducing the cost and complexity of cellular networks.

For this reason, the main focus of this thesis is measuring the merits of a SON in C-RAN, whereby various energy-efficient capacity scaling systems are proposed for the upcoming 5G and beyond.

First, a self-organising C-RAN is introduced, which dynamically adapts to varying network capacity demands. The number of active Base Band Units (BBUs) and Remote Radio Heads (RRHs) is scaled based on the Cell Differentiation and Integration (CDI) technique, according to the load demand, which is aimed at efficient resource utilisation without sacrificing the overall Quality of Service (QoS). A CDI algorithm is proposed in which semi-static CDI and dynamic BBU-RRH mapping for load balancing are jointly performed. Network load balance is formulated as an optimisation problem with constraints. Discrete Particle Swarm Optimisation (DPSO) is developed as an Evolutionary Algorithm (EA) to solve the optimisation problem and compared to the Exhaustive Search (ES) Algorithm. The CDI-enabled C-RAN shows significant throughput improvement compared to a fixed C-RAN. Specifically, an average throughput increase of 45.53% and average blocked users reduction of 23.149% is achieved. Additionally, a power model is proposed to estimate the overall power consumption of C-RAN. Approximately 16% power reduction is calculated in a CDI-enabled C-RAN when compared to a fixed C-RAN.

Second, since 5G in its preliminary deployment is going to coexist with 4G infrastructure, the CDI performance is tested for interworking between the 5G wireless access technology New Radio (NR) and Long Term Evolution-Advanced (LTE-A). There is an increase of 282.9% and 121.7% in throughput, whilst a reduction of 97.9% and 96.4% in the number of blocked users is observed in the NR CDI-CRAN network compared to fixed C-RAN and CDI-CRAN, respectively.

Third, a scheme for reducing interference and increasing the rate coverage of the conventional macrocell by exploiting the mmw band spectrum is proposed. Since the cell edge area has a lower coverage data rate due to high interference and less received power, it is served by an mmw band. Small mmw RRHs are overlaid onto a conventional macrocell at the cell edges in a Heterogeneous Cloud Radio Access Network (HC-RAN) architecture. The RRH clustering is formulated as a bin packing optimisation problem. This scheme has significantly higher coverage and rate compared to conventional HC-RAN. A gain of 20% and 40% is achieved in coverage and rate when compared to systems applying the Soft Frequency Reuse (SFR) technique and conventional HC-RAN.

Acknowledgements

First and foremost, I would like to express my sincerest gratitude to my supervisors, Professor Hamed Al-Raweshidy, for his continuous support and mentoring. I would like to thank him for his helpful advice, guidance, constructive-feedback on my research and unlimited support for my work. I could not have accomplished this work without his support and encouragement.

Also, I am deeply grateful to my mentor and friends Dr. Firas Sabir, Dr. Muhammad Khan, Dr. Ismail Hburi. Their advice, fruitful-collaborations and insightful-discussions have assisted me in my research. I appreciate that they shared their valuable experiences with me. I would like to acknowledge my colleagues at the Wireless Network and Communication Centre Wireless Networks and Communications Centre (WNCC), whom I have made throughout these past few years, their continuous support have been invaluable to me.

Finally, I would also like to extend my most profound gratitude to my family for their encouragement and precious moral support throughout the years of research. They always inspire me to pursue my study and encourage me to complete this work.

List of Publications

Journal Paper

1. **Zainab H. Fakhri**, M. Khan, Firas Sabir, and H. S. Al-Raweshidy, "A Resource Allocation Mechanism for Cloud Radio Access Network Based on Cell Differentiation And Integration Concept", in *IEEE Transactions on Network Science and Engineering, VOL. 5, NO. 3, PP. 261-274, Dec. 2018, DOI no.: 10.1109/TNSE.2017.2754101*.

2. M. Khan, **Zainab H. Fakhri**, Firas Sabir, and H. S. Al-Raweshidy, "Semi-Static Cell Differentiation And Integration With Dynamic BBU-RRH Mapping In Cloud Radio Access Network', in *IEEE Transactions on Network Service and Management, VOL. 15, NO. 1, PP. 289-303, MARCH 2018, DOI no.: 10.1109/TNSM.2017.2771622*.

Conference Paper

1. **Zainab H. Fakhri**, and H. S. Al-Raweshidy, "An Interference Mitigation Scheme For Millimetre Wave Heterogeneous Cloud Radio Access Network with Dynamic RRH Clustering", *IEEE International Symposium on Networks, Computers and Communications (ISNCC): Wireless and Mobile Networks, Istanbul, 2019, pp. 169-174. 978-1-7281-1244-2/19/©2019 IEEE*

Table of contents

List of Symbols

General Conventions

The max() is used for finding the highest number in a vector.

Mathematical Operators and Functions

$\lfloor . \rfloor$	Represents floor function
$\lceil . \rceil$	Represents the ceil function
$\sum$	The sum operation
$\bigcup$	Union of two sets
$\mathbb{E}$	Expected value
$\mathbf{R_{ij}}$	Element of the matrix $\mathbf{R}$ at row i, column j
$\mathbb{S}$	Sample set

Special Symbols

α, β	Control parameters of optimisation function
$\beta_{k_{r_i}}$	The channel gain between user k in the cell edge area and its serving RRH r_i
δ	Path-loss exponent

δ_{mmL}	LOS mmw channel path-loss exponent
δ_{mmNL}	NLOS mmw channel path-loss exponent
λ	Non-negative distribution intensity or rate
μ	Mean value
σ	Standard deviation
$\|\Delta\|$	Population/ Swarm size
ρ	Probability
γ	Signal-to-interference-and-noise-ratio
ϑ	Data rate achieved by user per PRB at given SINR
ϕ	Data rate demand of a user
η	Load
ψ	Load fairness index
χ	Bandwidth share
η_{PA}	Efficiency of power amplifier
$\mathscr{A}_{mmw}$	The mmw RRHs association probability
$\mathscr{A}_{\mu\text{w}}$	The MBS association probability
$\mathscr{C}(\mathscr{T}_c)$	The coverage probability
$\mathscr{C}_{HC-RAN}$	HC-RAN coverage probability
$\mathscr{C}_{mmw}$	The coverage probability of mmw RRHs
$\mathscr{C}_{\mu\text{w}}$	The coverage probability of MBS
$\mathscr{R}(\mathscr{T}_r)$	The rate coverage probability

$\mathscr{R}_{HC-RAN}$	HC-RAN rate coverage probability
$\mathscr{R}_{R_{mmw}}$	Rate coverage probability of mmw RRHs
$\mathscr{R}_{M_{\mu w}}$	Rate coverage probability of MBS
$\mathscr{T}_r$	Rate threshold
$\mathscr{T}_C$	SINR threshold

Symbols

A	Path loss constant that depends on the carrier frequency f_c
B_j	Biasing factor for user association
C	Number of cells
D_{ij}	Distance between RRH i and RRH j
$\mathsf{D}_{k_{in}}$	Distance between user k and RRH n in cell i
$\mathsf{D}_{k_{m_0}}$	The distance between user k and the serving MBS m in cell i
F	Objective function or fitness function
f_c	Carrier frequency
$\mathsf{g}^*_{k_{r_i}}$	Small scale fading channel between user k and its serving mmw RRH r_i
gbest^I	Global best position experienced by any particle upto I^{th} iteration
G	The directivity gain of the transmission uniform linear array antenna
$\mathsf{H}_{k_{in}}$	Composite fading channel between user k and RRH n of cell i
$\mathsf{h}^*_{k_{in}}$	Small scale fading channel between the RRH n and user k in cell i
$\mathsf{h}^*_{k_{m_i}}$	Small scale fading channel between the user k and MBS m_i.
$\mathsf{H}_{k_{mi}}$	The channel gain between the serving MBS m_i and user k

$I_{\max}$	Maximum number of iterations/generations
$l_{k_{m_i}}$	Large scale fading channel between the user k and MBS m_i.
$l_{k_{in}}$	Large scale fading channel between the RRH n and user k in cell i
$l_{k_{r_i}}$	The small and large scale fading channel, respectively, between the user k and its serving RRH r_i
K	Total number of users in the network
$l_{k_{r_i}}$	Large scale fading channel between the user k and its serving RRH r_i
$\mu\mathsf{W}$	Micro wave
N_0	The power of Additive White Gaussian Noise
$\mathsf{N}_{\mathrm{RB}}^{k}$	Total PRB required by user according to SINR and data rate requirement
$\boldsymbol{P}$	Transition matrix
P_c	Cross-over probability
P_m	Mutation probability
$\mathsf{P}_{k_{mi}}$	The transmit power between the serving MBS m_i and user k
P_{r_i}	The transmit power between serving the mmw RRH r_i and the user k
P_s	Selection probability.
pbest_j^I	Best position of particle j up to I^{th} iteration
P_{BW}	Bandwidth of a single physical resource block
P_{RB}	number of PRBs
R^I	Population/ Swarm Matrix at I^{th} iteration/generation
$\mathsf{S1}$	Data interface

$\mathbf{r}_j^I$	Particle/Chromosome solution vector positioned at index j at I^{th} iteration/generation
X2	control interface
U_i	Number of connected users in RRH i coverage area
Z_m	List of RRH served by BBU m

List of Acronyms

Acronyms

3GPP	3rd Generation Partnership Project
3G	The Third-Generation
4G	The Fourth-Generation
5G	The Fifth-Generation
5GPPP	5G Infrastructure Public Private Partnership
ADC	Analogue-to-Digital Conversion
AI	Artificial Intelligence
BBU	Base Band Unit
BS	Base Station
C-RAN	Cloud Radio Access Network
CAPEX	Capital Expenditure
CoMP	Coordinated Multi-Point transmission
CCO	Coverage and Capacity Optimisation
CDI	Cell Differentiation and Integration

CQI	Channel Quality Information
CT	Communication Technology
CRE	Cell Range Expansion
DAC	Digital-to-Analogue Conversion
DPSO	Discrete Particle Swarm Optimisation
eICIC	enhanced Inter-cell Interference Coordination
eMBB	enhanced Mobile BroadBand
EA	Evolutionary Algorithms
EPC	Evolved Packet Core
ES	Exhaustive Search
FFT	Fast Fourier Transform
GA	Genetic Algorithm
GBR	Guaranteed Bit Rate
GOPS	Giga Operations Per Second
HetNets	Heterogeneous Networks
HC-RAN	Heterogeneous Cloud Radio Access Network
ICIC	Inter-cell Interference coordination
IoT	Internet of Things
ITU	International Telecommunication Union
IMT	International Mobile Telecommunications
JT	Joint Transmission

KPI	Key Performance Indicator
LOS	Line of Sight
NLOS	Non-Line of Sight
LTE	Long Term Evolution
LTE-A	Long Term Evolution-Advanced
mmw	Millimetre Wave
mMTC	massive Machine-Type Communications
MAC	Medium Access Control
MBS	Macro Base Station
MIMO	Multiple-Input-Multiple-Output
MNO	Mobile Network Operators
NFV	Network Function Virtualisation
NR	New Radio
NGMN	Next Generation of Mobile Networks
NSA	Non-Stand Alone
OAM	Operations,Administration, and Maintenance
OFDM	Orthogonal Frequency Division Multiple Access
OPEX	Operational Expenditure
OTN	Optical Transport Network
PA	Power Amplifier
PHY	Physical layer

PRB	Physical Resource Block
PF	Proportional Fair
PSO	Particle Swarm Optimisation
RF	Radio Frequency
RAN	Radio Access Network
RR	Round Robin
RRC	Radio Resource Control
RRH	Remote Radio Head
ROC	Rank Order Centroid
SA	Stand Alone
SDN	Software-Defined Networks
SFR	Soft Frequency Reuse
SINR	Signal to Interference and Noise Ratio
SON	Self-Organising Network
QoS	Quality of Service
UE	User Equipment
URLLC	Ultra-Reliable Low Latency Communications

List of figures

List of tables

Chapter 1

Introduction

The Cloud Radio Access Network (C-RAN) architecture is a key enabler for meeting the Fifth-Generation (5G) requirements. It refers to the virtualisation of Base Station (BS) functionalities using cloud computing. This architecture can increase statistical multiplexing gain, capacity and energy efficiency, decrease the operation and maintenance cost as well as adapting to traffic fluctuation. For this thesis, improving the integration of the vital Self-Organising Network (SON) features in C-RAN to consolidate the cost efficiency of a 5G cellular network is considered. Specifically, addressing and analysing techniques for capacity scaling, load balancing and power saving are covered. This chapter has begun with a statement of the research challenges, next, the motivation behind this research is explained, with the thesis contributions being subsequently outlined.

1.1 Motivation

During the last few years there has been a tremendous increase in the number of mobile connected devices, especially smartphones. It is predicted that the number of connected devices will surpass the projected world population, reaching 12.3 billion devices by 2022. Consequently, this has led to the realisation of Internet of Things (IoT), which will solve many issues and will allow for the creating new industries. Moreover, the emerging data ravenous applications are producing an exponential increase in global mobile data traffic and it is expected to continue increasing, reaching seven fold in 2022 what it was in 2017 as shown in Fig. 1.1 [1]. This will push the existing system capacity beyond its limit.

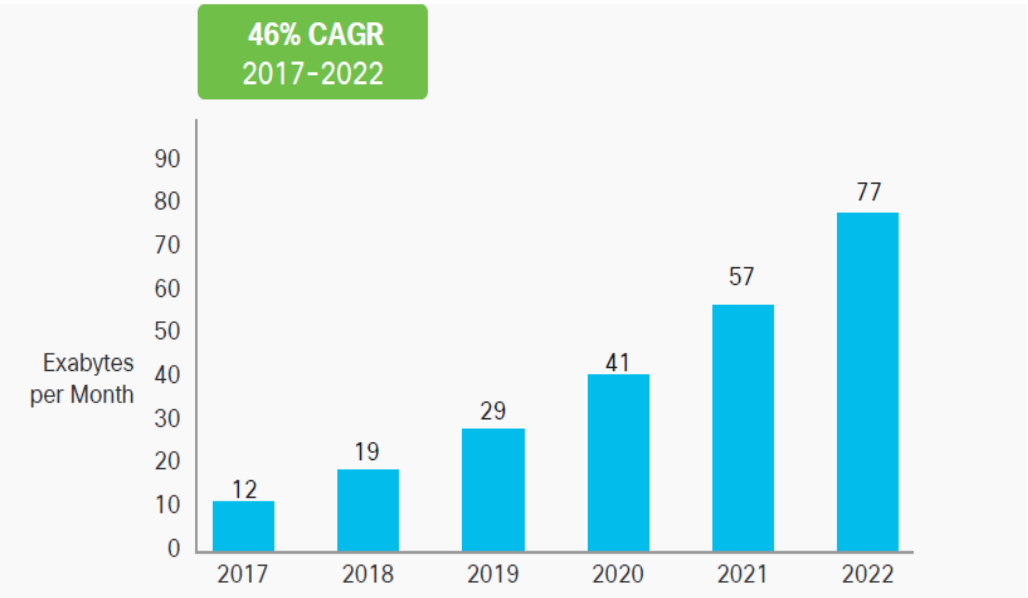


Fig. 1.1 Projected Global mobile data traffic [1].

Long Term Evolution-Advanced (LTE-A) is considered to be a wireless technology that can meet high data rate and vast capacity requirements. However, 5G requires even higher capacity values plus low latency and high reliability for specific use cases, such as virtual reality, massive IoT and enhanced vehicle to everything communication etc. 3rd Generation Partnership Project (3GPP) defines three main features for 5G, according to usage scenarios of International Mobile Telecommunications (IMT-2020) and beyond, as enhanced Mobile Broad Band (eMBB), massive Machine-Type Communications (mMTC) and Ultra-Reliable Low Latency Communications (URLLC) [2]. These requirements will push the existing system capacity beyond its limit. According to Shanon's equation, there are two approaches to dealing with the capacity deficiency of conventional networks. These approaches involve: increasing the available bandwidth efficiency as well as spectrum extension. In the first context, network densification by deploying a large number of low power (femto or pico) small cells is a promising approach. However, it leads to dramatic increases in energy consumption and results in increasing Capital (CAPEX) and Operational (OPEX) Expenditures due to additional deployment and maintenance. Moreover, cell size reduction induces increased handovers, which significantly raises the signalling and processing overhead associated with handover operations. Also, another proposed approach is through bandwidth extension by exploiting Millimetre Wave (mmw) frequencies. The mmw band has a wide bandwidth, which spans from 30-300 GHz. An mmw's cellular communication differs from conventional mobile sys-

tems in being short-range and intermittent, because these frequencies have high path losses and are vulnerable to blockage owing to shadowing by obstacles. So, directional communication using a high gain multiple antenna systems is used to compensate for the high path losses. As the mmw cell coverage is not continuous due to blockage vulnerability, then the User Equipment (UE) is required to find another unblocked channel or run a cell search all the time to find another cell and has to perform frequent handovers, even for stationary users. Moreover, mobility management will be more complicated. To avoid regular handovers and delay in re association with another small cell, the UE may adopt the reception of multi-beams from different directions. To this end, a centralised architecture C-RAN is vital to overcome the aforementioned issues by migrating the base band processing unit of multiple BSs to a Base Band Units (BBU) pool for centralised cooperation and coordination, while leaving the Remote Radio Heads (RRHs) at the cell sites with simple functionalities that make dense deployment feasible with minimum cost. Generally, the high capacity requirement of 5G application will bring high fluctuations in traffic load due to the variation of mobile users' spatial distribution and their data demand over time. This load fluctuation leads to inefficient resource utilisation that degrades the overall Quality of Service (QoS).

Hence, the motivation for this thesis is the need to integrate the self-organising feature of SON with the centralised capacity management characteristic of C-RAN, with the aim of higher QoS by exploiting re-configurable C-RAN architecture to realise energy-efficient operation for 5G networks.

1.2 Problem Statement

Centralised resource allocation, cooperative communication, clustering and spectrum reconfiguration are challenging research areas in C-RAN. Some of the research problems addressed by this thesis are outlined in this section.

- In the light of centralised cooperation and coordination that C-RAN supports, resource sharing and scheduling among RRHs of the same BBU pool must be neighbouring in a compact cluster. This cooperation has not been extensively studied, which gives a problem to solve regarding high capacity demand and the appropriate association of the BBU to UE.

- Network densification (which is reusing the available spectrum by deploying a large number of small cells to cover a certain geographical area or filling the coverage gaps) is considered a promising solution for increasing spectral efficiency to meet the enormous capacity demand for future networks. However, this approach imposes several challenges, for example, a growing number of handovers and increasing interference. Therefore, the frequent handover and power consumption due to the increase in BS number are essential problems to consider for future networks.

- In C-RAN architecture, BBU-RRH mapping can be modified according to the network capacity demand. This re-configurable BBU-RRH association leads to an increased number of handovers. Hence, optimal clustering of the RRHs is required, which means assigning them to the proper BBUs in the pool. Besides this, achieving high multiplexing gain without sacrificing QoS is another research problem to consider. Furthermore, this BBU-RRH association can be modified to realise power saving with minimum handovers, whilst simultaneously keeping a high QoS level is a problem to be solved.

1.3 Aim and Objectives

The primary aim of this thesis is to introduce a self-organising Cloud Radio Access Network in the light of 5G vision, which can dynamically and efficiently adapt to the varying capacity demand, while maintaining high QoS by adopting power saving, resource allocation and load balancing techniques. The main objectives to be met to fulfil this research aim are the following.

1. Surveying the literature to review the recent studies in centralised resource management as well as self-optimising SON techniques and Artificial Intelligence (AI) techniques employed in the cellular system to address coverage and load balancing problems.

2. Designing an intelligent power efficient technique for self-optimising C-RAN that imitates the stem cell division (mitosis, meiosis) in biological science. This scheme allows for efficient utilisation of the BBU resources in the cloud and proper deployment of the RRHs in the cell sites.

3. Selecting the appropriate Key Performance Indicators (KPIs) for measuring network performance.

4. Modelling the QoS as a C-RAN performance metric, which is presented by the selected KPIs and formulating an optimisation problem to maximise the network performance modelled by the performance metric of QoS.

5. Finding the optimum values of the KPIs for different network scenarios by applying the optimisation algorithm to the modelled network performance metric. i.e. QoS.

6. Formulating a power model for estimating the consumed power at a given load of the self-optimised C-RAN architecture.

1.4 Thesis Contributions

This thesis is an amalgam of several research disciplines. It integrates different notions to realise enhanced self-optimising network performance. In addition, AI is the basis of self-organising and machine learning network technologies. Since mMTC is one of the main 5G usage scenarios, consequently AI can be applied in next-generation wireless network standardisation.

First, a Cell Differentiation and Integration (CDI) algorithm that employs AI and the self-optimising feature of SON in C-RAN architecture to establish a cellular network that responds to the traffic load fluctuations and variable capacity demand is proposed.

The algorithm is inspired by cells splitting in biological sciences (mitosis, meiosis, fertilisation), where the stem cells are differentiated into cells of the same type or different type depending on their function. A C-RAN architecture is investigated, particularly focusing on the allocation of BBU cloud resources between the RRHs. A multi-objective optimisation problem is modelled concerning the maximisation of the QoS function, which is the proposed performance metric. The Genetic Algorithm (GA) and Swarm Intelligence are chosen to solve the optimisation problem among the various AI techniques as it is the most adopted, and it used in Long Term Evolution (LTE) self-optimising networks. Furthermore, a Discrete Particle Swarm Optimisation (DPSO) is also developed as an Evolutionary Algorithm (EA) to solve the optimisation problem and compared to GA as well as to an Exhaustive Search (ES) algorithm. The DPSO is found to be faster and less complex than GA and EA. It reaches the optimum solution for small networks and near optimum for large networks.

Second, the mmw frequency band is one of substantial features of 5G wireless access technology New Radio (NR). For this reason, the CDI algorithm performance is tested for small cells operation using one of the NR Frequency Range2 (FR2 bands , which are the mmw bands for 5G cellular access that spans 24.25GHz-52.6GHz [3]) for mobile access. A multi-objective optimisation problem is developed to maximise the network QoS with objective functions that match the mmw band operation. GA and DPSO algorithms are developed to solve the optimisation problem, and the results of the two algorithms are

analysed and compared to the ES results. Furthermore, the results of mmw-enabled CDI are compared to the CDI working at the Micro wave (μW) band for the same objective functions.

Third, 5G differs from the previous generations in being backwards compatible with LTE and WIFI, plus in its early deployment it is going to be integrated with LTE infrastructure in a non stand-alone operation mode. Furthermore, as eMBB is one of the main 5G usage scenarios, a scheme for exploiting the operation at the mmw band as NR FR2 to increase the data rate in conventional macrocell area is proposed. The mmw communication features are high directionality and noise-limited in some scenarios. Since the macrocell's edge area suffers from low data rate of one tenth of peak data rate due to high interference [4], mmw cells serve this area to improve the users' throughput. The proposed scheme is inspired by the soft frequency reuse technique, where the cell area is divided into an inner region and outer region. Each area is served with a different frequency band. The proposed scheme increases the spectral efficiency, coverage and rate of not only the edge area, but also the whole cell area.

1.5 Thesis Layout

The rest of this thesis is organised into five chapters. Each chapter starts by briefly introducing the contribution made in the chapter and finishes with a concise summary at the end.

Chapter 2 This chapter introduces a brief background and general overview of the main subject considered in this thesis. First, SON self-optimising techniques are reviewed and a concise introduction to C-RAN is also presented. After that, 5G operation scenarios are reviewed, the enabling technologies are considered. In addition, the 5G wireless access technology NR is briefly covered.

Chapter 3 In this chapter, an intelligent CDI algorithm for cell differentiation and integration is proposed, which benefits from the capacity routing feature of C-RAN to adapt to load fluctuation with time. The CDI algorithm is inspired from the biological process of cell splitting (mitosis and meiosis), where the stem cell splits into multiple cells

of the same or different types of cells. The intelligent CDI algorithm imitates this natural process by activating or deactivating the RRHs and BBUs in the network depending on the traffic demand at the end of each CDI cycle. An optimisation problem is formulated to reallocate the RRHs to the BBUs and balance the load, with two EAs GA and DPSO being developed to solve the problem. The system model and the algorithm performance are investigated, the results and algorithm complexity are analysed, and a conclusion is drawn. Furthermore, a power model is developed to estimate the power consumed in C-RAN and compared to CDI adopted C-RAN. The results of the power and energy efficiency, as well as the capacity results, are discussed and analysed.

Chapter 4 This chapter investigates the potential benefits of incorporating an mmw frequency band operation scenario for the small cells in the proposed CDI-enabled C-RAN. An optimisation problem with objective functions and constraints suitable for adopting high-frequency operation for small cells is formulated to increase the system QoS. The network throughput, blocked users and energy efficiency results of the proposed system are analysed and compared to the scenario of a lower frequency band operation.

Chapter 5 This chapter introduces a scheme to exploit one of the 5G enabling technologies, namely the mmw band in the early 5G deployment scenario of Non-Stand Alone (NSA). This scheme is based on the Soft Frequency Reuse (SFR) technique used in LTE for interference mitigation, but it targets the whole cell area, thereby enhancing the coverage, spectral efficiency and data rate. The obtained results of coverage and throughput are discussed and analysed at the end of this chapter.

Chapter 6 In this final chapter, the thesis findings are summarised and conclusions are drawn. Also, suggestions for future research directions are proposed, with some potential ideas being highlighted.

Chapter 2

Background

2.1 Introduction

As the primary motivation of this work is cellular system capacity enhancement due to the exponentially increasing traffic demand as shown in Fig. 1.1. Hence, according to Shannon's equation ($Capactiy = Bandwidth * log_{10}(1 + SINR)$), there are two approaches to facing the future capacity deficiency of conventional networks due to mobile data traffic growth in cellular networks, which involve: (i) increasing the available bandwidth efficiency and (ii) spectrum extension. The most promising techniques in increasing bandwidth efficiency are network densification and Multiple-Input-Multiple-Output (MIMO) technology that helps to manage interference as well as increasing spectral efficiency [5].

In the first context of capacity increasing, a large number of low power (femto or pico) small cells are expected to be deployed to fill the coverage gaps and increase the rate in the serving area of the macrocells. This brings in Heterogeneous Networks (HetNets), which have been standardised by a 3rd Generation Partnership Project (3GPP) [6]. However, large scale HetNets are facing several challenges, such as network management, interference, cost of small cells deployment and energy consumption. To deal with network densification issues, the Cloud Radio Access Network (C-RAN) is a promising solution. In this architecture, the small cells' baseband processing units are migrated, aggregated and shared among sites in a centralised virtualised Base Band Unit Pool (BBU). The Radio Frequency (RF) units and simple processing left at the cells sites, called Remote Radio Head (RRH), are interfaced with the BBU pool through the fronthaul. There are

several potential realisations of the fronthaul link between the BBU pool and RRHs depending on the cost, geographical area and the required Quality of Service (QoS). It could be wired, wireless or a combination of both wired and wireless. Furthermore, the fronthaul should offer high capacity and low latency to support the Fifth-Generation (5G) requirement of Ultra-Reliable Low Latency Communications (URLLC) [7].

A centralised cloud cooperative platform is also introduced to HetNet, where all the small cells as well as the macro base stations, are connected to the cloud for control and management, which gives rise to Heterogeneous Cloud Radio Access Networks (HC-RAN). For centralised cloud computing-based collaboration with the RRHs, high power traditional macro base stations are interfaced with the BBU pool through the data and control interfaces S1 and X2, respectively. HC-RAN is proposed to take full advantage of both HetNet and C-RAN [8–11].

The 5G era can be considered as moving to high operating frequencies and smaller cell size, which means a larger number of cells for the same geographical area. Furthermore, 5G, in its early deployment, will coexist and be integrated with Long Term Evolution-Advanced (LTE-A) radio access technology, which mostly uses a unity frequency reuse. Hence, inter-cell interference is high in such systems, especially for dense cell deployment. There are different techniques to address interference, such as Inter-Cell Interference Coordination (ICIC), enhanced Inter-cell Interference Coordination (eICIC) and Coordinated Multi-Point(CoMP). ICIC and eICIC [12, 13] reduce the interference at the cell edge, but they do not improve the throughput. This is due to the strict usage of resources in both the frequency domain (ICIC) and time-domain (eICIC). The other drawback is that the interference information is exchanged among the neighbour cells on a relatively long term basis. This delays the dynamic resource allocation for fast-changing channel conditions users, for example, User Equipment (UE) travelling fast or stepping into a shadowing area, as their information is not promptly reflected into cells cooperation. Moreover, the scheduling decisions that the scheduler makes would be less optimal, as it has to consider neighbour cells' interference. The other interference mitigation technique, which is the most advanced, is CoMP [14, 15]. This approach can

both reduce interference and increase the throughput of cell edge users, whilst also being able to respond to fast channel changes [16–18]. From the architecture point of view, the Joint Transmission (JT) algorithms of CoMP and eICIC are not efficient due to the high latency and low bandwidth of the X2 interface of the existing LTE-A. Even with C-RAN architecture, these techniques require more detailed instantaneous channel status signals, which leads to the need for ideal fronthaul in some scenarios and increased processing (more joules per bit).

Furthermore, energy saving is a crucial issue to consider as it increases the Operational Expenditure (OPEX). Moreover, the projected energy consumption in the Communication Technology (CT) sector in 2030 could be 51% of the global energy consumption as shown in Fig. 2.1, which is expected to contribute to 23% of the global greenhouse gas emissions [19], as shown in Fig. 2.2. For these reasons, Mobile Network Operators (MNOs) are interested in increasing capacity without an immense rise in consumed power.

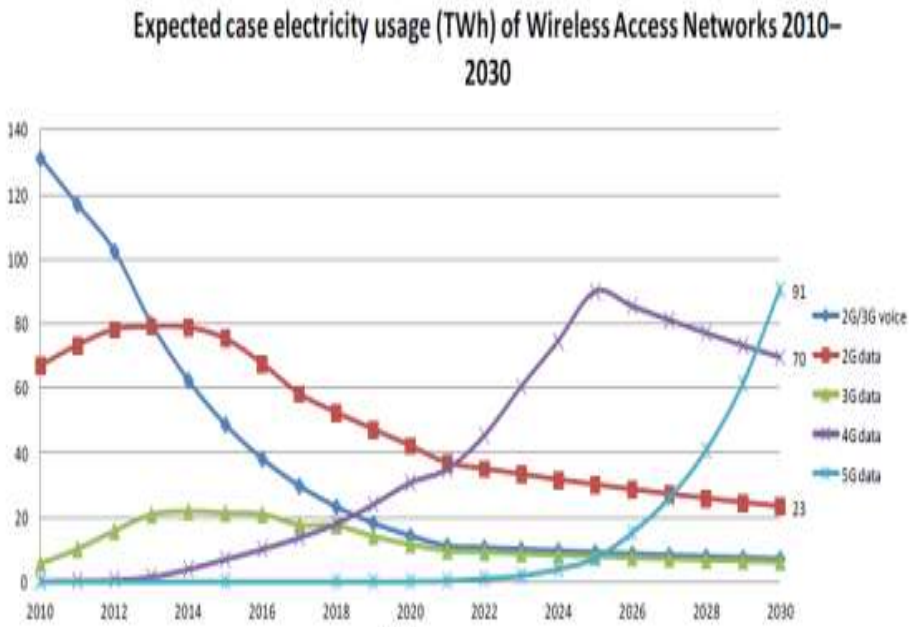


Fig. 2.1 CT projected wireless power consumption [19].

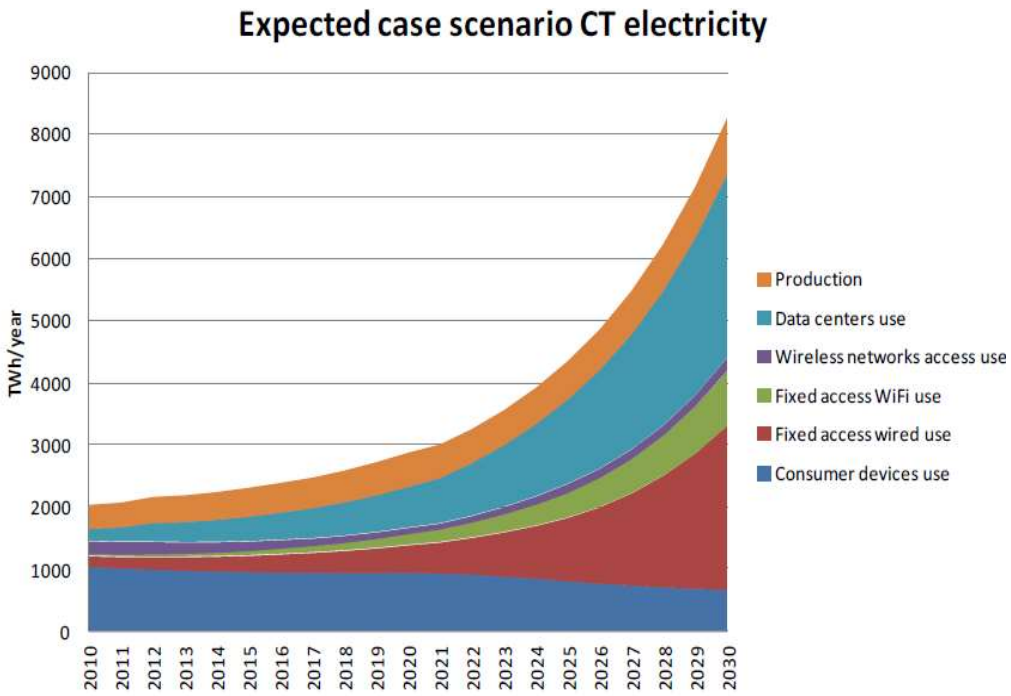


Fig. 2.2 CT projected power consumption [19].

Moreover, the operation complexity and OPEX could be the biggest challenges that face MNOs in 5G. To cope with similar challenges in the context of The Third-Generation (3G) and The Fourth-Generation (4G) networks, the 3GPP introduced the Self-Organising Network (SON) [20]. The Alliance of Next Generation Mobile Networks (NGMN) introduced various SON requirements and use cases [21]. The essential SON objectives can be divided into:

1) Adding autonomous, intelligence and adaptability feature into cellular networks;

2) Decreasing the Capital and Operation Expenditures (CAPEX& OPEX);

3) Enhancing the network performances regarding capacity, coverage, quality of service/experience, etc.

Furthermore, as the Radio Access Networks (RAN) transforms from distributed to centralised cloud-based architecture, nowadays SON can be conceived as the next generation RAN driving technology. It is aimed at simplifying management, enhancing spectral efficiency and lowering operation costs [22, 23]. The SON challenges have

been classified into useful use cases, that have been addressed by 3GPP, NGMN, 5G Infrastructure Public Private Partnership (5GPPP) and different EU projects [21]. For example, for project Self-Optimisation and Self-Configuration in Wireless Networks (SOCRATES) [24] methods, algorithms, coverage and load balancing in the Long Term Evolution (LTE) network were studied. SON is a critical enabler for improving Operations, Administration, and Maintenance (OAM) of the network. It targets cost reductions in installation and management of 4G and future 5G networks, through self -configuration, self-optimising and self-healing, as shown in Fig. 2.3 [25]. This thesis is particularly focused on self-optimisation regarding network Coverage and Capacity Optimisation (CCO) as well as load balancing.

2.2 Self-Organised Networks (SON)

The self-organisation approach, which is referred to as SON, is applied in cellular networks to maximise total performance. This is by bringing intelligence and autonomous adaptability features to mobile networks. Furthermore, it is a vital enabler for improving OAM functions. SON targets the reduction of installation and management cost by simplifying the tasks of operation using its capabilities to configure, optimise and heal itself. In short, SON's objective is to reduce human intervention, thus eliminating human error in network operation. Moreover, it enhances network capacity, coverage and QoS, while reducing CAPEX and OPEX at the same time [23].

2.2.1 SON Functions

SON involves a series of functionalities, which can be applied in several network parts aiming at various targets. These functionalities can be categorised into the following three parts [23, 26].

- **Self-configuration:** it is the network's ability to bring a new component into service with minimal human operator intervention. It also means applying self-configuration algorithms to configure new added, deleted, or modified nodes automatically;

• **Self-optimization:** this is the periodic process of dynamically and automatically tuning network parameters for optimal performance. It involves a set of mechanisms that optimise the network parameters during operation and executing optimisation algorithms based on the network measurements;

•**Self-healing:** it is a function that focuses on cellular network maintenance since it is prone to faults. Self-healing is responsible for automatic compensation of node failures to restore the degraded service.

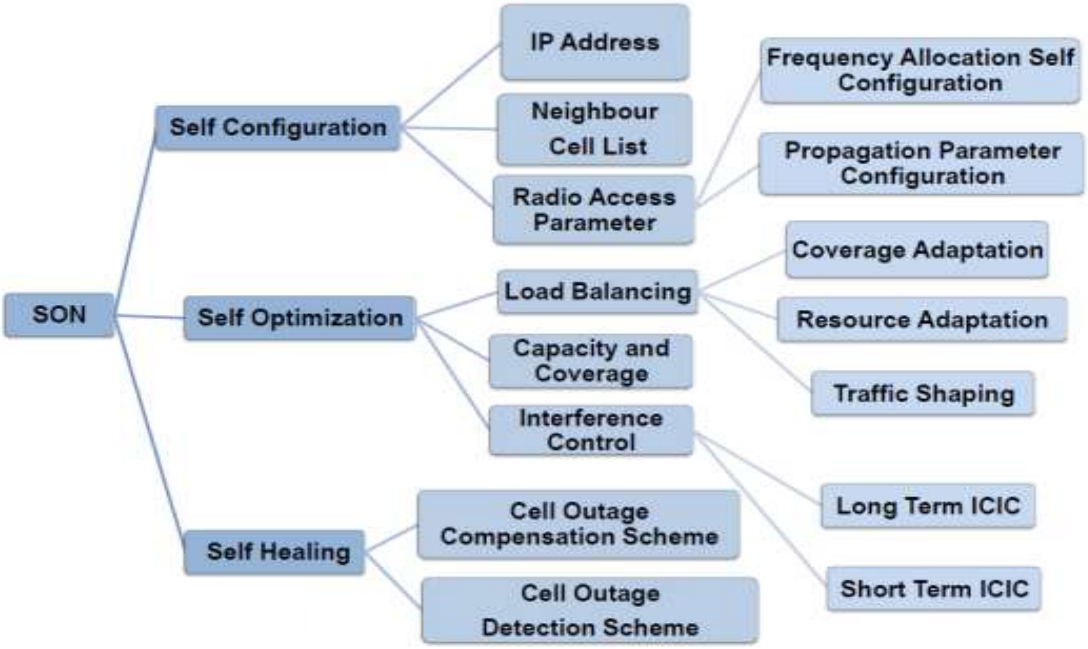


Fig. 2.3 Self organising network tends [18].

2.2.2 SON Implementation

SON implementation can be categorised into three types of architecture as follows [27, 23, 26]:

Centralised SON: Centralised SON, where the optimisation algorithm is established and executed in the management part of the network. It takes data and measurements from all the elements and nodes in the network, subsequently sending back updated, orders and parameters, according to the collected data. This architecture is more convenient for algorithm implementation since it considers the collected data

from all the network, which allows for a fully tuned system by optimising all the centralised functions. This architecture adapts slowly to changes in the network. However, it is more effective with network instabilities coming from the concurrent operation of SON different functions that have conflicting goals. In addition, the SON functions control is performed centrally; these different functions can be easily coordinated.

Distributed SON: Distributed SON pertains to where the SON algorithms run on the network parts. All self-organising data is exchanged directly between the parts of the network, which allows the SON functions to be more dynamic in adaptation to changes. However, the global optimum operations cannot be guaranteed for the whole network as all optimisation functions run at the cell level.

Hybrid SON: In hybrid SON architecture, the SON algorithms are partly executed on the management system of the network and partly on the other elements. This architecture integrates the advantages of both centralised and distributed SON architectures. However, it has also inherited the drawbacks of both architectures. Furthermore, the interface between the two architectures of SON functions is complex and challenging. The 5G SON implementation scenario is going to be a hybrid architecture [28].

2.3 Self-Organised Network for 5G

The primary objective of 5G SON is to improve the efficiency of OAM greatly and to help operators meet the increasing complexity of rapidly evolving wireless networks. Moreover, 5G SON should enhance network performance by using self-organising intelligence. Some of the 5G SON objectives are listed below.

Enabling Cloud RAN Currently, the RAN is transformed from distributed to cloud based architecture, in which most of the RAN functions are being centralised, while still leaving some of them at remote radio heads. Specifically, the functions with low latency requirement (seconds or higher) will be run in a centralised fashion in

the core network while those with high latency requirements (millisecond order) will still be run at the BBU cloud [25, 28]. Furthermore, 5G is expected to serve a massive number of connected devices, embrace other radio access technologies and deal with an unlimited number of new applications. All of this will increase network complexity and load. Consequently, this will bring high load fluctuations to the network. 5G SON combined with the new C-RAN architecture will address this challenge through automatic load monitoring and load prediction based on service statistics history. Then, the SON algorithm should respond by scaling the capacity and signalling, according to the load condition, so as to ensure optimal network performance [25, 28].

Improving OAM Efficiency The 5G vision is greatly enhanced users' experience without increasing the service costs. This enhanced users' experience in terms of connectivity and data rate coming from increasing the number of cells (network densification), antennas (using MIMO) and carriers (using Millimetre Wave (mmw) band). Furthermore, with the coexistence of multiple technologies, the heterogeneous nature of the architecture and different applications will contribute to network complexity escalation. This means higher CAPEX and OPEX, with the operators' revenues growth being insufficient to justify the increased network complexity. Accordingly, high efficiency OAM is crucial for 5G, especially the self-optimisation phase of SON. Furthermore, a proactive SON is an appealing approach for highly efficient OAM [25, 26, 28, 29].

Improving Energy Efficiency The network densification is one of 5G's primary themes and energy efficiency is one of its Key Performance Indicators (KPIs) [28]. However, it is not efficient from an energy and OPEX point of view that this massive number of cells to be always on. Instead, an effective strategy for most 5G cells could be of switching on when needed. This could be realised by embedding a centralised SON algorithm with an energy efficiency objective. The challenges that require addressing are: the compromise between the energy efficiency function

and the other SON functions, such as latency, signalling overhead, and the latency margin of cells coming back to active mode [22].

2.4 Centralised Radio Access Network (C-RAN) For 5G

The C-RAN in industry can be backdated to 2010, when it was proposed by IBM for more flexible and energy efficient networks[30]. After that, in 2011, the concept was utilised by the China Mobile Research Institute and the architecture was developed as well as the feasibility and challenges were explained [31]. C-RAN is a paradigm shift in the architecture of the next generation of wireless networks. It is based on the concepts of centralisation and virtualisation, where the base band processing functions are virtualised and moved to a remote centralised BBU pool and only RRHs are left at the cells' sites, these two entities being linked by the fronthaul, as shown in Fig. 2.4 [32]. The resource centralisation leads to statistical multiplexing gains, cost reduction and resource savings. This architecture is considered as one of the key solutions to meet the massive demand of capacity in future 5G wireless access networks [33].

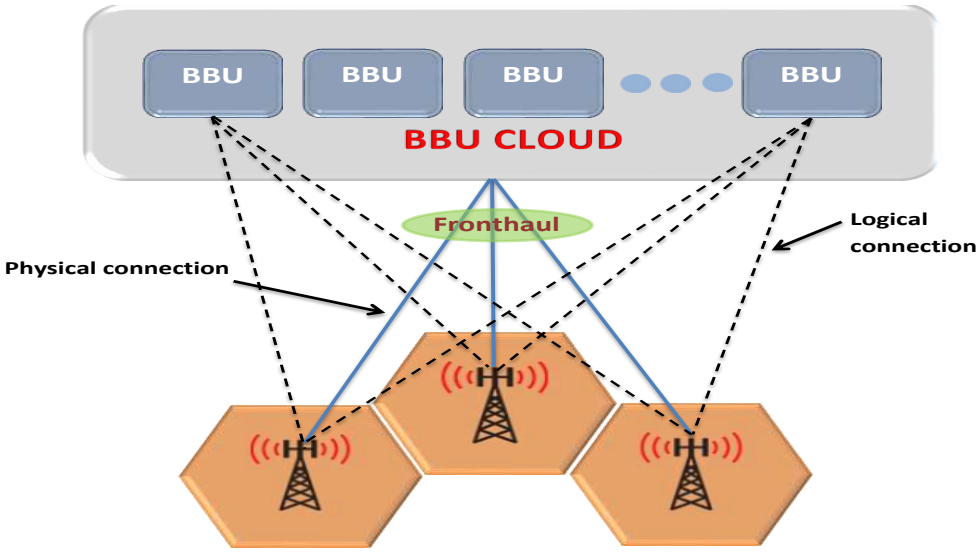


Fig. 2.4 C-RAN architecture.

There are two realisation scenarios of the C-RAN architecture depending on the functions splinting between RRH and BBU, which are fully centralised and partially

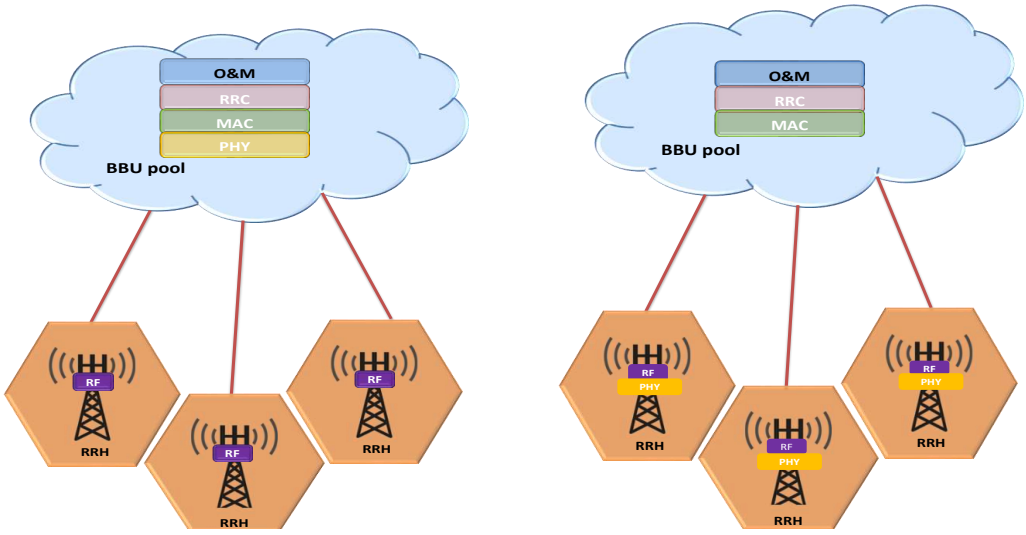
centralised as shown in Fig. 2.5. This function splitting depends on network requirements to alleviate the fronthaul high capacity overhead burden.

Fully centralised: With a fully centralised structure as shown in Fig. 2.5a, all the functions of the traditional base station would be moved to BBU pool, including the: i) Physical layer (PHY) functionalities; ii)Medium Access Control (MAC) [34] and iii) Radio Resource Control (RRC) [35]. Whilst the RRHs left with simple functions represent Analogue-to-Digital (ADC) and Digital-to-Analogue (DAC) conversions, power amplification, and filtering [36]. Accordingly, the BBUs are responsible for all the functions of managing and processing resources, which means that the C-RAN structure has significant merits of easy operation and maintenance. Moreover, the fully centralised implementation of C-RAN will bring to 5G cellular networks the features of scalability, enhanced resource sharing, ease of system upgrading, multi-standard operation and processing collaboration of multi cells. However, the fully centralised performance will be challenged by the capabilities of the fronthaul bandwidth overhead and the baseband signal transmission [37].

Partially centralised: Regarding partially centralised C-RAN, as shown in Fig. 2.5b physical layer functions, such as baseband processing and radio functions are located and performed at the RRHs, while other higher layers functions are moved into the centralised BBU. In this realisation of C-RAN, fronthaul bandwidth capacity requirements are lower, as the baseband processing is retained at the RRH. However, it also has some drawbacks, for instance, limited flexibility of network upgrades and is problematic regarding collaborative transmission for multi-cells.

In short, both fully and partially centralised C-RAN implementations have been developed from different perspectives, all targeting 5G requirements fulfilment. Hence, adopting either of the C-RAN architectures realisation scenarios depends on the network requirements. It is worth mentioning that with C-RAN architecture, the static association between the BBUs and RRHs is loosened. The RRHs will not be connected to specific

physical BBUs. Instead, each RRH is connected to any virtual BBU in the BBU pool using the technology of real-time virtualisation [37].



(a) Fully-centralised C-RAN. (b) Partially-centralised C-RAN.

Fig. 2.5 C-RAN architecture realisation.

2.4.1 C-RAN Architecture Advantages

C-RAN will add impressive features and enhance some of those already existing in traditional architecture, as summarised below.

Spectral efficiency enhancement and interference management: C-RAN architecture will improve the spectral efficiency of wireless networks by supporting efficient collaborative transmission, reception and interference mitigation techniques. For example, CoMP and ICIC strategies are performed simply and effectively between the RRHs connected to the same cloud. This is owing to the high speed $X2$ link between the BBUs as they aggregated together inside the BBU pool, which makes the sharing of CSI, control signals and data traffic among the BBUs fast and more efficient, resulting in low interference and high spectral efficiency [38, 4].

Decreased delays: One of the most significant features of C-RAN is the capability of performing different functions with reduced latency. In particular, the handovers time delay is noticeably decreased, because it is carried out in the BBU pool cloud,

not between the base stations at different sites. Also, the handover failure rate is reduced for the same reason. Similarly, less signalling information will be forwarded to the core network, which will contribute to lower the latency [39, 40].

Adaptability to fluctuating network traffic load: The mobile network traffic has a non-uniform distribution of load, which has a varying nature during the day time for the same geographical area, referred to as the tidal-effect [4]. However, the base stations are designed for peak load, which implicitly means inefficient use of processing power at off-peak intervals [40]. As the baseband processing in C-RAN is performed centrally in the BBU pool, the BBU resources can be dynamically and optimally allocated according to the immediate traffic demand.

Lowering CAPEX & OPEX: First, in C-RAN the cost of deploying new cells is much lower than with the traditional architecture, as deploying and operating the RRHs requires less time and cost compared to the traditional system. Second, the capability of sharing hardware in C-RAN will also contribute to CAPEX reduction. An analysis study has shown that C-RAN architecture CAPEX reduction can reach about 15% per kilometre [41]. In addition, aggregating all the computing and processing resources in the BBU pool cloud can significantly save the OPEX [42].

Power saving and energy effeciency enhancing: Generally, in C-RAN architecture the number of BBUs is less than the traditional architecture, thereby lowering the level of power consumption. In addition, the power consumed for the air conditioning is 90% decreased as there is no need for it at the RRHs side [43, 44]. Similarly, in HC-RAN the macrocells can offload some of the power-consuming data processing to the BBU cloud, which is also contributes to power saving. In a study carried out by ZTE, it was concluded that C-RAN architecture consumes 80% less power compared to the conventional architecture [45]. Another study by China Mobile states that 71% power consumption can be achieved by the C-RAN architecture compared to the traditional counterpart [31].

Easiness of network maintenance and upgrading: The resource centralisation characteristic of C-RAN will make the network upgrading and maintenance easier. To exemplify, if a new service needs to be covered, the RRHs are simply deployed and connected with the cloud to cover the areas or cell splitting for higher capacity. Also, cloud resource extension in C-RAN is easily performed by installing virtual resources in the cloud. Furthermore, for introducing a new standard in C-RAN, the hardware can be easily placed at a few centralised locations. Besides this, more frequent updates are easily realised in C-RAN compared to traditional systems [16, 32].

2.4.2 C-RAN Architecture Challenges

C-RAN is the next-generation wireless network architecture [46] owing to its many great advantages set out in the previous section. However, there are some challenges that need to be addressed by researchers and operators prior to the commercial deployment. The main challenges of the literature are listed below.

Fronthaul requirements: The fronthaul link between the RRHs and BBUs in the C-RAN architecture must support high bandwidth and low latency. This is due to the In-phase and Quadrature components (IQ) signal sent between the BBUs and RRHs. There are different realisations of fronthaul, which could be an optical fibre, wireless or a combination of both. Each fronthaul realisation has its pros and cons, for instance, a fibre optic link is suitable in terms of capacity, but there is the delay constraint relating to the fibre length plus deployment inflexibility and cost [7]. Clearly, the frame processing should be less than 1ms to meet the HARQ requirement. The fronthaul can be classified into ideal without constraints, and non-ideal with constraints of latency, bandwidth, and jitter. Optical fibre communication without constraints is considered to be the ideal fronthaul for C-RANs, because it can provide a high transmission capacity at the expense of high cost and inflexible deployment. On the other hand, wireless fronthaul using μW communication technologies or millimetre wave carrier frequencies is cost

effective and flexible to deploy, although it has less capacity than optical fibre fronthaul and other constraints. Since wireless fronthaul is cheap and flexible, this technology is anticipated to be prominent in practical C-RANs [7, 16, 47].

Cooperation between BBUs and RRHs clustering: Cooperation between the BBUs in the BBU pool is crucial to support techniques such as CoMP, as the user data and channel information must be shared among the BBUs. Moreover, the inter-connection between the BBUs should have high bandwidth, high reliability and low latency. In addition, the cells clustering is inessential for achieving statistical multiplexing gain. The RRHs must be optimally clustered and assigned to BBUs. Furthermore, optimisation of active BBUs in the pool and the associated RRHs is also important for realising optimal power saving in C-RAN as well as supporting inter cell interference management and cooperating transmission techniques, such as CoMP [16]. Several studies have addressed the dynamic switching of BBU-RRH association depending on traffic demand [48–51].

Virtulisation of base station: The virtualisation of network functions is an important enabler technology, which can reduce CAPEX and OPEX by 30% [52]. The virtualisation in communication networks was first employed in the core network, which was later extended to the radio access domain, but this is still in its early stages. Moreover, the virtualisation of wireless networks is more challenging because of the real-time characteristics of the wireless communication system, for example, mobility, broadcast, attenuation, interference, time-varying channels, coverage, etc.

In general, there are three challenges that face the virtualisation of a wireless access system that needed to be addressed: first, efficient wireless resource sharing and distribution among the different operators of the virtual network. Second, the interference resulting from the utilisation of the same resources must be managed.

Third, technical and managerial issues must be considered before any wireless network virtualisation deployment. Furthermore, to enable 5G cellular network

end-to-end virtualisation, both the core network and radio access network should be virtualised. The wired network virtualisation could be performed with the core network virtualisation of a mobile network. However, with the virtualisation of RAN or the base stations, the distinctive characteristics of the wireless access network must be considered. For example, the end-users' group dynamics, mobility, and varying channel conditions need different virtualisation mechanisms from wired ones. For this reason, the virtualisation of wireless network resources will be harder for the mobile operator [37, 40].

2.5 Heterogeneous Cloud Radio Access Network (HC-RAN)

The RAN architecture of a 5G cellular network is more heterogeneous compared to the those of traditional LTE/LTE-A networks. As the capacity requirement of 5G is very high, improvements in system capacity and spectral efficiency will be needed to fulfil the requirements of end-users beyond 2020.

One of the most promising techniques to achieve this goal is deploying small cells over the coverage area of macrocells, which is called a heterogeneous network. These small cells could use the microwave band or millimetre wave band for cellular access. A new RAN architecture has been introduced called the Heterogeneous C-RAN or HC-RAN, where both the control and user planes are decoupled to improve C-RAN performance. Both heterogeneous network and C-RAN full benefits are combined in HC-RAN, resulting in enhanced spectral and energy efficiencies, while increasing the data rates.

The HC-RAN architecture involves two layers of cells. The high power macro Base Stations (BSs) cells and the low power small cells or RRH cells [53]. The macrocells function is the control of signalling and enhancing network coverage. In addition, the small cells and RRHs target is to increase the network capacity and guarantee the QoS requirements of end-users. The HC-RAN architecture is depicted in Fig. 2.6. The centralised cloud computing-based BBU pool and the decoupling of control and user

planes all lead to improved and more efficient management of HC-RAN networks [37].

To improve network performance, for example, extending network coverage and enhancing system capacity, new RRHs need to be deployed close to the user and linked to the BBU pool. HC-RAN is similar to C-RAN in that a large number of RRHs with RF and simple processing are connected to a centralised BBU pool to achieve cooperative communication and energy efficiency gain. Furthermore, the high power macro base stations are interfaced with the BBU pool via the S1 and X2 links, in order to mitigate the cross-tier interference between the RRHs and high power base stations coexisting in the same geographical area. By involving the high power base stations, the fronthaul requirements are reduced in the H-CRAN. For example, system broadcasting data and the control signalling are sent to the end-users through macro base stations, which eases the capacity and time delay burden in fronthaul. Furthermore, at low RRH traffic load, some RRHs could be switched to sleep mode to enhance energy efficiency, with the BBU pool managing all the RRHs in this mode [9, 37].

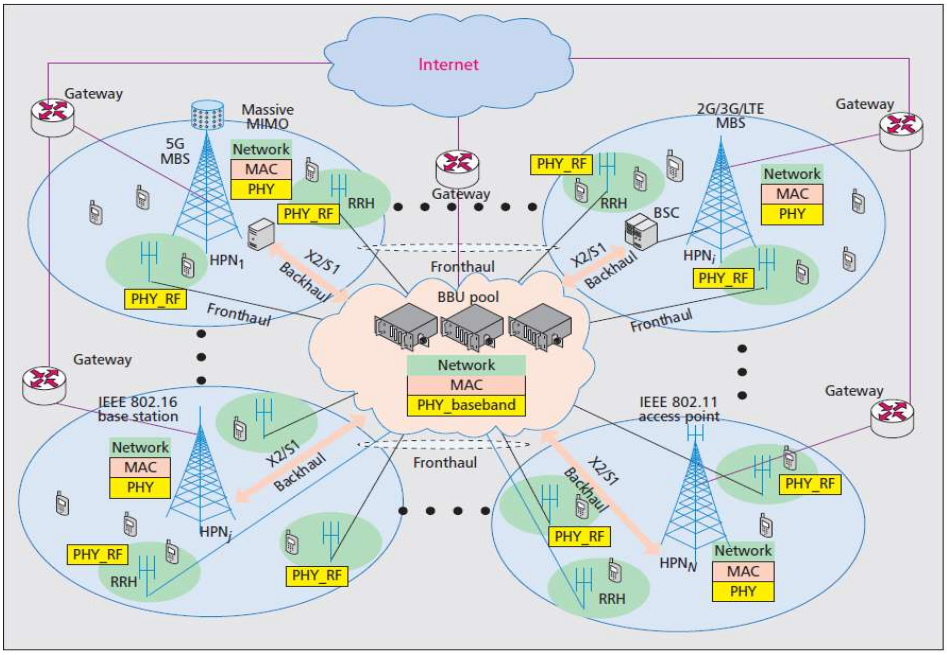


Fig. 2.6 Heterogeneous cloud radio access network architecture [9].

2.6 RRH Clustering in C-RAN

C-RAN architecture offers centralised, scalable, energy-efficient, and intelligent resource management. Traditional architecture was based on the logical one-to-one association, whereby each BBU is mapped to only one RRH. Accordingly, the BBU radio resources might be underutilised as one RRH may not utilise all of the BBU radio resources. Subsequently, the C-RAN architecture changed this fixed association between the BBUs and RRHs. One BBU can now be allocated to multiple RRHs, which means N RRHs can share the resources of a single BBU. Hence, the number of concurrently active BBUs is reduced, which typically enhances network energy efficiency[54]. Similarly, clustering is defined as determining the number of RRHs associated with one BBU at a time period. As mentioned above, this concept increases network flexibility, facilitates cooperative communication as well as interference mitigation techniques, alleviates the capacity requirement of the fronthaul and reduces BBU pool power consumption. There are several clustering techniques and algorithms in the literature targeted at finding the best C-RAN configuration by converting it into an optimisation problem so as to increase system performance. Moreover, some clustering techniques dynamically change the RRHs grouping in clusters, according to the load conditions [55]. Two of the clustering techniques are explained below.

RRHs clustering using multi-objective optimisation In this approach, the Multi-Objective Optimisation clustering ensures a trade-off and a balance between maximising or minimising the objective function, while grouping the RRHs into clusters [56].

RRHs clustering using Bin Packing approach In this approach, the bin packing problem is to pack objects into a finite number of containers or bins. The objects are the RRHs with an average demand d, which needs to be allocated to the BBUs (bins or containers) with a maximum capacity C. Hence, the demand of the RRHs allocated to the same BBU must not exceed its maximum capacity. The objective of this approach is to maximise or minimise a certain function while grouping the RRHs into clusters [57].

2.7 5G Wireless Access Technology: New Radio (NR)

The 5G wireless access technology, known as New Radio (NR), will address a variety of usage scenarios from enhanced Mobile BroadBand (eMBB) to Ultra-Reliable Low Latency Communications (URLLC) through to massive Machine-Type Communications (mMTC). NR will meet the performance requirements set by the International Telecommunication Union (ITU) for International Mobile Telecommunications for the year 2020 (IMT-2020). The 3GPP is a global standard-development organisation, which has been developing 5G NR over the past few years. There are two modes of NR operation, first, in its early deployment, it will coexist with LTE in a Non-Standalone (NSA) NR and subsequently, take the form of Stand Alone (SA) NR. The 5G requirements by each usage scenario are listed in Table. 2.1 and the KPIs for each usage scenario are explained in the following subsection.

Table 2.1 5G NR main requirements [2, 28]

Usage scenario	KPI	Target values	
		DL	UL
eMBB	Peak data rate	20 Gbps	10 Gbps
	Peak spectral efficiency	30 bps/Hz	15 bps/Hz
	Control plane latency	10 ms	
	User plane latency	4 ms	
	Average (TRxP) spectral efficiency (bps/Hz)	3 times higher than IMT-Advanced*	
	Area traffic capacity (bps/m²)	Related to average (TRxP) spectral efficiency	
	User experienced data rate (bps)	Related to 5% user spectral efficiency	
	5% user spectral efficiency (bps/Hz/user)	3 times higher than IMT-Advanced*	
	Target maximum mobility speed	500 km/h	
	Mobility interruption time (Also related to URLLC and mMTC)	0 ms	
	Network energy efficiency (Also related to URLLC and mMTC)	Required as design principle (No quantitative requirement)	
	UE energy efficiency (Also related to URLLC and mMTC)	Required as design principle (No quantitative requirement)	
	Bandwidth (Also related to URLLC and mMTC)	No requirement from 3GPP	
mMTC	Coverage	Max coupling loss 164 dB	
	UE battery life	Beyond 10 years	
	Connection density	1,000,000 device/km²	
	Latency of infrequent small packets	10 s	
URLLC	User plane latency	0.5 ms	
	Reliability	$1\text{-}10^{-5}$ success probability for 32 bytes within 1 ms user plane delay	

TRxP: Transmission Reception Point

2.7.1 KPIs of NR Usage Scenarios

The KPIs of NR usage scenarios and the required values [28, 29] are discussed below and summarised in Table. 2.1

• **eMBB KPIs:** eMBB and data hungry applications, such as video streaming, virtual reality immersive gaming are the main driver of the need for higher system capacity, better coverage, and higher data rates. In order to get experience on mobile devices as from fixed optical fibre, the 5G target improvement in the main KPIs includes three times improvement in spectral efficiency, according to the IMT-Advanced requirement.

Furthermore, in using the millimetre wave frequency band and dual connectivity the peak data rate target values are 20GHz/s for downlink and 10GHz/s for uplink, while the data rate target values are 100MHz/s for and 50MHz/s for downlink and uplink, respectively. This improvement in spectral efficiency and peak data rate is owing to the use of inter-working between the mmw and microwave frequency bands as well as enabling technologies, such as multi user-MIMO [2, 58, 59].

• **mMTC KPIs:** mMTC can also be referred to as the Narrow Band Internet of Things (NB-IoT). The 5G target value for devices connection density is 1,000,000 per km^2. Furthermore, there is to be 20dB improvement in coverage and a terminal battery life of 10-15 years [2, 58, 59].

• **URLLC KPIs:** The URLLC 5G vision is to provide data delivery with extreme reliability and very low latency. This usage scenario is targeting industry application and services with latency-critical requirements, very high reliability and security, for example, autonomous driving and tactile internet applications. The target 5G KPIs values are 99.999% success probability, 0.5ms control plane latency and –4dB signal to interference plus noise ratio [2, 58, 59]. All the 5G NR KPIs target values are summarised in Table. 2.1.

2.7.2 5G NR Deployment Stages

The legacy 3G and 4G LTE was deployed using new RAN and core network, whereas 5G NR will start with NR NSA operation mode by accommodating NR by Evolved Packet Core (EPC) as shown in Fig. 2.8. Next, after providing NR by EPC, which accommodate NSA NR, the 5G NR core network will be deployed while still utilising the LTE through different scenarios, as shown in Fig. 2.7.

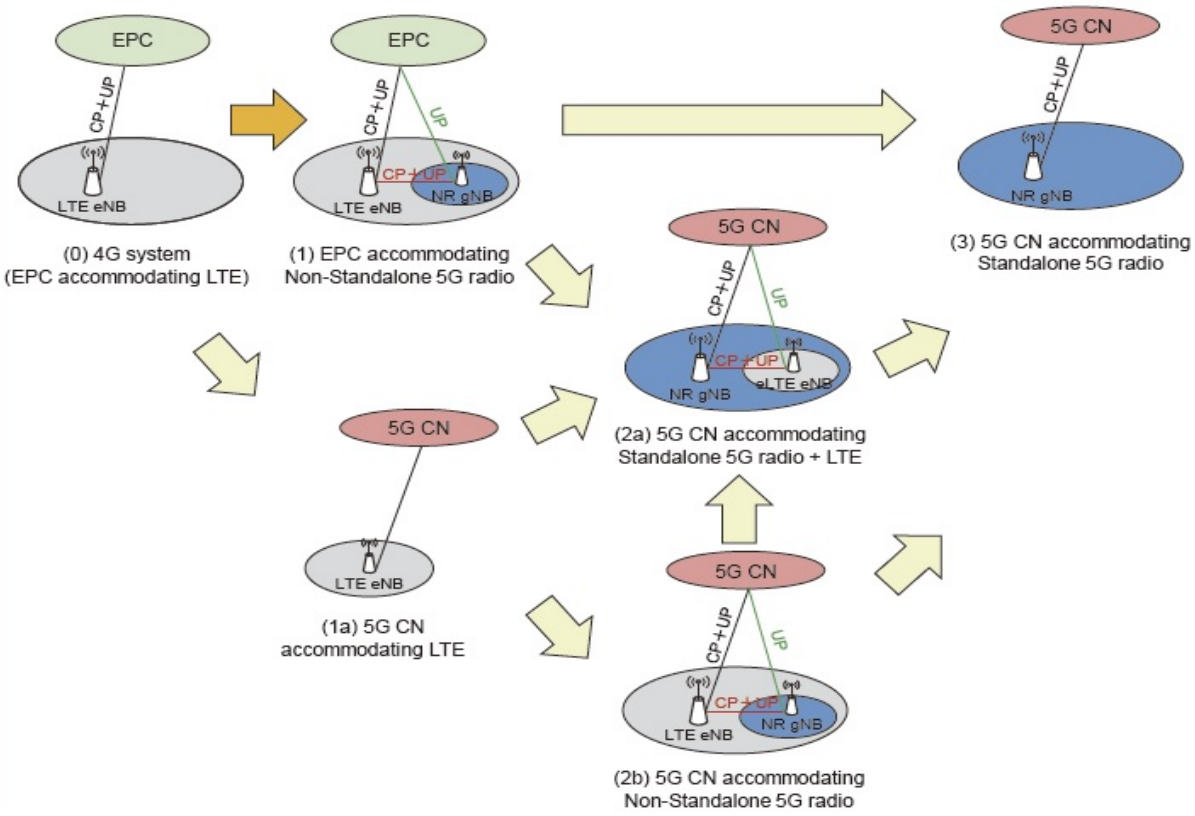


Fig. 2.7 5G NR deployment stages [2].

2.7.3 Non-Standalone Operation of NR

A significant feature of NR is that it supports NSA operation, which provides NR service in the LTE and LTE-A area with no need for having only an NR area. The existing LTE and LTE-A covering wide areas use the 2GHz and 800MHz frequency bands, while 5G in its early deployment stage is expected to work in the mmw frequency band. The operators will expand their service area by adding NR within the service area of LTE-A, according to the capacity demand. NSA will use carrier aggregation technology from LTE release 12 in order to coexist with LTE. This technology aggregates different carrier bands together, where macrocells and small cells use different bands for improving user throughput [60–62]. In brief, LTE carrier aggregation technology will be extended to be used between LTE and NR. Consequently, the user equipment should support dual connectivity. Additionally, in NSA operation, LTE will be used for initial access and mobility handling [58, 63]. 5G NSA advantages are:

- Maintaining the high probability of coverage for areas already using LTE/LTE-A services.
- Using EPC is convenient for eMBB use case.
- Reducing the number of new elements and testing in the 5G early deployment phase.

2.7.4 NR Architecture

The 5G RAN is described by 3GPP in [46] and shown in Fig. 2.8, where C-RAN architecture is adopted to be the RAN architecture of 5G NR with a functional split between the RRHs and BBUs. In the NSA scenario, the NR base station is denoted as en-gNB and is linked to the LTE eNB using the $X2$ interface. $X2$ is used in LTE for connecting multiple eNBs together. However, it has been extended to connect en-gNB to eNB in the NSA scenario. Additionally, in NSA NR, en-gNBs are connected to EPC using $S1$ interface. On the other hand, in the SA scenario, the service will be provided by gNB only, as shown in Fig. 2.8. That is, gNBs are connected to the new 5G core network via the NG interface, whereas multiple gNBs are connected together using the Xn interface [46, 64].

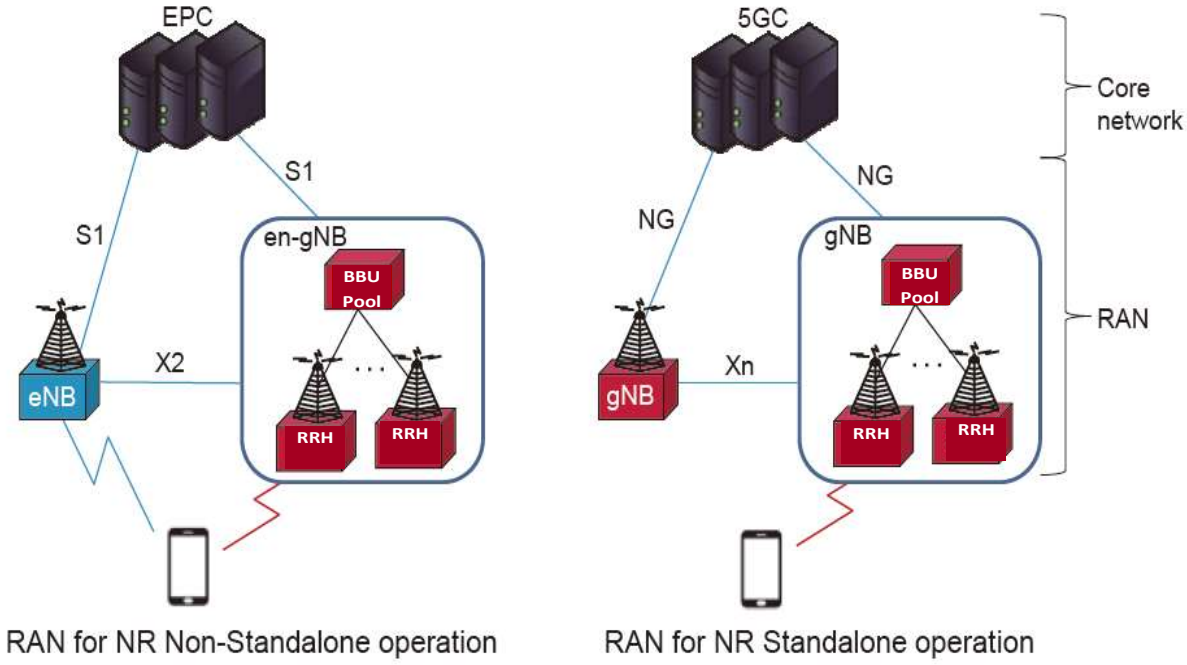


Fig. 2.8 5G Radio access network. [64].

2.8 State of the Art

Artificial Intelligence (AI) techniques facilitate a network to automatically re-configure system parameters for optimum network performance and adaptively learn about the necessary system parameters to perform upgrades and maintenance routines along with recovering from failures. Since AI, is the basis of self-organising and machine learning network technologies, it can lead to a significant paradigm shift by driving the ongoing efforts in next-generation wireless network 5G standardisation [65].

Furthermore, the benefits of AI techniques while designing load balancing SON algorithms are inevitable. Among numerous AI techniques, the Genetic Algorithm (GA) [66, 67] and Swarm intelligence [68] are the most embraced learning algorithms inspired the process of gene evolution and the natural actions of swarms of ants, a shoal of fish, a flock of birds etc, respectively. Many algorithms have been designed to mimic the behaviour of natural organisms. However, Particle Swarm Optimisation (PSO) [69] remains the backbone of swarm intelligence on which all other algorithms are built. Both GA and PSO are widely discussed in studies related to network planning, interference management, routing and coverage optimisation problems [70–73]

Most recent studies on resource management in C-RAN mainly focus on schemes related to RRH-UE mapping and only limited work addresses the BBU-RRH configuration schemes. Some related works on the former schemes are briefly discussed in [74–76]. In [74], the authors propose a QoS-aware radio resource optimisation solution for maximising downlink system utility in C-RAN. User grouping, virtual base stations clustering, and beamforming for multiuser, multicell distributed MIMO networks were investigated. In line with this work, the authors of [75] propose an efficient resource allocation scheme in heterogeneous C-RAN. A weighted minimum mean square error (WMMSE) approach is used to solve network-wide beamforming vectors optimisation and identify proper RRH-UE clusters. Moreover, minimising the number of active BBUs is formulated as a bin packing problem for energy saving. The work of [76] expresses a mixed integer non-linear programming (MINLP) problem aiming joint RRH selection to minimise power consumption via beamforming, where the transport network power

is determined by the set of active RRHs. Regarding the BBU pool in C-RAN, some studies are described in [77–80]. A joint-scheduling strategy for resource allocation in C-RAN is proposed in [81] where the time/frequency resources of multiple base stations are jointly optimised to schedule network users concurrently for network throughput improvement. However, the authors did not consider BBU-RRH mapping and focused mainly on joint scheduling in C-RAN. The authors of [77] initially investigated semi-static and adaptive BBU-RRH switching schemes for C-RAN. The authors of [78] then proposed a lightweight, scalable framework that utilises optimal transmission strategies via BBU-RRH reconfiguration to cater dynamic user traffic profiles. A dynamic BBU-RRH mapping scheme is introduced in [79] using a borrow-and-lend approach in C-RAN. Overloaded BBUs switch their supported RRHs to underutilised BBUs for a balanced network load and enhanced throughput. The work in [82] address a blocking probability based load balancing problem in C-RAN via evolutionary algorithms. However, power saving in C-RAN was not addressed. Furthermore, there have been attempts to develop Network Function Virtualisation (NFV) and Software Defined Network (SDN) solutions for C-RAN [83–85]. Moreover, an in-depth review of the principles, technologies and applications of C-RAN describing innovative concepts regarding physical layer, resource allocation, and network challenges together with their potential solutions are highlighted in [86, 87].

To sum up, the existing resource allocation mechanisms does not take full advantage of the centralised BBU pool concept in C-RAN. This thesis extends the scope of C-RAN by introducing the concept of Cell Differentiation and Integration (CDI) with dynamic BBU-RRH mapping for load balancing and efficient resource utilisation. The system model in this thesis allows combining self-optimising feature of SON and capacity routing ability of C-RAN for a more centrally managed network operations. The problems of capacity increasing and load balancing in cellular networks have been addressed using SON in Numerous studies via different approaches. When a burst of traffic or load imbalance is detected, the network responds by autonomous adjustments in operating parameters. Cell Range Expansion (CRE) in LTE-A is a technique to offload users from

macrocells to small cells by extending the coverage area of small cells. Users associate to small cells, despite receiving the strongest signal macrocell. In 3GPP specifications, coverage and capacity optimisation is a crucial feature of SON. Authors in [88] compared online and offline SON solutions for simultaneous capacity and coverage optimisation to maximising network performance. In [89] a SON management mechanism is proposed for load balancing and capacity increase in future network. The suggested Automated Load Balancing and Capacity Enhancement (ALBCE) mechanism optimise the coverage area of macro and small cells based on the load in the cells with a minimum number of handovers constraint. In the same context, [90] investigates reducing the coverage area to achieve Mobility Load Balancing (MBL), where the handover thresholds are adjusted depending on traffic conditions that lead to increasing or decreasing the users due to the change in virtual transfer areas among adjacent cells. Yet, if handover parameter adjustment is incorrect, this leads to additional assignments in the network causing handover ping-pongs/delays and a weak radio link. Furthermore, SON autonomous adjusting of other operating parameters such as antenna tilt [91] is investigated in other studies. [92] extended the SON automation to other functionalities achieving more performance advantages. [93] presented a comprehensive survey with a detailed description of self-organisation schemes suggested in the literature for future cellular networks. As mmw frequency communication has distinct characteristics of sensitivity to blockage and directional signal transmission, low latency is a crucial feature of small cell mmw link. However, for traditional SON the healing and optimisation operations start after detecting a problem. This will impose an unavoidable delay, which will be against the low latency requirements of the mmw link. To this end, [26] suggested a proactive SON instead of the legacy reactive SON to cope with the ultra-low latency of the mmw link. Moreover, centralised architecture (C-RAN) is inevitable to overcome the issues mentioned above of mmw link by migrating the processing unit to a BBU pool and leaving the radio unit (RRH)s in the cell sites for low latency centralised cooperation. Several studies also exploit the advantages of C-RAN architecture for mmw communication such as in precoding [94], scheduling [95], and receiver design [96]. In [97] the perfor-

mance of mmw has been analysed. The realisation of centralised SON is discussed in [98]

On the other hand, many research studies on enabling technologies for C-RAN exist. Here, some related studies on BBU-RRH mapping along with RRH-UE association are briefly described. In [99], the authors propose a cross-layer framework for downlink multi-hop C-RAN to improve throughput performance by optimising network resources. Also, RRH's beamforming vectors, user RRH association, and network coding based routing are optimised in an overall design. In [100], the authors attempt to solve a joint RRH and precoding optimisation problem which aims to minimise network power consumption in a MIMO based user-centric C-RAN. [101] describes the traffic adaptation and energy-saving potential of TDD-based heterogeneous C-RAN by adjusting the logical connections between BBUs and RRHs. The authors of [102] recently investigated an RRH clustering design and proposed a Spectrum Allocation Genetic Algorithm (SAGA) to improve network QoS via efficient resource utilisation.

During the last few years, there has been a dramatic increase in the number of mobile connected devices, especially smart phones. It is predicted that their number will surpass the projected world population, reaching 11.6 billion devices by 2021[103]. Moreover, the emerging data ravenous applications have led to an exponential increase in global mobile data traffic and this is expected to continue, increasing seven fold in the coming five years, which will push the existing system capacity beyond its limit. The millimetre wave is an attractive solution for tackling the future capacity deficiency because of its wide bandwidth, which spans from 30-300 GHz. A millimetre wave's cellular communication differs from conventional cellular systems in being short range and intermittent, because these frequencies have high path losses and are vulnerable to blockage owing to shadowing by obstacles. So, directional communication using a high gain multiple antenna system is used to compensate for the high path losses[104]. For this reason, in its early deployment mmw base stations will have to coexist in small cells with a conventional μw BS in a hybrid or multi-band HetNet architecture for a ubiquitous, reliable and robust communication link that meets the strict QoS requirement of the

5G[105]. Yet, this architecture has some drawbacks as the mmw cell coverage is not continuous due to blockage vulnerability, so the User UE is required to run cell search all the time to find another small cell and perform frequent handovers even for stationary users. Moreover, the mobility management will be more complicated. To avoid the regular handovers and delay of re-association with another small cell, the UE may adopt to receive multi-beams from different directions of BSs[106]. Furthermore, the centralised architecture will help to overcome the aforementioned issues. In H-CRANs, the BBU pool is interfaced with μw BSs for load balancing, handover and mobility management between the mmw RRHs and μw BSs using the centralised cloud computing-based platform [106–108, 104, 109, 110]. Moreover, the centralised architecture enables small cell clustering for efficient resource utilisation for reliable cooperated mmw communication. Accordingly, much work has been carried out in this direction. For example, in [111] a UE connectivity to multi-RRHs approach is investigated to maximise the users throughput reliability. Performance analysis of mmw C-RAN for two UE-RRH association scenarios are proposed in [94], which are nearest neighbour participation and best channel participation. The outage probability is derived and the result show that blockage and high pathloss decreases intra cluster interference.

Several latest studies have addressed heterogeneous mmw cellular networks as enabler for 5G. Authors in [106] classify heterogeneity in mm Wave cellular networks into two different types: spectrum heterogeneity and deployment heterogeneity. In the first type of UEs use mmw frequencies for data communication while μw frequencies used for control signals communication. The other type of heterogeneity introduced two deployments scenarios stand alone and integrated. All the tiers in stand alone scenario are working in mmw band whereas in integrated scenario μw and mmw cells work together. A Hybrid mmw HetNet for interference mitigation in HetNets is introduced in [112], where the V band to be used for short range high speed point to point and the E band used in connecting the base stations. [107] proposed a complete cellular network architecture where the mmw and μw BSs are connected to a centralised RAN (C-RAN) for efficient resource management, cell cooperation and coordination. In [113] a hybrid mmw and μw cellular

system is explored where the uplink-downlink coverage and rate distribution is evaluated where mmw UEs are offloaded to μw band on bad mmw channel conditions. A similar approach is considered in [108] in which uplink and downlink association is decoupled and opportunistically offloaded to mmw cells. An analysis framework for the mmw cellular network using stochastic geometry is presented in [114]. It especially considers real data derived path-loss and blockage models. The work in [115] investigates the effect of BSs cooperation on coverage in downlink mmw heterogeneous network. The rate and coverage of hybrid mmw and μw cellular system with blockage model of actual building location in suburban environment is explored in [116] and the noise limited property of 5G hybrid mmw channel is highlighted in the analysis. [47] tested the performance of mmw in HetNet and introduced a novel 3GPP backward compatible mmw frame structure, this achieved about 13Gb/s aggregated network throughput which is favourable for 5G. Towards this end, [117] explored the potential advantages and addressed the challenges of mmw HetNets. Also a new dual mode small cell design is proposed to exploit both mmw and μw bands advantages where the coverage area is divided into three spaces inner, middle and outer. The inner space is served by mmw, the middle space is served using both mmw and μw and the outer space is served by μw. The authors in [118] suggested a new joint mmw and μw scheduling framework which is a user application context aware scheduler for dual mode small BSs in HetNet. The scheduling decision is made based on considering several user application and network context information to maximise user application QoS. Furthermore, to tackle the problem of mobility management in integrated mmw-μw cellular system, [119] proposed a comprehensive frame work where cached enabled mobility management problem is formulated and dual mode base stations capability is used to reduce the probability of hand over failure for mobile UEs. In a relay based multi-band HetNet the mmw propagation properties is exploited in a joint recourse allocation formula that is optimised in [120] to increase the overall network throughput. Optimal load balancing in cellular mmw HetNet has been studied in [121], where the load on macrocells is reduced by offloading users to mmw small cells through optimising the biasing factor to increase the network rate and coverage. Signal to Interference and

Noise Ratio (SINR), rate and coverage is derived analytically for the downlink K-tier mmw HetNet. Also, the effect of biasing factor used for the small cells on the overall energy efficiency of network is analysed in [122]. A non-uniform HetNet and a cell association scheme that selectively switching off specific small cells is proposed in [123]. The cell area is divided into two subspaces; inner space where the users are associated to the macro BS whereas the users in the outer space are connected to either macro BS or small BS depending on the scheme of Biased/unbiased maximum received power.

To sum up, the suggested mmw and μw heterogeneous network in literature does not take the full advantage of the noise limited property of mmw band cells to combat interference in traditional μw network. This thesis suggests a scheme to exploit the mmw band properties to combat interference in cell edge area.

Chapter 3

A Resource Allocation Mechanism for C-RAN Based on Cell Differentiation and Integration [1]

Adaptability to load variation by scaling the capacity is one of the appealing Cloud Radio Access Network (C-RAN) features by the means of centralised management and processing. In this regard, in this chapter, a self-organising cloud radio access network is introduced, which dynamically adapts to varying capacity demands. The Base Band Units (BBUs) and Remote Radio Heads (RRHs) are scaled semi-statically based on the concept of Cell Differentiation and Integration (CDI), while dynamic load balancing is formulated as an integer-based optimisation problem with constraints. A Discrete Particle Swarm Optimisation (DPSO) is developed as an Evolutionary Algorithms (EA) to solve the load balancing optimisation problem. The performance of DPSO is tested based on two problem scenarios and Exhaustive Search (ES) algorithm. The DPSO delivers optimum performance for small-scale networks and near optimum performance for large-scale ones.

[1] A part of this chapter has been published in *IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING*, October 2018 [124] and IEEE TRANSACTIONS ON NETWORK SERVICE AND MANAGEMENT, March 2018 [125]

3.1 Introduction

In the past few years, the proliferation of personal hand-held mobile computing devices such as tablets and smartphones, along with the growing volume of data-demanding services and applications, has produced and a great need for wireless access and high-speed data transmission. Internet access anywhere and everywhere has triggered the formation of radio *hot-spot* networks. The major challenge in cellular networks is managing the available resources in a way to achieve 1) Optimum returns on investment, 2) User's service demands satisfaction, and 3) High levels of network Quality of Service (QoS). Unaware of the cell load, a User Equipment (UE) associates itself to the cell providing the strongest signal. The spatial distribution of users and their capacity demands vary with respect to time, causing unbalanced traffic loads and wasteful utilisation of network's resources. Therefore, it is important to self-optimise the network resources dynamically. To overcome the aforementioned challenges, C-RAN [87, 126, 127] has been proposed as a novel architecture that can address some significant challenges the Mobile Network Operators (MNOs) are facing with today. C-RAN architecture is composed of three parts: 1) The BBUs collected into a virtualised BBU cloud/pool for centralised processing, 2)The RRHs in the radio access network, and 3) An Optical Transport Network (OTN) that connects the BBUs to the RRHs. C-RAN can achieve significant cost and energy savings by dynamically scaling the BBUs with respect to changing traffic conditions [128] and adjusting the logical BBU-RRH links using suitable resource allocations schemes. C-RAN with self-optimising ability can provide MNOs with a flexible network regarding network dimensioning, adaptation to non-uniform traffic, and efficient utilisation of network resources. However, before a full commercial C-RAN deployment, several challenges need to be addressed. Firstly, the front-haul technology used must support enough bandwidth for delivering delay sensitive signals. Secondly, the proper BBU-RRH assignment in C-RAN to not only support collaboration technology like Coordinated Multi-Point transmission (CoMP) [129] but also enabling dynamic load balancing and power saving in the network. The main motivation of this chapter is to exploit the capacity routing ability of C-RAN by employing self-optimisation

for efficient resource utilisation with high levels of QoS and a balanced network load. Inspired by the concept of cell splitting in biological sciences, a two-stage design is proposed for real-time BBU-RRH mapping and power saving in C-RAN. The main contributions of the proposed scheme is as follows:

1) The proposed mechanism monitors the load on each cell in a given geographical area and divides it into multiple small cells and vice versa if the load in a cell exceeds or falls a certain threshold.

2) The fitting number of BBUs required to serve all RRHs in the given geographical area is assigned based on the actual load on the network. A key challenge of initial BBU-RRH mapping before identifying an optimum BBU-RRH mapping is also addressed.

3) An EA is proposed to find the optimum BBU-RRH configuration to balance the network load for enhanced QoS. Therefore, the resources can be utilised efficiently.

3.2 System Model

3.2.1 Proposed C-RAN Architecture

A self-optimised C-RAN architecture is presented in Fig.3.1. The BBUs are decoupled from the RRH and migrated to a centralised BBU-pool, whereas the RRHs are left on the cell sites. A Self-Organising Network (SON) controller is introduced inside the BBU cloud which monitors the BBU-pool resource utilisation as well as controlling the switch. Since an optical switch can only support one-to-one switching, soft switching (one-to-one and one-to-many) is enabled indirectly by using optical splitters and multiplexers [78]. The SON controller dynamically assigns the required number of BBUs in the BBU pool to the independent RRHs based on traffic demands. However, each BBU allocates its radio resources Physical Resource Block (PRBs) only to the RRHs assigned to it at a particular time.

At extremely low traffic load conditions, only a high power Macro Base Station (MBS) serves the given geographical area. As the traffic load increases and the MBS

reaches its load limit, the geographic area is differentiated into **C** equally sized small cells serving the same coverage area. Each **C** cell can further differentiate into **c** more small cells by activating the CDI supporting RRHs deployed to accommodate capacity demands. The actual number of RRHs are determined by the coverage area, users density, and other environment-related factors. However, both **C** and **c** are considered to be seven as a reasonable example.

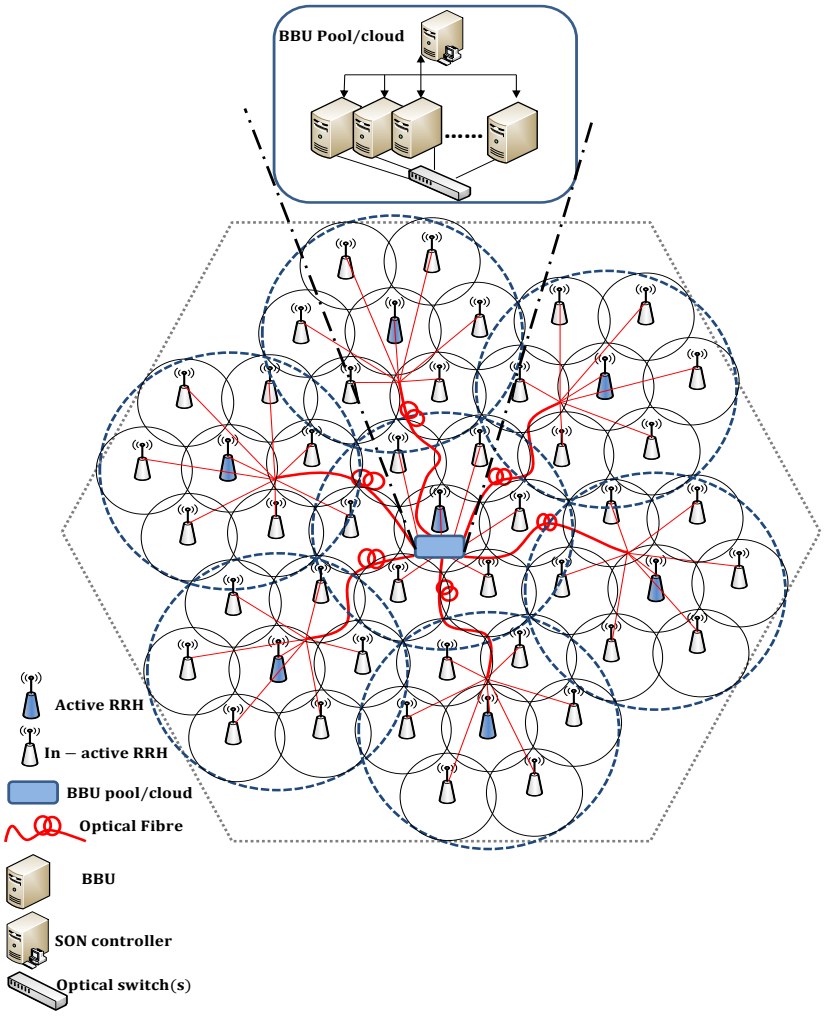


Fig. 3.1 Structure of a cloud radio access network with a SON controller.

Furthermore, the CDI concept is realised by considering three tiers of RRHs deployment as shown in Fig.3.2, i.e., tier-3 RRH deployment imitates a high-power base station serving a macrocell as in traditional cellular systems. Tier-2 and tier-1 represents a structure with universal frequency reuse, where each cell is surrounded by a continuous tier of $6+E$ and $6\times[1+j]+E$ cells, respectively. Where E represents the

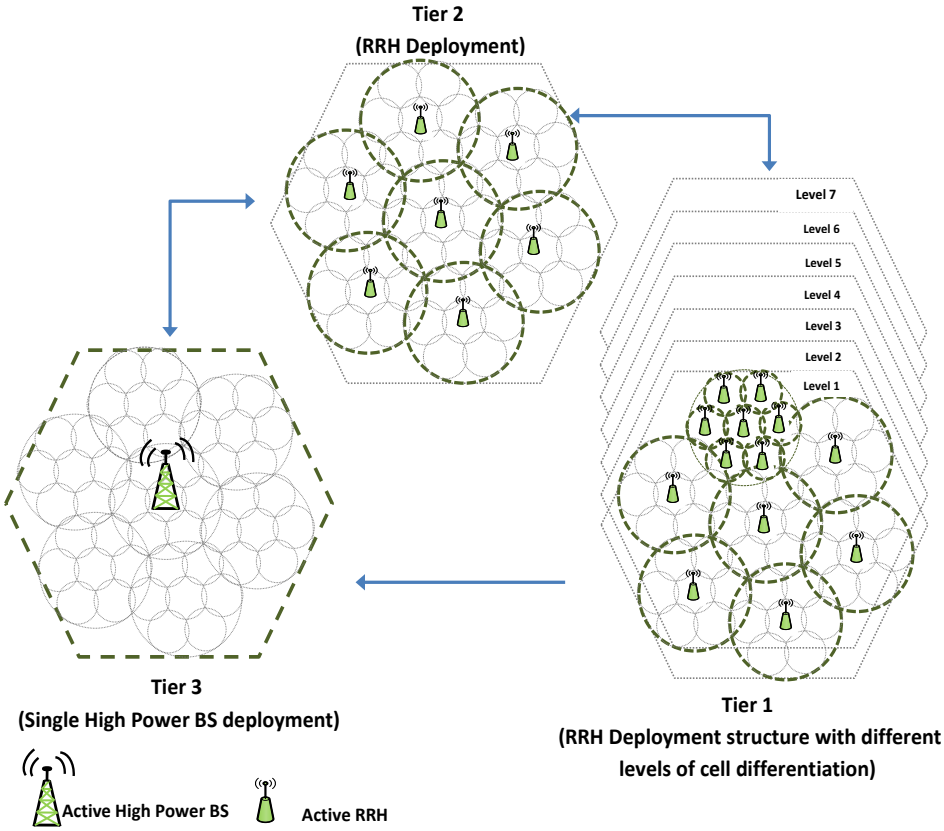


Fig. 3.2 Cell differentiation and integration with multiple tiers of RRH deployment.

number of other external macrocells and j accounts for the level of differentiation. A set $\mathbf{S}_i = \{\text{RRH}_{i1}, \text{RRH}_{i2}, ..., \text{RRH}_{ic}\}$ is maintained for each cell C_i in tier-2 RRH deployment, which contains a group of RRHs responsible for differentiating cell C_i into c small cells provided that the sum of transmit powers of all RRHs covers C_i coverage area.

The central RRH of each cell C_i is represented as RRH_{i1}. Where i represents the cell number in tier-2 RRH structure. The SON server is responsible for cell differentiation and integration with proper BBU-RRH configurations, whereas the optical switch is in charge of realising the settings via server commands. Note that, with small cell deployment in C-RAN, a high inter-cell interference is inevitable. Therefore, a clustering based interference mitigation technique is adopted to avoid network performance degradation. RRHs served by the same BBU are grouped together based on a proximity property [82].

3.2.2 Channel Model

In this study, Guaranteed Bit Rate (GBR) users with QoS requirements are considered. The frequency reuse factor is 1, and the time-frequency resources are equal for all BBUs. The basic unit of time-frequency resources that can be allocated to users per time slot (0.5 ms) of an Long Term Evolution (LTE) subframe is known as the PRB. Each PRB consists of 12 consecutive sub-carriers with a sub-carrier spacing of 15 kHz, corresponding to 0.5 milli-seconds in time domain and 180 kHz in frequency domain. Let **M** and **N** represent the number of active BBUs and RRHs in the network, respectively, such that K_{in} represents the total number of users in cell i served by RRH n. Each user reports Channel Quality Information (CQI) to its serving BBU every two subframes (i.e., 2 milli-seconds) for proper PRB assignment. The channel model considered in this work is a composite fading channel which involves path-loss and both small and large scale fading [130], given as:

$$\mathrm{H}_{k_{in}} = \mathrm{h}^{*}_{k_{in}} l_{k_{in}} \left[\mathrm{AD}^{-\delta}_{k_{in}}\right] \tag{3.1}$$

where $\mathrm{h}^{*}_{k_{in}}$ and $l_{k_{in}}$ represent the small and large scale fading channel between the RRH n and user k in cell i, respectively. The small scale fading is assumed to be a Rayleigh random variables with a distribution envelop of zero-mean and unity-variance Gaussian process. $\mathrm{AD}^{-\delta}_{k_{in}}$ reflects the path-loss between RRH n and user k in cell i, where A is a constant which depends on the carrier frequency f_c and $\mathrm{D}_{k_{in}}$ is the distance between user k and RRH n in cell i and a path-loss exponent of δ. In this work, a path-loss of $(\mathrm{A}, \delta) = (1.35 \times 10^7, 3)$ is considered [131, 130]. The large scale fading is assumed to be lognormal random variable with a standard deviation of 10dB and is typically modelled with a probability density function of [131]:

$$\rho(l) = \frac{\zeta}{\sqrt{2\pi}\sigma_l l} \exp\left[-\frac{(10\log_{10} l - \mu_l)^2}{2\sigma_l^2}\right] \tag{3.2}$$

where $\zeta = 10/\ln 10$, and μ_l and σ_l are the mean and the standard deviation of l, both expressed in decibels.

The instantaneous Signal-to-Interference-and-Noise-Ratio (SINR) γ based on CQI received from user k in cell i served by RRH n at time-slot t is expressed as

$$\gamma_{k_{in}}(t) = \frac{\mathrm{H}_{k_{in}}(t)\mathrm{P}_{in}(t)}{\mathrm{N}_0 + \sum_{j\in C}\sum_{a\in c, a\neq n}\mathrm{H}_{k_{ja}}(t)\mathrm{P}_{ja}(t)} \tag{3.3}$$

where $\mathrm{P}_{in}(t)$ and $\mathrm{H}_{k_{in}}(t)$ are the transmit power and channel gain between the serving RRH n of user k at time-slot t in cell i . N_0 is the power of Additive White Gaussian Noise per PRB and $\sum_{j\in C}\sum_{a\in c, a\neq n}\mathrm{H}_{k_{ja}}(t)\mathrm{P}_{ja}(t)$ represents the inter-cell interference power received from all other active RRHs a at time-slot t in cells j except the serving RRH n of user k in cell i.

Assuming the best modulation coding scheme for a given radio channel, the highest data rate achieved by user k served by RRH_n in cell i for a given SINR at time-slot t can be expressed by Shannon formula:

$$\vartheta_{k_{in}}(t) = \log_2(1 + a\gamma_{k_{in}}(t)) \tag{3.4}$$

where a is the constant bit error rate (BER) defined as $a = -1.5/ln(5\times 10^{-6})$ [132]. The total PRB required by the user can be now be determined by the achievable throughput of the user k at a given SINR, the demanded data rate ϕ_k of user k, and the bandwidth P_{BW} of a single PRB (i.e., 180 KHz) from the following:

$$\mathrm{N}^k_{\mathrm{RB}}(t) = \left\lceil \frac{\phi_k(t)}{\mathrm{p}_{\mathrm{BW}}.\vartheta_{k_{in}}(t)} \right\rceil \tag{3.5}$$

where p_{BW} represents the bandwidth of a PRB and the notion $\lceil . \rceil$ is the ceil function.

3.3 Dynamic BBU-RRH Configuration and Formulation

For a Self-optimising C-RAN architecture shown in Fig.3.1, it is essential to balance the network load amongst the active BBUs by proper BBU-RRH configuration. After each CDI cycle, the network may reconfigure itself by scaling the BBUs and RRHs with respect to traffic load. However, during the process, the BBU-RRH mapping might not satisfy the QoS requirement. Therefore, If the BBU-RRH configuration at time

t is known then it is necessary to adjust the BBU-RRH configuration at time $t+1$ to adaptively balance the variance in traffic demands. Note that, the time between t and $t+1$ is longer than that of a subframe (i.e., one millisecond) and is called the load balancing cycle. A user location indicator vector $\mathbf{u} = \{u_1, u_2, ..., u_K\}$ is defined which shows users association with RRHs such that $u_k = \{r_{in} | r_{in} \in \mathbb{Z}^+ : i,n = 1,2,3,...C\}$, where $u_k = r_{in}$ if user k is associated with RRH_n of cell C_i. To indicate RRHs association with BBUs, a vector $\mathbf{r} = \{r_{11}, r_{12}, ..., r_{in}\}$ is defined, where $r_{in} \in \{1,2,...,M\}$ and $r_{in} = m$ indicates RRH_n of cell C_i is being served by BBU_m. Whereas, $r_{in} = 0$ indicates that RRH_n of cell C_i is not active. If the user location indicator vector $\mathbf{u}$ is given, then the problem is to identify the new RRH allocation vector $\mathbf{r}$.

3.3.1 Number of BBUs required in the network

The required number of BBUs to serve the offered traffic load at a particular time t can be calculated using actual load $\eta(t)$ on the network. Let $\eta_m(t)$ be the load on BBU_m at time period t, which is represented as

$$\eta_m(t) = \frac{\sum_{k=1}^{K} I_{m,k}(t) N_{RB}^k(t)}{P_{RB}} \tag{3.6}$$

Where $I_{m,k}$ is a binary indicator such that $I_{m,k} = 1$ if user k is served by BBU_m. However, an important constraint $\sum_{m=1}^{M} I_{m,k} = 1, \forall k$ defines that each user k is served by only one BBU at time period t. Note that, all BBUs are assigned the same number of PRBs (P_{RB}). Another important constraint is that $\sum_{k=1}^{K} I_{m,k}(t) N_{RB}^k(t) \leq P_{RB}, \forall m$, which states that the number of PRBs assigned to users served by the same BBU should not exceed the BBU PRB limitation. The total load on the network at time t is represented as the aggregated load on each active BBU at time t, which is given by

$$\eta(t) = \sum_{m=1}^{M} \eta_m(t) \tag{3.7}$$

Now the number of required BBUs (M) in the network at a particular time t can be given as:

$$\text{No. of BBUs= M} = \begin{cases} \lceil \eta(t) \rceil & if\ \eta(t) < \text{M}_{\text{total}} \\ |\text{M}_{\text{total}}| & if\ \eta(t) \geq \text{M}_{\text{total}} \end{cases} \tag{3.8}$$

where M_{total} is the total number of BBUs available in the BBU pool and the notation $\lceil . \rceil$ is the ceil function. Moreover, the load contributed by an active RRH_n of cell i in the network is given by

$$\eta_{\text{RRH}_{in}}(t) = \sum_{k=1}^{\text{K}} \text{I}_{k,in}(t)\text{N}_{\text{RB}}^{k}(t) \tag{3.9}$$

Network performance determined by Key Performance Indicators (KPIs).

3.3.2 Key Performance Indicator for Load Fairness Index

In this study, a Jains fairness index ψ is monitored, which determines the level of load balancing in the network at a particular time t and is defined as:

$$\psi(t) = \frac{\left(\sum_{m=1}^{\text{M}} \eta_m(t)\right)^2}{\text{M}\left(\sum_{m=1}^{\text{M}} \eta_m^2(t)\right)} \tag{3.10}$$

where M is the number of active BBUs. The range of ψ is in the interval $[\frac{1}{M}, 1]$, with a higher value representing a highly balanced load distribution amongst all active BBUs. Therefore, maximising ψ is one of the objectives to achieve a highly balanced load in the C-RAN.

3.3.3 Key Performance Indicator for Average Network Load

Minimising the average network load can avoid handovers between BBUs for users with poor channel conditions in the system. A user(s) associated to an RRHs may have imperfect channel conditions with more PRBs requirement to meet desired data rate. Failure to meet the user's PRB demand, the BBU has to perform a handover operation. Therefore, to avoid unnecessary handovers, minimising the average network load is considered as a second objective and is given as

$$\eta_{ave}(t) = \frac{\sum_{m=1}^{\mathrm{M}} \eta_m(t)}{\mathrm{M}} \tag{3.11}$$

where $\eta_m(t)$ is the load on a BBU_m defined in eq. (3.6) and M is number of active BBUs calculated from eq. (3.8).

3.3.4 Key Performance Indicator for Handovers

Network transition to a new BBU-RRH configuration may require significant forced handovers. An increased number of forced handovers in the system is undesirable and leads to performance degradation. Allocating an RRH to a new BBUs at a particular time results in forced handovers of all users associated with the RRH. Since inter-BBU handovers not only involves BBUs but a signalling overhead between the Serving Gateway (S-GW) and Mobility Management Entity (MME). Therefore, it is desirable to achieve a new optimum BBU-RRH configuration with a minimum required handovers. A handover index $\mathrm{h}(t)$ is monitored as a third objective for load balancing problem and is given as:

$$\mathrm{h}(t) = \frac{1}{2}\left(\frac{\sum_{m=1}^{\mathrm{M}} \sum_{k=1}^{\mathrm{K}} |\mathrm{I}_{m,k}(t) - \mathrm{I}^{\circ}_{m,k}(t)|}{\mathrm{K}}\right) \tag{3.12}$$

where $\mathrm{I}^{\circ}_{m,k}(t)$ is a binary variable that indicates a user's association to BBU in previous BBU-RRH configuration i.e., $\mathrm{I}^{\circ}_{m,k}(t)$=1, if user k is served by BBU_m in previous BBU-RRH configuration.

3.4 RRH Clustering

Proper BBU-RRH association can provide enhanced flexibility in C-RAN network management. However, an important limitation to consider is the reliable operation of C-RAN regarding BBU-RRH mapping for high system performances. Neighbouring RRHs must be assigned to the same BBU [79] to support advance LTE-A features like CoMP and to avoid unnecessary handovers among network cells. RRHs clustering is an approach to support CoMP for interference mitigation in LTE-A and C-RAN [133]. Therefore, this work considers that the RRHs served by the same BBU forms a compact cluster.

This compactness and consistently connected RRHs in a cluster, not only minimises frequent handovers among cells but also reduces the inter-cell interference among them. This is because compact an RRH group shares fewer common boundaries with other RRH groups. Therefore, the RRHs proximity is defined by introducing a binary variable A_{ij}, where A_{ij}= 1, if RRH_i and RRH_j are adjacent else A_{ij}= 0. If a cluster has multiple RRHs, then the RRHs in that cluster must be adjacent and connected. To formulate the connectedness of a cluster and proximity of the RRHs, let S1 be any proper subset of the set of RRHs served by BBU_m (Z_m), such that $S1 \subset Z_m,\ S1 \neq \emptyset$, and $S1 \neq Z_m$. Let S2 be another subset of Z_m such that, $S2 = Z_m - S1$, i.e., S2 is the complementary set of S1. To confirm that the RRHs in Z_m are connected, the following property must be satisfied.

$$\sum_{i \in S1} \sum_{j \in S2} A_{ij} \geq 1 \tag{3.13}$$

For proper BBU-RRH configuration, a QoS function is needed which is the weighted combination of KPIs defined in eqs. (3.10), (3.11), and (3.12). The multiple objectives are combined into a single QoS objective function. This work represents QoS as the following maximisation problem with constraints:

$$\begin{aligned} \text{Max} \quad & \text{QoS}(t) = \alpha\psi(t) - \beta\eta_{ave}(t) - (1-\alpha-\beta)\text{h}(t) \\ s.t. \quad & \text{C}_1 : \sum_{i \in S1} \sum_{j \in S2} A_{ij} \geq 1, \forall S1, S2 \in Z_m, \forall m \in \{1,2,...M\} \\ & \text{C}_2 : \sum_{k=1}^{K} I_{m,k}(t) N_{RB}^{k} \leq P_{RB}, \forall m \in \{1,2,...M\} \\ & \text{C}_3 : \sum_{m=1}^{M} I_{m,k}(t) = 1, \forall k \in \{1,2,...,K\} \end{aligned} \tag{3.14}$$

Both α and β are control parameters of the QoS function. The main objective is to maximise the QoS function.

3.5 Power Model for C-RAN

This section explains the necessary aspects needed to assess the power consumption of C-RAN. However, a more detailed description of the components involved in a C-RAN

power model is given in [43]. The three most important parts considered for the power model are described as follows.

3.5.1 BBU Power Estimation Model

The BBU performs a different set of functions (I_{BB}) which includes scheduling of PRBs, Forward error correction, Fast Fourier Transform (FFT) and Orthogonal Frequency Division Multiple Access (OFDM) specific processing, filtering, modulation/demodulation, and transport link related functions, etc. These features can be measured in Giga Operations per Second (GOPS) and then translated into power figures. About 40 GOPS per Watt is estimated as the power cost of a large BBU [134]. The power model for the BBU can be given as:

$$\mathrm{P}_{\text{BBU}} = \sum_{i \in I_{\text{BB}}} \mathrm{P}^{\text{ref}}_{i,\text{BBU}} \mathrm{A}^{x_i^A} \mathrm{W}^{x_i^w} \tag{3.15}$$

where $\mathrm{P}^{\text{ref}}_{i,\text{BBU}}$ in Watts represents the power consumption of BBU with respect to BBU functions. A is the number of antenna chains/ Radio Frequency (RF) transceivers with x_i^A scaling exponent. W is the bandwidth share used in transmission with a scaling exponent x_i^w. In [135], the authors model BBU operations with exact scaling components and reference values to calculate BBU power consumption, shown in Table.3.1

Table 3.1 BBU operations and their scaling values with transmit antennas and bandwidth

Processing type, i	GOPS	$\mathrm{P}^{\text{ref}}_{i,\text{BBU}}$ [W]	x_i^A	x_i^w
Time Domain Processing	360	9.0	1	0
Frequency Domain Processing	60	1.5	2	1
Forward Error Correction	60	1.5	1	0
Central Processing Unit	400	10.0	1	0
Common Public Radio Interface	300	7.5	1	0
Leakage	118	3.0	1	0

3.5.2 RRH Power Estimation Model

An RRH consist antenna chains/ RF transceivers, each with its own Power Amplifier (PA). The PA is main element of consideration as it consumes most of the power within

an RRH. The power consumption of a PA is affected by its power efficiency (η_{PA}). The power consumed by the PA can be given as $\text{P}_{\text{PA}} = \frac{\text{P}_{\text{TX}}}{\eta_{\text{PA}}(\sigma_{\text{feed}})}$, where P_{TX} is the output power of the PA, which depends on the bandwidth share (χ), i.e., the actual number of physical resource blocks (N_{RB}) used for transmission and the output power of the antenna P_{out} ($\text{P}_{\text{TX}} = \text{P}_{\text{out}}\chi$). σ_{feed} represents the feeder loss. Moreover, the RF transceiver units of an RRH are responsible for functions like signal modulation/demodulation, voltage controlled oscillation and mixing, AC-DC and DC-AC conversions, and low noise, gain amplification. The power consumed by an RRH can be modelled as:

$$\text{P}_{\text{RRH}} = \sum_{a=1}^{A} (\text{P}_{\text{PA}} + \text{P}_{\text{RF}}) \tag{3.16}$$

where $a \in \{1, ..., A\}$ denotes the number of antenna/RF chains for Multiple-Input-Multiple-Output (MIMO).

3.5.3 Optical Transceiver Power Estimation Model

In C-RAN architecture, the *front-haul* connectivity with high bandwidth, low cost, and low latency requirements for transport networks is challenging. Several factors influence the operation of optical transceivers such as the technology used, the operating conditions, and the output power required, which in turn affect the power consumption. From a power consumption perspective, the optical transceivers can be divided into two modules. The optical transmitter module, in which the OFDM electrical signals are modulated over optical carriers using an external or direct modulated lasers. And a receiver module which detects the optical OFDM signals either by direct detection or coherent detection. The power consumption of the optical transceiver as described in [136] can be given as:

$$\text{P}_{\text{TRANS}} = (\text{P}_{\text{laser}} + \text{P}_{\text{driver}} + \text{P}_{\text{I/O}})_{\text{TX}} + \left(\text{P}_{\text{PD}} + \text{P}_{\text{amp}} + \text{P}_{\text{I/O}}\right)_{\text{RX}} \tag{3.17}$$

where $\text{P}_{\text{laser}}, \text{P}_{\text{driver}}, \text{P}_{\text{I/O}}, \text{P}_{\text{PD}}$, and P_{amp} are the powers consumed by direct-modulated laser, electronics driving the laser, the electrical input/output interface, photodetector, and the trans-impedance and limiting amplifiers. This study consider a Point-to-Point (PtP) transceivers rather than point-to-multipoint, because the PtP link loss is driven by

distance and used operating wavelength only, i.e., the link loss of PtP is as low as 6dB with a 20 km network reach [137].

The total power consumption of a C-RAN ($P_{C\text{-}RAN}$) can be estimated by summing the power consumed by three main parts of the network along with power consumed by other components P_{OTHER} such as power conversions(AC-DC, DC-DC) and cooling, i.e.,

$$P_{C\text{-}RAN} = \sum(P_{BBU} + P_{TRANS_B}) + \sum(P_{RRH} + P_{TRANS_R}) + P_{OTHER} \tag{3.18}$$

Where P_{TRANS_B} and P_{TRANS_R} indicates the power consumption of PtP transceivers located at each BBU and RRH, respectively. According to [138], base stations with a total power consumption ≤ 500 Watts do not require a cooling system. This can be applied to RRH in C-RAN if its components (i.e., PA, RF, and optical transceiver) require an overall power less than 500 Watts. In this work, the cooling power for RRH is ignored considering supply power as the only overhead.

From [43], the supply power required for a base station can be estimated as an affine function of transmitting power. The power consumption can be expressed by a load-dependant part that linearly increases with a power gradient (slope) Δp and a static load independent part P_{static} as shown in Fig.3.3. Moreover, the supply power reaches a maximum P_1 when the transmitting power reaches the maximum limit P_{max}. A base station may enter an idle mode (sleep mode), with minimum power consumption (P_{sleep}) when it is not transmitting. The total supply power for a base station can be formulated as:

$$P_{supply}(\chi) = \begin{cases} P_1 + \Delta p \; P_{max}(\chi - 1) & \text{if } 0 < \chi \leq 1 \\ P_{sleep} & \text{if } \chi = 0 \end{cases} \tag{3.19}$$

where $P_1 = P_{static} + \Delta p \; P_{max}$. χ is a scaling parameter which indicates the bandwidth share, i.e., $\chi = 1$ indicates that the system is transmitting with full power and bandwidth whereas $\chi = 0$ represents an idle system. The basic power model presented in eq. (3.19) is parameterised to understand the contribution of different parameters. Parameters

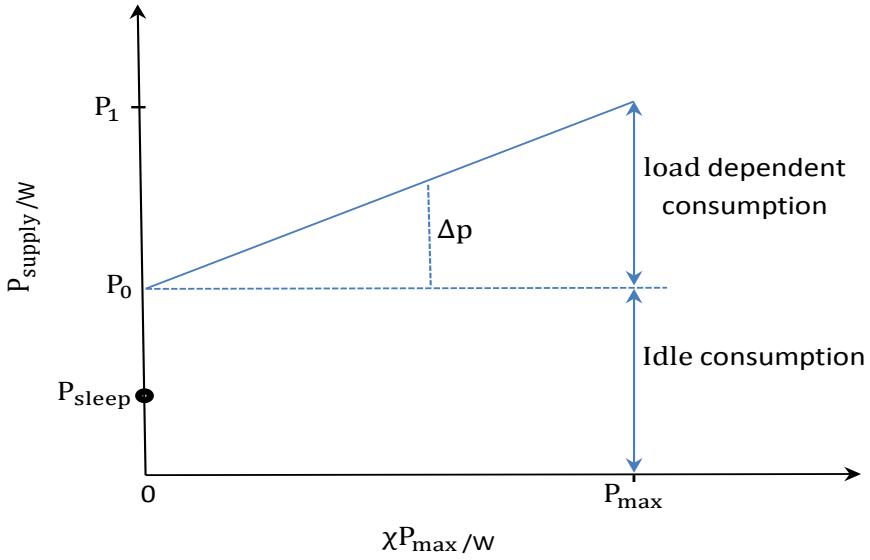


Fig. 3.3 The required base station supply power as a function of transmitting power.

which are assumed to be constant or having negligible effects are also highlighted. The following approximations are made:

• Both the BBU and RF power consumption, linearly scales with the number of Antennas (A) and bandwidth (W), i.e., $P_{BBU} = A(\frac{W}{BW_{TOTAL}})P^{pm}_{BBU}$ and $P_{RF} = A(\frac{W}{BW_{TOTAL}})P^{pm}_{RF}$. Where P^{pm}_{BBU} and P^{pm}_{RF} are parameterised power consumption of BBU and RF, respectively.

• Each antenna unit of an RRH has a PA. The power consumed by a PA depends on the maximum power transmission per antenna unit $(\frac{P_{max}}{A})$ and its efficiency (η_{PA}). Losses between the antenna and PA are known as feeder losses (σ_{feed}) which may be ignored since PAs are placed close to the antennas [139].

• The loss factors of DC-DC, AC-DC conversions, main supply units (MS), and cooling power consumption for the BBU pool are approximated by $\sigma_{DC,POOL}$, $\sigma_{MS,POOL}$, and $\sigma_{COOL,POOL}$. Whereas for the RRHs, the loss factors are approximated by $\sigma_{DC,RRH}$ and $\sigma_{MS,RRH}$. Moreover, the optical fibre losses between BBUs and RRHs are approximated by a loss factor $\sigma_{optical}$.

• Power consumption of the optical transceivers linearly scales with the number of BBUs and RRHs.

If the power consumed by a single BBU serving a single RRH is:

$$P_1 = P_{BBU} + P_{RRH} \tag{3.20}$$

$$P_1 = \frac{A(\frac{W}{BW_{TOTAL}})P^{pm}_{BBU} + P_{TRANS_B}}{(1-\sigma_{DC,POOL})(1-\sigma_{MS,POOL})(1-\sigma_{COOL,POOL})} + \frac{A(\frac{W}{BW_{TOTAL}})P^{pm}_{RF} + (P_{max}/A.\eta_{PA}) + P_{TRANS_R}}{(1-\sigma_{DC,R})(1-\sigma_{MS,R})(1-\sigma_{optical})} \tag{3.21}$$

Then the total power consumed by all active BBUs and RRHs in a C-RAN network can be modelled as

$$P_{supply} = \sum_{m=1}^{M} \left(P_{BBU} + \sum_{n \in Z_m} P_{RRH} \right) \tag{3.22}$$

where M represents the number of active BBUs in the network and Z_m represents the list of RRHs handled by BBU_m

3.6 Cell Differentiation and Integration (CDI) Algorithm

According to the intuitive analysis above, a CDI algorithm is proposed in this section and Fig.3.4. Network information is collected in the first step and analysed for proper cell differentiation and integration. The algorithm seeks to utilise the network resources efficiently by calculating the necessary number of BBUs and RRHs to serve capacity demands at the end of each CDI cycle. Apart from a single BBU required to serve load requirements, proper BBU-RRH configuration is adjusted at the end of optimisation step by comparing the analysed and optimised QoS values.

Note that the QoS metrics can be different depending on load intensity and the number of active BBUs and RRHs in the network. For the optimisation part of the algorithm, a DPSO algorithm is developed as an EA to solve the BBU-RRH configuration problem and is explained in the next section. The optimisation process continues until the CDI cycle is completed. Note that, the CDI algorithm shown in Fig.3.4 is triggered at the beginning of each CDI cycle.

The pseudo-codes for semi-static cell differentiation and integration are given in Algorithm 1 and Algorithm 2, respectively. An important consideration is the first

Algorithm 1: Pseudo-code for Semi-static Cell Differentiation

Input : Current network load $\eta(t)$ from (3.7)
BBU-RRH mapping vector **r**
Required number of BBUs from (3.8)

if *No. of active BBUs =1* **then**
  if $\eta(t) \geq |P_{\text{RB}}|$ **then**
    -Activate required No. of BBUs
    -Differentiate cell into tier-2 RRH structure by BBU-RRH mapping using Algorithm 4
    -Update BBU-RRH mapping vector **r**
    for *i=1 to C* **do**
      -Select set S_i
      -Compute $\eta_{\text{RRH}_{i1}}(t)$ from (3.9)
      if $\eta_{\text{RRH}_{i1}}(t) > |P_{\text{RB}}|$ **then**
        - R$\leftarrow S_i$ {Add S_i to R}
        -Diffentiate cell C_i by activating all RRHs in S_i and map them to BBUs according to Algorithm 4.
        -Update BBU-RRH mapping vector **r**.
      end
    end
  else
    -No cell differentiation required.
    -Tier-3 RRH structure remains.
  end
else
  if *No. of active BBUs* $\leq$ *No. of required BBUs* **then**
    if *All possible RRHs deployed in the network are active* **then**
      -Activate the required No. of BBUs.
      -Cells can not be differentiated further.
      -Update BBU-RRH mapping vector **r**
    else
      -Activate required number of BBUs
      for *i= 1 to C* **do**
        -Select set S_i
        -Compute $\eta_{\text{RRH}_{i1}}(t)$ from (3.9)
        if $\eta_{\text{RRH}_{i1}}(t) > |P_{\text{RB}}|$ **then**
          - R$\leftarrow S_i$ {Add S_i to List R}
          -Differentiate cell C_i further to tier-1 RRH structure by mapping newly activated RRHs to active BBUs using Algorithm 4
          -Update BBU-RRH mapping vector **r**
        end
      end
    end
  end
end

Algorithm 2: Pseudo-code for Semi-static Cell Integration

```
Input : Current network load η(t) from (3.7)
        BBU-RRH mapping vector r
        Required number of BBUs from (3.8)
if No. of active BBUs =1 then
    -No cell integration required.
    -A high-power BS serves the geographical area.
else
    if No.of required BBUs=1 then
        -Integrate all cells into tier-3 RRH structure, i.e., a high power BS should
         serve the geographical area.
        -Switch-off remaining BBUs.
        -Update BBU-RRH mapping vector r.
    else
        for i=1 to C do
            -Select set S_i
            for j=1 to end of S_i do
                -Compute load η_RRH_ij(t) from (3.9)
                -Sum=Sum+η_RRH_ij(t)
            end
            if Sum ≤ P_RB then
                -Integrate all cells by switching-off all RRHs in set S_i except
                 RRH_i1.
                -Offload RRHs to required number of BBUs according to
                 Algorithm 5.
                -Update BBU-RRH mapping vector r.
            end
        end
        -Run Algorithm 5 {Case of BBU reduction and no integration}
    end
end
```

association of RRHs to the required number of BBUs during cell differentiation and integration, before identifying a proper BBU-RRH mapping in the optimisation phase. Algorithm 4 and 5 are supporting algorithms for Algorithm 1, and 2 which covers all possible cases of initial BBU-RRH assignment during differentiation or integration of cells along with cases where the number of BBUs are increased, decreased or remain unchanged. The initial BBU-RRH mapping is necessary for utilising the available BBU resources in an efficient manner so as to prevent high blocking rate. The blocking rate of

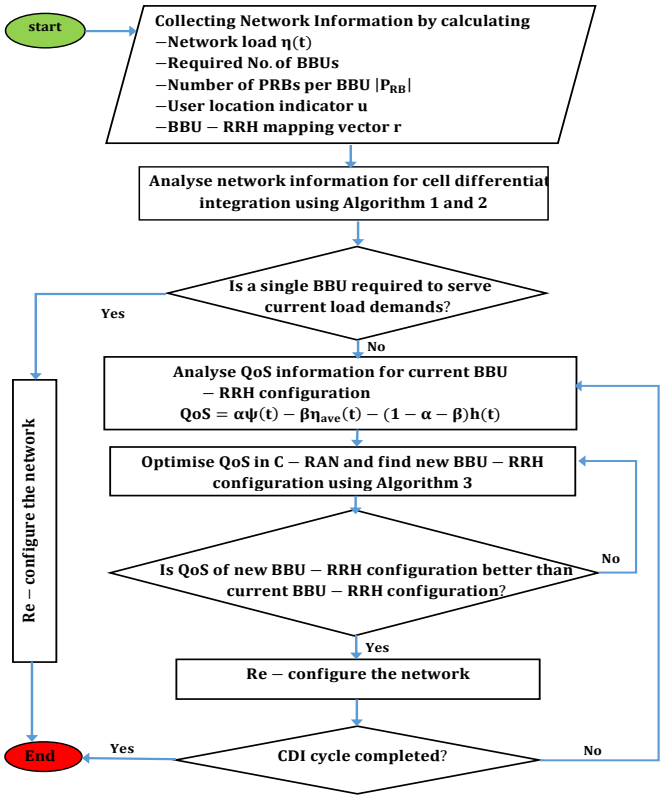


Fig. 3.4 Block diagram of CDI Algorithm for one CDI cycle.

the network at time t can be measured as

$$\text{Blocking rate} = \left[1 - \frac{\sum_{m=1}^{\mathrm{M}} \sum_{k=1}^{\mathrm{K}} \mathrm{I}_{m,k}(t)}{\mathrm{K}}\right] \times 100 \tag{3.23}$$

where $\mathrm{I}_{m,k}(t)$ as discussed earlier, is a binary indicator such that $\mathrm{I}_{m,k} = 1$, if user k is served by BBU_m at time t. Note that, users are served based on the choice of scheduler used by a BBU. Moreover, the amount of resource shortage (or PRB shortage) in the network based on users PRB demand can be estimated as follows

$$\text{Resource Shortage} = \sum_{m=1}^{\mathrm{M}} \max\Big[(\eta_m(t) - 1), 0\Big] \times 100 \tag{3.24}$$

Note that, the CDI algorithm triggers Algorithm 1 and Algorithm 2 sequentially, i.e., Algorithm 2 is triggered immediately after the Algorithm 1 is executed. In the interest of simplicity and understanding, the CDI algorithm is divided into separate pseudo-codes.

Algorithm 3: Initial RRH association to active BBUs during cell differentiation.

Input : List A of newly activated BBUs
List R containing sets of RRHs supporting cell differentiation

if *A is not empty* **then**
  for *m=1 to No. of active BBUs* **do**
    -Compute $\eta_m(t)$ from (3.6)
    if $\eta_m(t) \leq$ *lower limit* **then**
      A$\leftarrow$ BBU$_m$\{Add BBU$_m$ to List A\}
    end
  end
  I=1;
  while *not the end of List R* **do**
    -Select Ith set from list R
    $m = 1$;
    for *j=1 to end of set* S_i **do**
      if $m > |A|$ **then**
        $m = 1$
      end
      $BBU_m \leftarrow RRH_{ij}$\{Map RRH_{ij} to BBU_m except R_{1j}\}
      $m = m + 1$;
    end
    I=I+1;
  end
else
  for *m=1 to No. of active BBUs* **do**
    -Compute $\eta_m(t)$ from (3.6)
    if *lower limit* $\leq \eta_m(t) \leq$ *Upper limit* **then**
      A$\leftarrow$ BBU$_m$\{Add BBU$_m$ to A\}
    end
  end
  if *A is still empty* **then**
    A $\leftarrow$ All active BBUs
  end
  -Sort A in increasing order of BBU loads
  I=1;
  while *not the end of List R* **do**
    -Select Ith set from List R
    $m = 1$;
    for *j=1 to end of set* S_i **do**
      if $m > |A|$ **then**
        m=1;
      end
      BBU$_m \leftarrow$ RRH$_{ij}$\{Map RRH$_{ij}$ to BBU$_m$ except RRH$_{1j}$\}
    end
    I=I+1;
  end
end

Algorithm 4: Initial RRH association to active BBUs during cell integration.

```
Input : List A of No. of active BBUs
        No. of required BBUs
        BBU-RRH mapping vector r
if No. of required BBUs < |A| then
    for m=1 to |A| do
        for i=1 to C do
            for j=1 to c do
                -Select RRH_ij from BBU-RRH vector r
                if RRH_ij = m then
                    Z_m ← RRH_ij
                    {Z_m is a List of RRHs handled by BBU_m}
                end
            end
        end
    end
    -Sort List A in decreasing order of BBU loads
    for m=1 to end of List A do
        if m ≠ |No. of required BBUs| then
            Ā ← BBU_m{Ā is a List of required BBUs}
        else
            B̄ ← BBU_m
            {B̄ is a List of BBUs to be switched off}
        end
    end
    -Sort List Ā in increasing order of BBU loads
    for i=1 to end of List B̄ do
        -Select i^th BBU from List B̄
        -Select List Z_m of the i^th BBU
        m=1;
        for j=1 to end of Z_m do
            if m> |Ā| then
                m=1;
            end
            -Select RRH at j^th index in List Z_m
            -Select BBU at m^th index of List Ā
            -BBU_m ← RRH^j{Assign RRH j to BBU m}
            m++
        end
    end
    -Switch off all BBUs in List B̄
end
```

3.6.1 Discrete Particle Swarm Optimisation (DPSO)

Particle Swarm Optimisation (PSO) is a robust optimisation technique inspired by social behaviour of flocking organisms. PSO method uses *Swarm Intelligence* for solving global optimisation problems [140]. PSO utilises a population (or swarm) of particles, where each particle represents a solution, namely a BBU-RRH association vector $\mathbf{r}$. As the QoS represented in eq. (3.14) is considered as the main objective function (or fitness function), PSO seeks to maximise the QoS function by finding the best solution vector $\{\mathrm{r}_{11}, \mathrm{r}_{12}, ..., \mathrm{r}_{in}\}$. PSO operates on a group of particles (or solutions) to probe the solution space in a random way with different velocities. The vector $\{\mathrm{r}_{11}, \mathrm{r}_{12}, ..., \mathrm{r}_{in}\}$ is viewed as the particle position in the n-dimensional solution space while discovering the optimal solution can be viewed as particles probing the solution to search for the optimum position. To direct the particles to their best fitness values, the velocity of each particle is changed stochastically at each iteration. The velocity update of each particle j depends on the historical best position experience (pbest) of the particle itself and the best position experience of neighbouring particles, i.e., the global best position (gbest). Therefore, every particle in the swarm tends to direct itself towards the best solution at each iteration. Since the solution vector $\mathbf{r}$ is real-valued, the standard PSO algorithm can not be applied directly to solve the discrete optimisation problem. Therefore, a Discrete PSO is developed to solve the QoS maximisation problem defined in eq. (3.14). The DPSO algorithm is described in Fig. 3.5 and the following steps:

Step 1: Generate initial population $\mathbf{R}^0$ with population/swarm size of $|\Delta|$. Where $\mathbf{R}^0$ consists of N-bit particles (BBU-RRH mapping solutions). Where $\mathbf{N}$ is taken according to the number of active RRHs in the network and the superscript 0 represents the initial iteration number $\mathbf{I} = 0$. The best position of each particle $\mathbf{pbest}_{\mathrm{j}}^0 = \mathbf{r}_{\mathrm{j}}^0, 1 \leq \mathrm{j} \leq |\Delta|$ are initialised with a random velocity of $\mathbf{V}_{\mathrm{j}}^0$ for each particle.

Step 2: Calculate the fitness values for each particle (BBU-RRH mapping solution) in the current swarm/population using the fitness function F defined as QoS in eq. (3.14) and identify the global best position achieved i.e., $\mathbf{gbest}^0 = \underset{1 \leq \mathrm{j} \leq |\Delta|}{\operatorname{argmax}} \mathrm{F}(\mathbf{pbest}_{\mathrm{j}}^{\mathrm{I}})$.

Step 3: Update particle j position by updating its velocity. The velocity update equation is given as

$$\begin{aligned} &\mathbf{v}_{\mathrm{j}}^{\mathrm{I}} = w\mathbf{v}_{\mathrm{j}}^{\mathrm{I}-1} + c_1\varepsilon_1\left(\mathbf{pbest}_{\mathrm{j}}^{\mathrm{I}} - \mathbf{x}_{\mathrm{j}}^{\mathrm{I}}\right) + c_2\varepsilon_2(\mathbf{gbest}_{\mathrm{j}}^{\mathrm{I}} - \mathbf{x}_{\mathrm{j}}^{\mathrm{I}}) \\ &1 \leq \mathrm{j} \leq |\Delta| \end{aligned} \tag{3.25}$$

where $\mathbf{x}_j^{\mathrm{I}}$ is the current position of particle j in iteration I and $\varepsilon_1, \varepsilon_2$ are random numbers chosen between the range $[0-1]$. Both c_1 and c_2 are acceleration constants that pulls the particle towards best position. Values in the range 0-5 are chosen for c_1 and c_2. The inertial weight w represents the effect of preceding velocity on the updated velocity. Larger and smaller value of w are used for global exploration and local search expedition in the search-space, respectively. However, choosing an optimum value for w can assist a balanced proportion between global and local exploration of the search space. Usually values between 0-1 are selected for w [141]. A value of 0.9 for w is selected in this study. The new position of particle j for the next iteration $\mathrm{I}+1$ will be:

$$\mathbf{x}_{\mathrm{j}}^{\mathrm{I}+1} = \mathbf{x}_{\mathrm{j}}^{\mathrm{I}} + \mathbf{v}_{\mathrm{j}}^{\mathrm{I}} \tag{3.26}$$

Step 4: Update the iteration counter $(\mathrm{I} = \mathrm{I}+1)$. If the convergence criteria is satisfied then end else go to step 5.

Step 5: Update particle j's personal best position as

$$\mathbf{pbest}_{\mathrm{j}}^{\mathrm{I}} = \begin{cases} \mathbf{pbest}_{\mathrm{j}}^{\mathrm{I}-1} & \text{if} \quad \mathrm{F}(\mathbf{r}_{\mathrm{j}}^{\mathrm{I}}) \leq \mathrm{F}(\mathbf{pbest}^{\mathrm{I}-1}) \\ \mathbf{r}^{\mathrm{I}-1} & \text{if} \quad \mathrm{F}(\mathbf{r}_{\mathrm{j}}^{\mathrm{I}}) > \mathrm{F}(\mathbf{pbest}_{\mathrm{j}}^{\mathrm{I}-1}) \end{cases} \tag{3.27}$$

Step 6: Update global best position achieved by:

$$\mathbf{gbest}^{\mathrm{I}} = \begin{cases} \underset{1\leq j\leq|\Delta|}{\operatorname{argmax}}\, \mathrm{F}(\mathbf{pbest}_{\mathrm{j}}^{\mathrm{I}}) & \text{if}\, \mathrm{F}(\mathbf{pbest}_{\mathrm{j}}^{\mathrm{I}}) > \mathrm{F}(\mathbf{gbest}^{\mathrm{I}-1}) \\ \mathbf{gbest}^{\mathrm{I}-1} & \text{otherwise} \end{cases} \tag{3.28}$$

Step 7: Repeat all steps starting from step 1.

Algorithm 5: Discrete Particle Swarm Optimisation (DPSO) Algorithm

-I=0;
-Generate initial Swarm with random position and velocity $(r_1^I, r_2^I, ..., r_{|\Delta|}^I)$.
for $i = 1$ *to* $|\Delta|$ **do**
    {Initialise best positions of each particle}
    pbest_i=r_i^I
end
-QoS=0;
-{Find QoS of each particle using (3.14) i.e.,}
for *i=1 to* $|\Delta|$ **do**
    -Select r_i^I from pbest_i
    if $f(r_i^I) > QoS$ **then**
        QoS=$f(r_i^I)$
    end
end
gbest=QoS;
-Update positions and velocities of all particles in the swarm using equation (4.15) and (4.16)
while $I < I_{max}$ **do**
    for *i=1 to* $|\Delta|$ **do**
        -Select r_i^I from the swarm
        F_1=$f(r_i^I)$
        -Select r_i^I from pbest
        $F_2 = f(r_i^I)$
        if $F_1 \geq F_2$ **then**
            $\text{pbest}_i = r_i^I$
            QoS=F_1
            if *QoS>gbest* **then**
                gbest=QoS;
                $x = i$;
            end
        end
        -Update positions and velocities of all particles in the swarm using equation (3.25) and (3.26)
    end
    I=I+1;
end

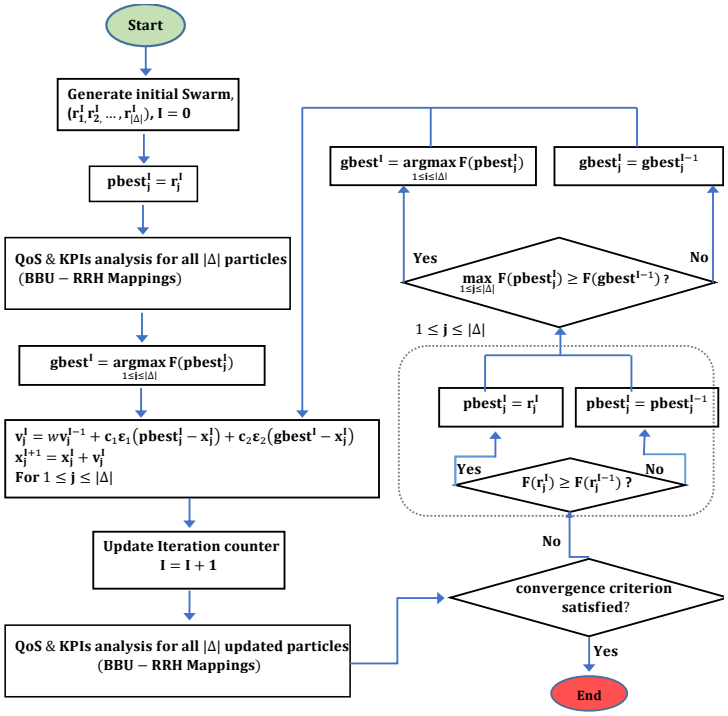


Fig. 3.5 Block diagram of DPSO algorithm.

3.7 Computational Results and Analysis

To make the simulation more realistic, the user arrivals in Fig.3.7 follows a Poisson process with rate λ. However, due to the dynamic spatial and temporal nature of user traffic, the user arrival is modelled as a time-inhomogeneous process. This is achieved by multiplying the time-homogeneous Poisson process with traffic intensity parameter λ and the rate function $f(t)$ shown in Fig.3.6. The rate function is unit-less and reshapes the traffic from constant intensity to an analogous time varying profile that reflects typical traffic patterns in a real cellular network. If users arrive in the system following a Poisson process with intensity λ users/min, with a constant service time of h (60 sec), then the number of users at time t is calculated as $\mathrm{K}(t) = \chi h f(t)$. Where $\chi \sim \mathrm{Poiss}(\lambda)$ is a random variable with mean λ (i.e., $\lambda = 200$).

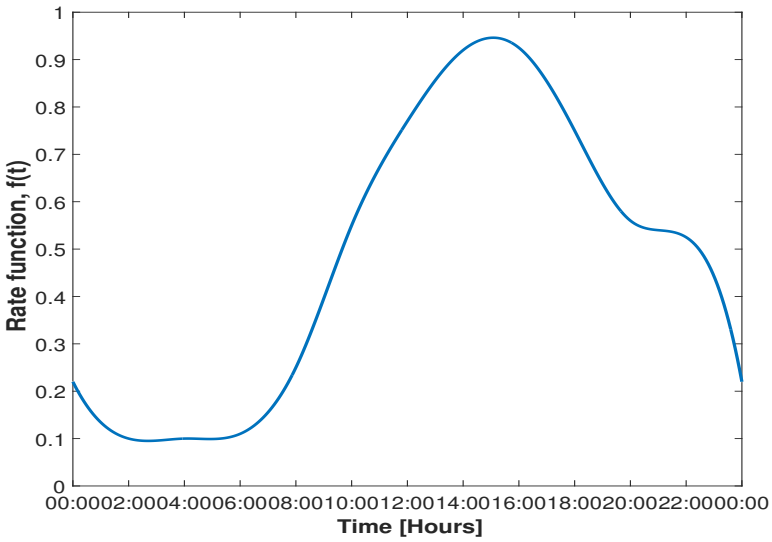


Fig. 3.6 Rate function for time in-homogeneous user arrivals.

Moreover, different data rate requirements are assumed for end users based on 3rd Generation Partnership Project (3GPP) standard simulation parameters [142] i.e., 4-25 kbps for audio, 32-384 kbps for video, 28.8 kbps for data, and 60 kbps for real-time gaming services. Based on uniform user distribution and network load shown in Fig.3.7, an actual number of active BBUs and RRHs with respect to time is shown in Fig.3.8.

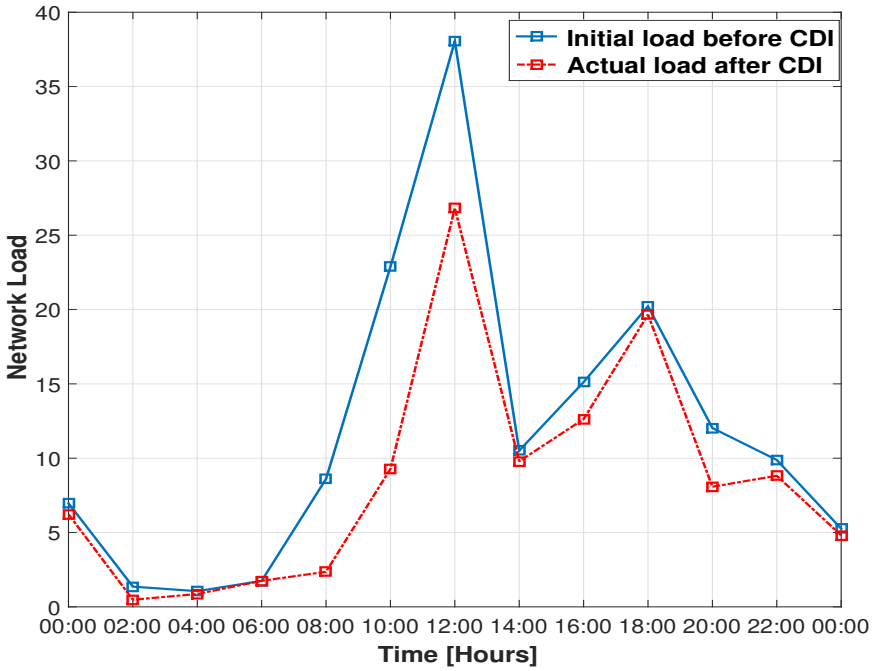


Fig. 3.7 Actual network load with respect to time.

The BBU-RRH association vector $\mathbf{r} = \{r_{11}, r_{12}, ..., r_{in}\}$ is maintained and updated after each CDI cycle. Newly activated RRHs and BBUs in the network are mapped according to Algorithm 4 and 5. In this work, a maximum of 49 RRHs and 5 BBUs are deployed in the network to support semi-static cell differentiation and integration. The initial BBU-RRH mapping at the beginning of a CDI cycle might degrade the network QoS. Therefore, dynamic BBU-RRH mapping is proposed to identify proper BBU-RRH mapping.

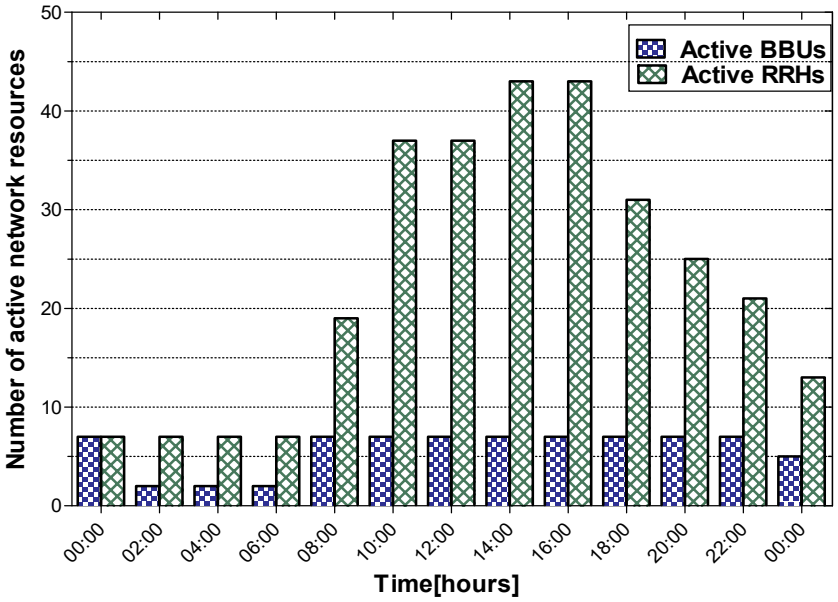


Fig. 3.8 Number of active BBUs and RRHs with respect to network load/time.

Before going to a more thorough analysis of the proposed CDI concept, the efficiency of DPSO over two different problem scenarios, P_1, P_2, and compared with ES algorithm. Both scenarios consists of 5 active BBUs with 19 active RRHs including two differentiated cells (Tier 1, level 2, RRH structure) for P_1, and 49 active RRHs (Tier 1, level 7, RRH structure) for P_2, respectively. The aim is to analyse DPSO performance for small and large networks. User distribution within each cell is uniform where 6 and 25 users are considered for non-dense and high dense cells, respectively.

For Monte Carlo analysis, the DPSO and ES algorithms are repeated 50 times with different initial BBU-RRH settings for each problem and results obtained are averaged. The load fairness index, averaged network load, and handover index are represented in

Fig.3.10, Fig.3.11 and Fig.3.12 , respectively, over 200 iterations for both P_1 and P_2. The optimum values shown in the figures and Table.3.2, are achieved by exhaustively searching for all possible solutions M^N) using ES algorithm, which helps in demonstrating the improvement in each iteration of the DPSO algorithm. Note that, ES algorithm is independent of iterations.

Table 3.2 Computational Results for DPSO and ES

		P_1 (19 RRH)	P_2 (49 RRH)
Quality of Service	DPSO	0.599142	0.588970793
	ES	0.599142	0.592940793
Load Fairness Index	DPSO	0.984797	0.97168
	ES	0.984797	0.97556
Average Network Load	DPSO	1.506	1.4962
	ES	1.506	1.4663
Handover Index	DPSO	0.38095	0.38748
	ES	0.38095	0.383659

Fig.3.9 shows that the DPSO algorithm converges to the optimum solution in P_1 with a Convergence Rate (CR) of 0.825. CR is defined as the number of times the DPSO finds a best or optimum solution during the entire number of iterations. This implies that over 200 iterations, the optimum solution is achieved 165 times for P_1. For P_2, the CR of DPSO algorithm is 0.12. However, the optimum solution is not reached over 200 generations. DPSO algorithm achieves the best value 24 times i.e., after 176 iterations and $176\times|\Delta|$ fitness evaluations, which is still 99.53% of the optimum value achieved by ES algorithm after an enormous 5^{49} (M^N) fitness evaluations.

Fig.3.10 shows that the DPSO algorithm converges to the optimum load fairness index value after 13^{th} iteration in P_1. However, in P_2, the optimum value can not be found over 200 iterations, and the best load fairness index value is achieved after only 176 iterations and $176\times|\Delta|$ fitness evaluations, which is 99.57% of the optimum value found by ES algorithm. ES algorithm performs 5^{49} fitness evaluations to find the optimum value which is a considerable amount of fitness evaluations.

Figs. 3.11 and 3.12 displays the convergence of DPSO algorithm to the optimum value for average load value and handover index in both P_1 and P_2 . In P_1, optimum are achieved

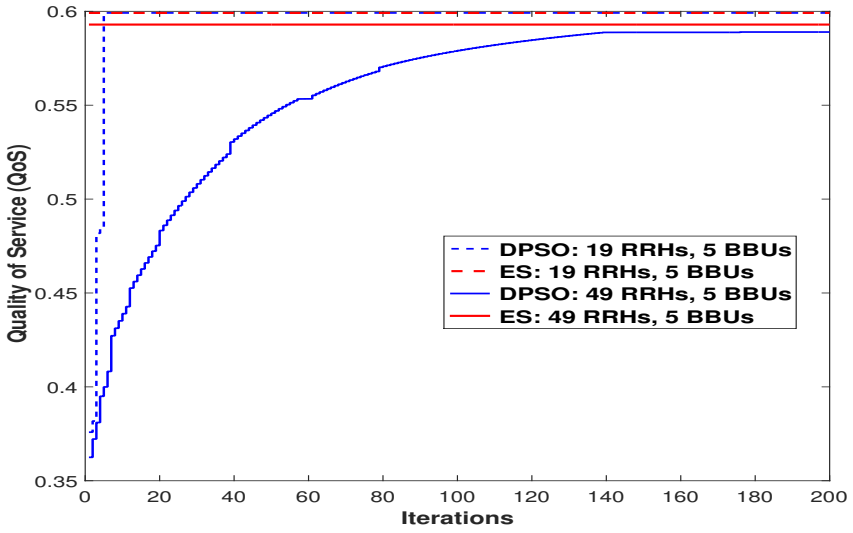


Fig. 3.9 QoS values for DPSO and ES.

after 12 and 38 iterations for average network load and handovers, respectively. For P_2, the DPSO algorithm could not find the optimum value over 200 iterations. However, the best possible value achieved for average network load and handover index are 98% and 99.01% of the optimum value found by ES algorithm, respectively. ES algorithm determines the optimum value after performing 5^{49} enormous fitness evaluations whereas the DPSO algorithm performs $67\times|\Delta|$ and $145\times|\Delta|$ to find the best value for average network load and handover index, respectively.

Note that, the α and β control parameters in eq. (3.14) are selected by performing an ES algorithm to identify the optimal BBU-RRH setting for P_1. Both α and β values are orderly set to 0, 0.1, ..., 1 with a constraint $\alpha+\beta\leq 1$ as shown in Fig.3.13. An optimal BBU-RRH setting is found using ES algorithm for each pair of α and β. It is observed that setting a higher value for load fairness index (until $\alpha=0.8$) not only reduces the resource shortage but also improves network balance. Setting values for $\alpha>0.8$ results into improper BBU-RRH mapping which implies that maximising network load balance is overly considered compared to minimising average network load and handovers, resulting into an increased resource shortage. This work considers $\alpha=0.8$ and $\beta=0.1$ which means assigning a 10% weight to handover minimisation.

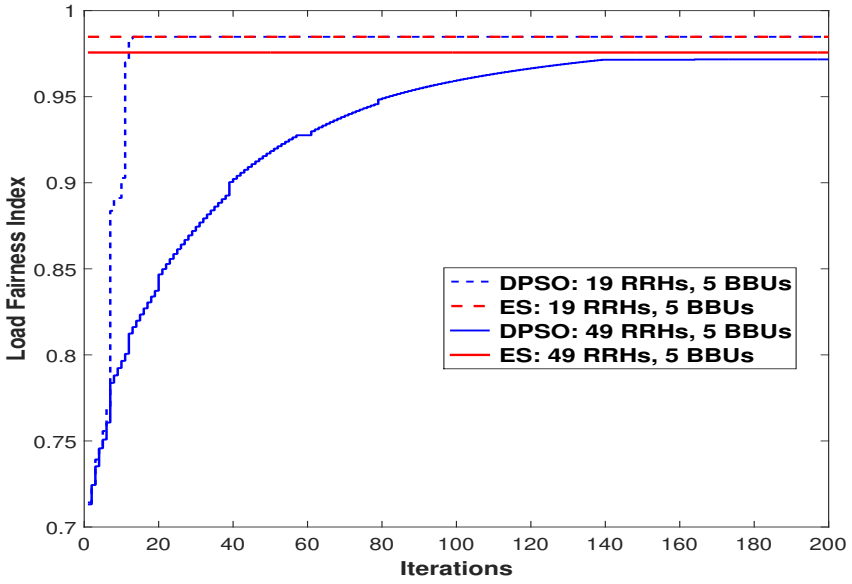


Fig. 3.10 Load fairness index values for DPSO and ES.

For a more thorough analysis, the proposed CDI concept is compared to a fixed C-RAN scenario (F-CRAN). The BBU cloud holds five BBUs in both cases. However, the fixed C-RAN scenario does not support cell differentiation or integration, and only 7 RRHs serves the entire macrocell coverage area. The dynamic BBU-RRH mapping is enabled in the fixed C-RAN scenario which shows 5^7 possible BBU-RRH mapping solutions to choose from at the beginning of each CDI cycle. The number of possible BBU-RRH mapping solutions for CDI scenario at the start of each CDI cycle is M^N, where M and N represents the number of active BBUs and RRHs, respectively. Moreover, an increasing user arrival is considered in the network with random data rates requirement as explained earlier. However, a Monte-Carlo analysis is performed, where 100 uniformly distributed users are envisaged for each instance, and the average of all distributions are taken into account regarding network load, throughput, blocked users, and resource shortage analysis. Figs.3.14 and 3.15 shows the relative performances regarding average blocked users and average network throughput with Proportional Fair (PF) and Round Robin (RR) scheduling techniques. Since the CDI algorithm includes 2 phases of BBU-RRH mapping, i.e., the initial BBU-RRH assignment during cell integra-

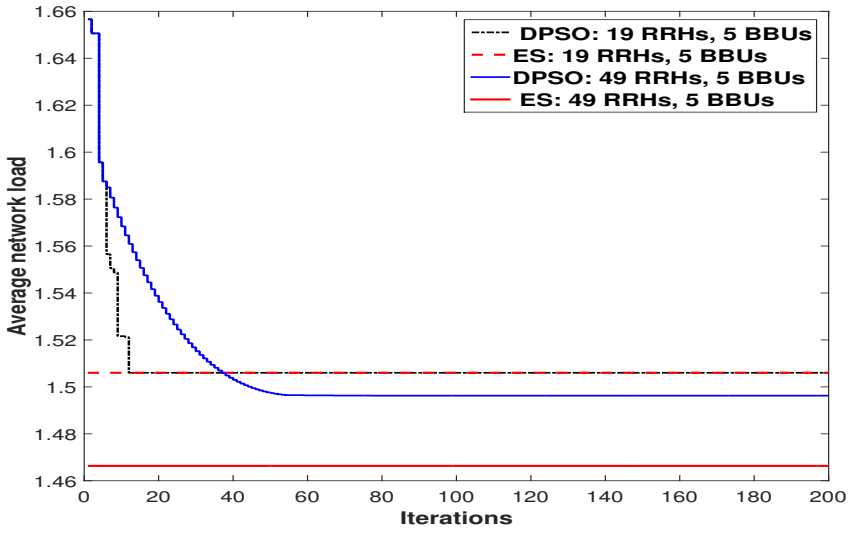


Fig. 3.11 Average network load for DPSO and ES.

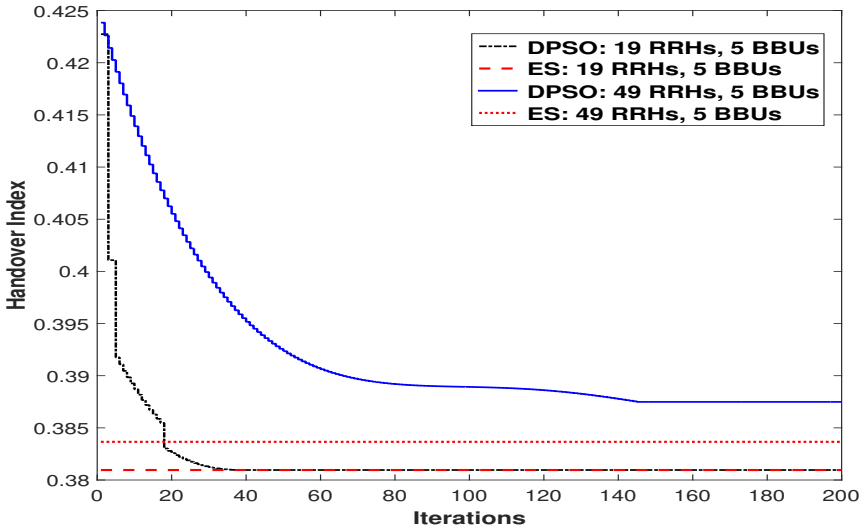


Fig. 3.12 Average handovers for DPSO and ES.

tion/differentiation and the optimum BBU-RRH setting achieved by DPSO in the second step, the results of both phases are analysed.

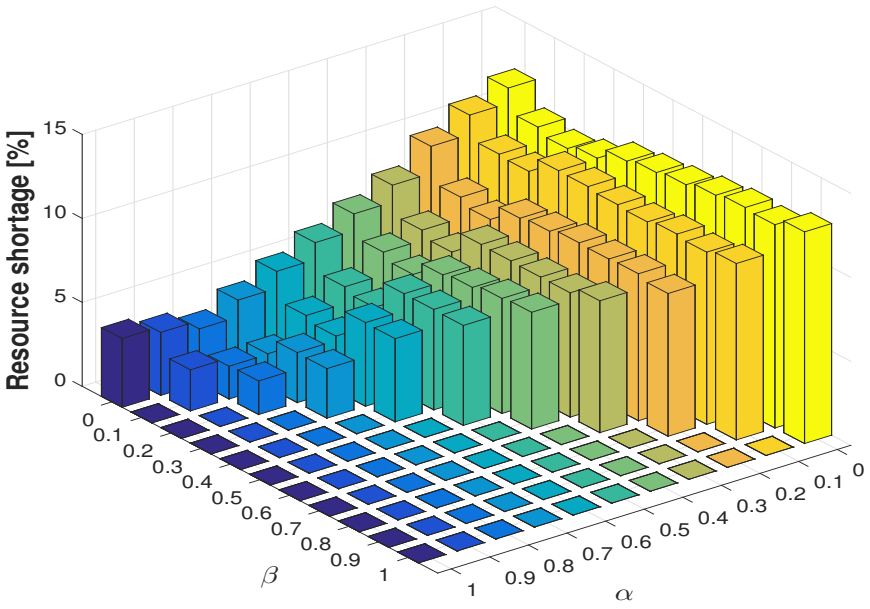


Fig. 3.13 Resource shortage for different α and β.

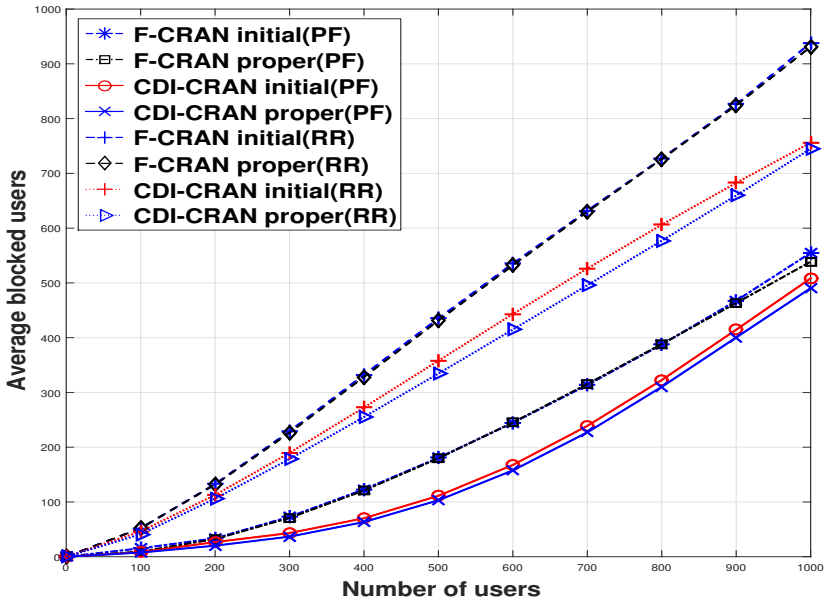


Fig. 3.14 Average blocked users for fixed and CDI-enabled C-RAN.

The simulation results demonstrate the advantage of using CDI-enabled C-RAN (CDI-CRAN) instead of a F-CRAN setting. When a fixed C-RAN is considered, the average blocked users in the network are much higher with significantly lower average throughput, using any scheduling technique, as shown in Figs. 3.14 and 3.15, provided

that the dynamic BBU-RRH mapping is also enabled. However, an interesting observation is the significant drop in the averaged blocked users and the necessary increase in average network throughput in CDI-CRAN compared to F-CRAN. This indicates that during cell differentiation, an overloaded cell divides into multiple smalls cells, and not only reduces the user to RRH distances but also the PRB demands resulting from high SINR and low path loss values. A further decrease in average network load is observed after proper BBU-RRH mapping, providing a balanced network load across the active BBUs. Note that, cell differentiation increases the number of RRH interferers in the network. However, RRHs served by the same BBU does not contribute to the overall interference experienced by users served by the same BBU.

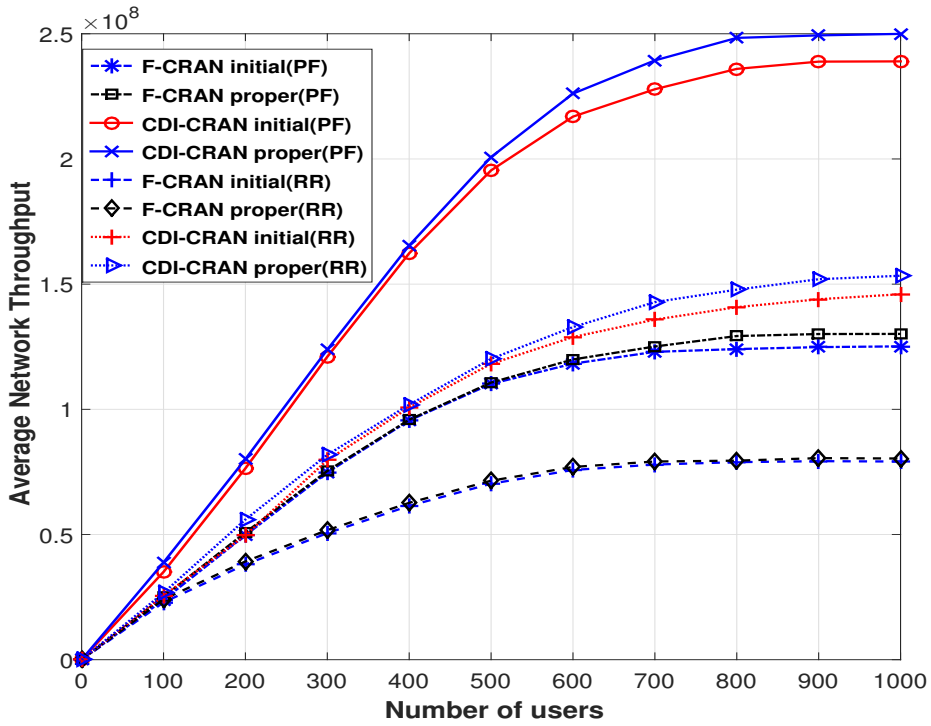


Fig. 3.15 Average network throughput for fixed and CDI-enabled C-RAN.

From the results shown in Fig. 3.15, it is observed that the average network throughput increases by 45.53% in the CDI-CRAN compared to F-CRAN, both enabled with PF schedulers. Whereas with RR schedulers, an increase of 42.102% is observed. Moreover, the average throughput difference between initial and optimum BBU-RRH mapping in a CDI-enabled C-RAN, with PF and RR scheduling is 4.0219% and 4.126%, respectively.

This indicates efficient resource utilisation during cell differentiation and integration, ensuring minimum blocked users until a proper BBU-RRH setting is identified. Note that, the initial BBU-RRH mapping supported by Algorithm 4 and 5, is an important consideration in the overall CDI concept (as explained earlier). About 23.149% reduction in the average number of blocked users with PF scheduler and 20.903% with RR is observed in Fig.3.14. Moreover, the average resource shortage drastically decreases in the CDI-CRAN, compared to fixed C-RAN as shown in Table.3.3 provided that both scenarios have an equal amount of resources available (i.e., 5 BBUs, 5×100 PRBs). A 76.57% decrease in average PRB shortage is estimated with CDI-CRAN compared to F-CRAN.

Table 3.3 Comparison results for fixed and CDI-enabled C-RAN

	Blocking Rate[%]				Resource Shortage	
	Initial		Proper		Initial	Proper
	PF	RR	PF	RR		
F-CRAN	35.99	81.66	35.34	81.10	16.44×10^3	16.42×10^3
CDI-CRAN	26.87	67.33	25.809	62.13	38.79×10^2	38.47×10^2

Fig.3.16 shows the average power consumed by the C-RAN network for both CDI and fixed setting for different schedulers. Despite the fact that the geographical area is served with more RRHs in CDI-CRAN, the total power consumed by the network is still lower. An average decrease of $\approx 15.28\%$ and $\approx 16.02\%$ in the average power consumption is estimated in the CDI-CRAN with PF and RR schedulers, respectively, compared to a fixed C-RAN setting. Fig.3.17 shows the SINR thresholds versus the probability of coverage results. The CDI activated C-RAN performs well compared to fixed C-RAN regarding coverage performances for 1000 users, provided that no interference mitigation techniques are applied in this work. Note that, the probability of coverage still has a linear behaviour for SINR thresholds within the range of 0dB to 5dB for both cases.

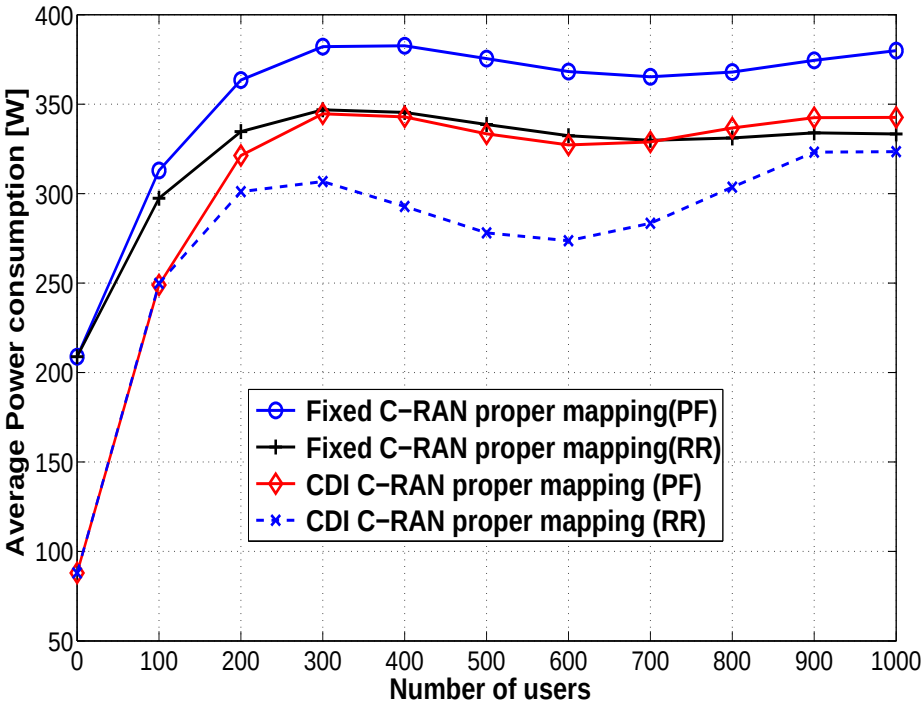


Fig. 3.16 Average power consumed by fixed and CDI-enabled C-RAN.

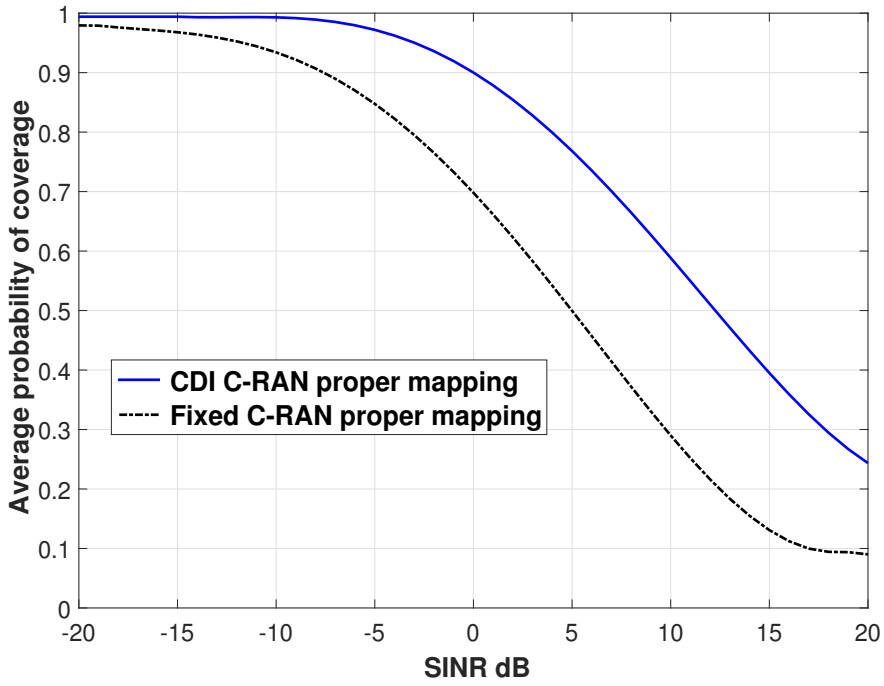


Fig. 3.17 Probability of coverage versus SINR threshold.

3.8 Summary

The concept of cell differentiation and integration in C-RAN is examined with an objective to utilise network resources efficiently without degrading the overall network QoS. An energy efficient C-RAN network is considered to accommodate traffic by scaling the BBUs and RRHs as well as maintaining a balanced network via proper BBU-RRH mapping, formulated as a constrained integer programming problem. A CDI algorithm is developed for C-RAN and tested for comparison with a fixed C-RAN setting. Computational results based on Monte Carlo analysis shows an average throughput increase of 45.53% with 23.149% decrease in average blocked users and 76.57% reduction in average resource (PRBs) shortage. The CDI algorithm hosts a DPSO algorithm which is developed to find optimum BBU-RRH configuration dynamically. The performance of DPSO is tested and compared to ES. Using two benchmark problems, the DPSO delivered noticeably faster convergence compared to ES, which makes the CDI algorithm more reliable for a self-organised C-RAN. It could be noticed from the results that the network performance with PF scheduler is better than RR scheduler in terms of throughput and blocked users. RR scheduler guarantees fairness between the users as it schedules all the users regardless their channel condition which would be on the expense of network throughput. On the other hand, PF scheduler maximise the total network throughput while at the same time allowing all users at least a minimal level of service. Moreover, the power model for C-RAN is proposed to estimate the overall network power consumption. It is noticed that despite deploying a higher number of RRHs (49) in a given geographical area for CDI enabled C-RAN, the power consumption of a fixed C-RAN for the same geographic area is still higher by $\approx$ 15.28% and 16.02% for PF and RR schedulers, respectively.

Chapter 4

Load Aware Self-Optimised 5G Network Exploiting Millimetre-Wave NR and C-RAN Architecture

The Fifth Generation (5G) of cellular networks will integrate and enhance the existing wireless networks. Regarding the 5G wireless access technology, known as New Radio (NR), its preliminary deployment scenario will be a Non-Standalone (NSA) operation, where the existing Fourth-Generation (4G) infrastructure will support the 5G networks. Very high capacity is the common requirement of the 5G usage scenarios, which involve enhanced Mobile BroadBand (eMBB), Ultra-Reliable Low Latency Communications (URLLC) and massive Machine-Type Communications (mMTC). Furthermore, key enablers of 5G, to realise this goal, are operating in the wide band Millimetre Wave (mmw) frequency and the Cloud Radio Access Network (C-RAN) architecture. C-RAN enables centralised efficient resource management as well as coordinated joint transmission between small cells, to overcome the limited coverage of mmw. The high-capacity requirement of 5G applications will bring higher fluctuation in traffic load due to the spatio-temporal variation of mobile users' data demand. In this chapter, a self-optimised C-RAN is proposed, which dynamically adapts to the varying capacity demands by interworking between mmw and μw bands for efficient resource utilisation. At high traffic conditions, when a macrocell's resource usage reaches its limit, it is divided into compact clusters of mmw small cells to increase capacity. When traffic

conditions are low, the small cells are reintegrated again into a macrocell based on the Cell Differentiation and Integration (CDI) technique. The CDI technique performance is examined for C-RAN mmw and μw small cells, being subsequently compared to a fixed C-RAN.

4.1 Introduction

According to the mobile traffic analysis shown by Cisco [143], the mobile data traffic is observed to have grown exponentially by 18 folds over the past five years. Moreover, this mobile traffic tsunami has a compound annual increase rate of 66%. The rapid increase in wireless devices along with bandwidth-hungry internet applications are the prime drivers of such drastic growth of wireless communication demands and has pushed the existing system capacity demand beyond its limits. To cope with ever-increasing data traffic, it is of serious interest for Mobile Network Operators (MNOs) to focus on commercial wireless communication technologies since the current wireless spectrum is almost saturated. A promising approach for additional spectrum that has the potential for 1000 fold increase in data rates, capacities, frequency-range, as well as latencies of less than 1 ms for all connected devices is mmw communication [144]. 5G wireless access technology known as NR, operates in two frequency ranges as defined in 3rd Generation Partnership Project (3GPP) Release 15, which are FR1 (450 MHz – 6 GHz known as sub-6 GHz) and FR2 (24.25 GHz – 52.6 GHz, known as the millimetre wave) [63].

However, the realisation of a mmw in mobile networks imposes some challenges. This includes opening up Gigahertz (GHz) of spectrum encompassing mmw frequencies, i.e., 30-300 GHz in the future for the next generation (5G) communication networks. Besides, the potential of mmw cellular systems comes with a set of new technical challenges. First, mmw transmission systems require massive arrays of highly directional antennas to mitigate the effects of path losses in high-frequency propagation environments. Since the density of user traffic is uncertain and unevenly distributed over time, C-RAN is flexible to adapt to changing user traffic environment by dynamically adjusting the Remote Radio Heads (RRHs) to Base Band Units (BBUs) configuration. However, the implementation

of load balancing among BBUs is not adequately investigated by both academia and industry.

Both C-RAN and mmw technologies are recognised as the critical enablers for the 5G networks. Although mmw is sensitive to severe transmission attenuation which makes it suitable for short-range wireless communication. C-RAN can fully utilise mmw by a dense deployment of small cell RRHs. Moreover, C-RAN can optimise its network resources according to the varying traffic environment by dynamically adjusting the logical connections between RRHs and BBUs [145]. However, to ensure secure network management, a centrally managed C-RAN architecture with cost-effective RRHs is highly desirable. A centralised C-RAN network can enable smooth handovers with excellent resources utilisation [82]. Moreover, the Self-Organising Network (SON) management mechanisms proposed by Next Generation of Mobile Networks (NGMN) and 3GPP can be employed for C-RAN to make the planning, configuration, maintenance, and optimisation easy and faster [146]. SON is a crucial driver to maximise network performance. The principal idea is to integrate them with intelligence and autonomous adaptability to enhancing network performance, regarding network capacity, coverage and Quality of Service (QoS). It aims at reducing the cost of installation and management by simplifying operational tasks through the capability to configure, optimise and heal itself [147].

In this chapter, a Radio Access Network (RAN) system design is proposed which couples three potential candidates, i.e., mmw, C-RAN, and SON, for 5G communication into multi-user mmw C-RAN systems. Although there are three types of SON architectures a) Centralised, b) Decentralised and c) Hybrid [148]. However, the centralised SON architecture is more suitable for C-RAN and highly manageable regarding the implementation of SON algorithms compared to distributed and hybrid SON architectures. It enables the SON algorithms to jointly optimise multiple network parameters, therefore, allowing a globally tuned system. Moreover, the features of SON include self-configuration, self-optimisation and self-healing, but this study emphasise on a self-optimised C-RAN system focusing on enhancing network QoS.

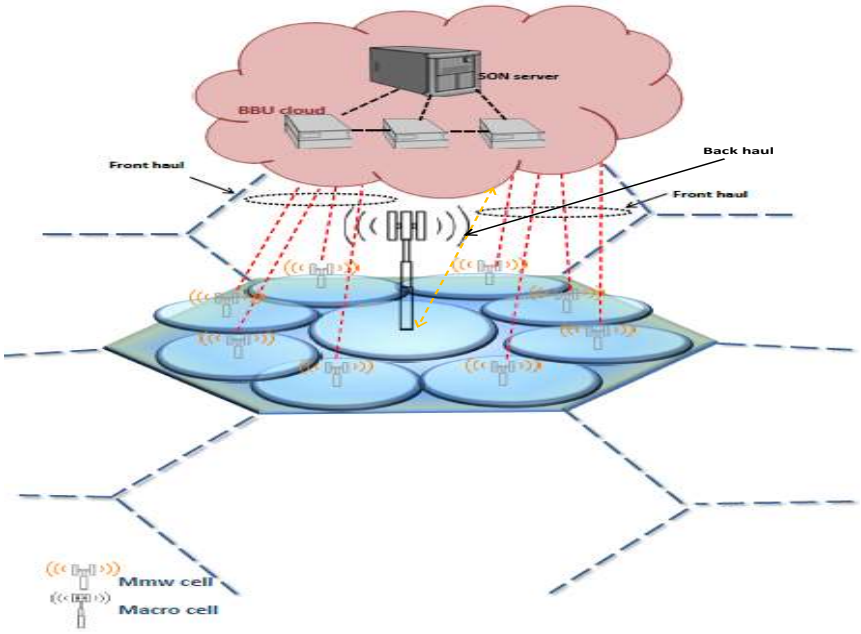


Fig. 4.1 Structure of the proposed cellular network.

Note that, the proposed work is an extension of the previous chapter. The aim is to integrate mmw technology into C-RAN to minimise the number of BBUs and RRHs required to serve a given geographical area. Extending the available spectrum is a solution to realise 5G use cases in an energy-efficient way rather than deploying more RRHs in the system.

4.2 System Model

4.2.1 Proposed Architecture

The proposed architecture is shown in Fig. 4.1. All the small cells' baseband units are decoupled from the low power RRHs and aggregated together in a centralised BBU pool, where the BBUs are coordinate and cooperate in the cloud for efficient resource utilisation. This also is aimed at overcoming the issues mentioned above of blockage and frequent handover of the mmw channel, and for robust reliable communication. The RRHs are left on the cell sites with a simple Radio Frequency (RF) and are connected to the BBU pool via front haul. There are several potential realisations of the front haul

link between the BBU pool and RRHs depending on the cost and the required QoS, which can be wired, wireless or a combination of the two. Furthermore, the fronthaul should offer high capacity and, low latency to support any substantial escalation of data traffic. A SON controller is introduced inside the BBU cloud, which monitors the BBU-pool resource utilisation as well as controlling the switch. The SON controller dynamically assigns BBU radio resources to the RRHs based on traffic demands. At extremely low traffic load conditions, only a high power Macro Base Station (MBS) serves the given geographical area. As the traffic load increases and the MBS reaches its load limit, the geographic area is differentiated into C equally sized small cells serving the same coverage area; C is considered to be seven as a reasonable example in this work. Furthermore, the CDI concept is realised by considering two tiers of cell deployments as shown in Fig. 4.2, i.e., tier-1 deployment imitates a single high-power base station serving a macrocell as in a conventional cellular systems. Tier-2 of the RRHs deployment represents a structure with small cells that uses NR FR2 as a radio access technology. A set $\mathbf{S} = \{\text{RRH}_1, \text{RRH}_2, ..., \text{RRH}_c\}$ is maintained for the tier-2 RRH deployment, which contains a group of RRHs responsible for differentiating the macrocell M into C small cells. The RRH serving each cell c_i is represented as RRH_i, where i represents the cell number in tier-2 RRH structure. The SON server is responsible for cell differentiation and integration with proper BBU-RRH configurations, whereas the switch is in charge of realising the settings via server commands. Note that, with small cell deployment in C-RAN, inter-cell interference is inevitable even in mmw frequency band. Hence, a clustering based interference mitigation technique is adopted to avoid network performance degradation.

4.2.2 Channel Model

In this study, a downlink wireless network is considered, which contains two classes of cells, as shown in Fig. 4.2. The first class is a tier-1 of single MBS in the middle of the cell. The second class is tier-2 of **N** mmw picocells which cover the cell area after differentiation with **M** BBUs inside the cloud. There are **K** users uniformly distributed

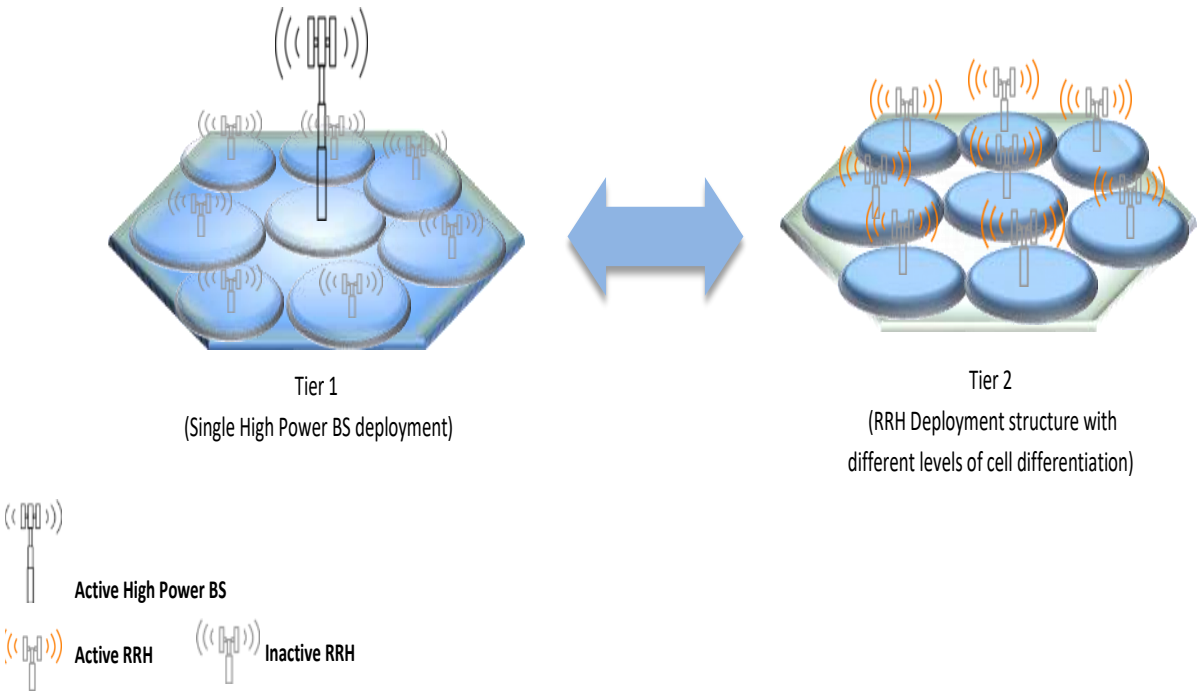


Fig. 4.2 Cell differentiation and integration.

across this cell area. The basic unit of time-frequency resources that can be allocated to users is known as the Physical Resource Block (PRB). Each PRB consists of 12 consecutive sub-carriers, with the PRB bandwidth being 180 kHz for the μw frequency band and 720 kHz for the mmw frequency band FR2[3].

The Signal-to-Interference-and-Noise-Ratio (SINR) received by a user k served by the MBS is represented as:

$$\gamma_{k_{Mi}} = \frac{\mathrm{H}_{k_{m_i}} \mathrm{P}_{m_i}}{\sum_{j \in M, j \neq i} \mathrm{H}_{k_{m_j}} \mathrm{P}_{m_j} + \mathrm{N}_0} \tag{4.1}$$

where, $\mathrm{P}_{k_{mi}}$ and $\mathrm{H}_{k_{mi}}$ are the transmit power and channel gain between the serving MBS m_i and user k. N_0 is the power of Additive White Gaussian Noise and the channel gain is a composite fading channel, which involves path-loss and both small and large scale fading, given as:

$$\mathrm{H}_k = \mathrm{h}^*_{k_{m_i}} l_{k_{mi}} \left[\mathrm{AD}^{-\delta_M}_{k_{m_i}} \right] \tag{4.2}$$

where, $\mathrm{h}^*_{k_{m_i}}$ and $l_{k_{m_i}}$ represent the small and large scale fading channel, respectively, between the user k and MBS m_i. The small scale fading is assumed to be Rayleigh random

variables with a distribution envelope of zero-mean and a unity-variance Gaussian process. The large scale fading is assumed to be a lognormal random variable with a standard deviation of 10dB [131], $\mathrm{AD}_{k_{m_i}}^{-\delta_M}$ reflects the path-loss between an MBS m_i and user k, where A is a constant that depends on the carrier frequency f_c and can be calculated from:

$$\mathrm{A}(f_c) = (\frac{C}{4\pi d_0 f_c})^{\delta} \tag{4.3}$$

where, d_0 is a reference distance of 1 metre and C is the speed of light. $\mathrm{D}_{k_{m_i}}$ is the distance between user k and MBS m in cell i, whilst a path-loss exponent of δ_M = 3 is considered for an urban area cellular radio environment [149]. $\sum_{j\in M, j\neq i} \mathrm{H}_{k_{m_j}} \mathrm{P}_{m_j}$ represents the inter-cell interference power received from all other MBS cells , except for the serving one.

In this paper, it is assumed that after the macrocell is differentiated into C small cells, the RRHs use NR mmw access (FR2) to serve the users. Therefore, for a user k served by an RRH r, a blockage probability must be considered. The blockage model of the mmw channel is adopted from reference [150]. In this model the blockage is expressed as the probability of the link (D) between the user k and the serving RRH r is blocked. The longer the link, the higher the probability that it will be blocked by one or more obstacles. The of line of sight (LOS) probability function $P(D)$ in the network is derived from stochastic blockage models, where the blockage parameters are characterised by random distributions. In this paper, the blockages are modelled as a rectangle Boolean scheme as in [150]. It is shown that the LOS probability function $P(D)$ exponentially decays with the link length. Consequently, the probability that RRH is LOS with user k is represented by $p(D) = e^{-\beta D}$. β is parameter that is calculated from the density and average size of the blockages. $1/\beta$ is called the LOS range and $1/\beta = 141.4$ metre[151]. The Non-Line Of Sight (NLOS) or (blockage) probability is expressed as $1 - P(D)$. Since the mmw channel is affected by blockage, there are different values of path-loss parameters for LOS and NLOS users in the mmw channel model [150]. The SINR received by user k

served by mmw RRH r_i is represented by the following equation:

$$\gamma_{k_{R_i}} = \frac{\mathrm{G}_{r_i}\beta_{k_{r_i}}\mathrm{P}_{r_i}}{\sum_{j\in R, j\neq i}\mathrm{G}_{r_j}\beta_{k_{r_j}}\mathrm{P}_{r_j}+\mathrm{N}_i} \tag{4.4}$$

where, G_{r_i} is the directivity gain of the transmission uniform linear array antenna and its array pattern is approximated using a sectored antenna model. The transmit antenna pattern is assumed to be 10 dB, -10 dB and 30^o, which represent the main lobe directivity gain, back lobe gain, and main lobe beamwidth, respectively [151]. $\beta_{k_{r_i}}$ is the channel gain between RRH r_i and user k , which is a composite fading channel that involves small scale fading, large scale fading and path loss, given by:

$$\beta_{k_{r_i}} = \mathrm{g}^*_{k_{r_i}} l_{k_{ri}} \left[\mathrm{AD}^{-\delta_{mm}}_{k_{r_i}}\right] \tag{4.5}$$

where, $\mathrm{g}^*_{k_{r_i}}$ and $l_{k_{r_i}}$ represent the small and large scale fading channel, respectively, between the user k and its serving RRH r_i. The small scale fading is assumed to be a Nakagami with normalised gamma distribution [131] with parameters $\mathrm{N}_L = 3$ and $\mathrm{N}_N = 2$. The large scale fading is assumed to be a lognormal random variable with a standard deviation of 3.1 dB for LOS and 8.2 dB for NLOS [104]. $\mathrm{AD}^{-\delta_{mm}}_{k_{r_i}}$ reflects the path-loss between the serving RRH r_i and user k, where A is a constant, which depends on the carrier frequency f_c and can be calculated from equation (3). The LOS and NLOS path loss exponents are $\delta_{mm_L} = 2$ and $\delta_{mm_N} = 4$ respectively. $\sum_{j\in R, j\neq i}\mathrm{G}_{r_j}\beta_{k_{r_j}}\mathrm{P}_{r_j}$ represents the inter-cell interference power received from all the RRHs, except for the serving RRH and N_i is the thermal noise power.

4.3 RRH-BBU Association and Configuration

In the proposed self-organised mmw based C-RAN, it is essential to reconfigure the BBU-RRH configuration after a certain time interval, due to changing traffic environments. Since the traffic is increased or decreased at each RRH coverage area, it is inevitable to dynamically re-assign the underlying RRHs to proper BBUs in the cloud to satisfy network limitations. A network's performance determines its QoS, which is measured

by values expressed by a single or multiple objectives. The SON server in the proposed architecture is responsible for identifying optimum BBU-RRH configuration by utilising collected data from the RRHs. Network vendors specify different objectives to evaluate their network's performance. The QoS metrics are then defined and mapped to a set of objectives, which is represented by a weighted normalised function. Let the BBU-RRH configuration at time t, then identifying the BBU-RRH allocation at time $t+1$ that adaptively balances the network load and expected network gain is the prime goal. To identify BBU-RRH association, an allocation indicator vector $\mathbf{r} = \{r_1, r_2, r_3, ..., r_N\}$ is defined, $r_n = m$ when RRH_n is assigned to BBU_m. A user location indicator $\mathbf{u} = \{u_1, u_2, u_3, ..., u_K\}$ is defined, where $U_k = n$ when the received uplink power of user k is at RRH_n is higher than all other RRHs. This shows that user_k is inside RRH_n coverage area. Two Matrix variables **S** and **I** are defined, which are functions of the two vector variables **r** and **u** such that:

$$\underset{\substack{i\in\{1,2,...,N\}\\ k\in\{1,2,...,K\}}}{\mathbf{S}^i_k(\mathrm{r},\mathrm{u})} = \begin{cases} 1 & \text{if } \mathrm{r}_{uk} = \mathrm{r}_i \\ 0 & \text{otherwise} \end{cases} \tag{4.6}$$

$$\underset{\substack{m\in\{1,2,...,M\}\\ k\in\{1,2,...,K\}}}{\mathbf{I}_{m,k}(\mathrm{r},\mathrm{u})} = \begin{cases} 1 & \text{if } \mathrm{r}_{uk} = \mathrm{m} \\ 0 & \text{otherwise} \end{cases} \tag{4.7}$$

where $\mathrm{r}_{uk} \in \{1,2,...,\mathrm{M}\}$, $\mathrm{u}_k \in \{1,2,...,\mathrm{N}\}$, and $\mathrm{k} \in \{1,2,...,\mathrm{K}\}$. The main problem is to achieve a new BBU-RRH configuration (**r**) for a known user location indicator (**u**). Therefore, the BBU-RRH association vector (**r**) is defined as the primary vector variable for the problem. For example, if 3 BBUs are utilised to service 5 RRHs, $\mathbf{r} = \{2,3,3,1,1\}$ indicates that RRH 1,2,3,4, and 5 are assigned to BBU 2,3,3,1, and 1, respectively. Also, the number of possible BBU-RRH configurations is 3^5. Each BBU-RRH configuration shall affect the load-fairness Index, the handover Index, and the proximity index. An Exhaustive search algorithm (ES) is the only effective search algorithm to identify an optimal solution (optimum BBU-RRH allocation). The number of possible BBU-RRH configuration increases exponentially with the number of BBUs (M) and RRHs (N).

Therefore, it is not feasible to utilise ES to determine the optimal solution since the execution time of the algorithm also increases exponentially. The possible solutions for ES is $|\mathrm{M}|^{|\mathrm{N}|}$. Therefore, the use of an evolutionary algorithms is proposed in Section VI to solve the BBU-RRH optimisation problem.

The SINR recieved by a user k in tier-1 cell deployment (macrocell) is defined in eq. (4.1). The average spectral efficiency (ASE) of a user k can be obtained in its simple form by:

$$\begin{aligned}\vartheta_{kMi}(\mathbf{S}) &= \log_2(1+\gamma_{k_{Mi}}) \\ &= \log_2(1+\frac{\mathrm{H}_{k_{m_i}}\mathrm{P}_{m_i}}{\sum_{j\in M, j\neq i}\mathrm{H}_{k_{m_j}}\mathrm{P}_{m_j}+\mathrm{N}_0})\end{aligned} \tag{4.8}$$

Similarly the ASE of a user k in tier-2 cell deployment structure can be achieved by:

$$\gamma_{k_{R_i}}(\mathbf{S}) = \log_2(1+\frac{\mathrm{G}_{r_i}\beta_{k_{r_i}}\mathrm{P}_{r_i}}{\sum_{j\in R, j\neq i}\mathrm{G}_{r_j}\beta_{k_{r_j}}\mathrm{P}_{r_j}+\mathrm{N}_i}) \tag{4.9}$$

Since the achievable ASE is, in fact, the attainable throughput when divided by the user assigned bandwidth. Therefore, the number of resource blocks required by user k to achieve the throughput requirement ϕ is given as:

$$\mathbf{N}^k_{\mathrm{RB}} = \left\lceil \frac{\phi_k}{\mathrm{p}_{\mathrm{BW}}.\vartheta_k(\mathbf{S})} \right\rceil \tag{4.10}$$

where p_{BW} represents the bandwidth of the resource blocks. And the notation $\lceil . \rceil$ is the ceil function.

Now, the BBU-RRH allocation problem is formulated as an integer programming for load balancing amongst the BBUs and to minimise the average load of the network as well as reducing the number of handovers. Moreover, the number of BBUs required in tier-2 deployment structure is also formulated.

4.3.1 The Network Load and Number of Required BBUs

The average load contribution of each BBU from eq. 3.6 which is found by summing the PRB demand of all users associated to BBU m divided by the P_{RB}. The total network load can be found by aggregating the average load of each active BBU as given in eq. 3.7 and the average network load is found from eq. 3.11. Now, the number of required BBUs to satisfy the traffic demand can be found from eq. 3.8.

4.3.2 Key Performance Indicators for Quality of Service Function

The network performance is evaluated by the Key Performance Indicator (KPIs). The SON server determine the optimum BBU-RRH configuration depending on these KPIs by using the number of active BBUs to achieve high QoS with respect to load demand. Following are the KPIs considered by the SON server for the BBU-RRH mapping problem.

Load Fairness Index The level of load balance in the network is evaluated using Jain's fairness index ψ defined in eq. 3.10. Hence, maximising load fairness index ψ is considered as an objective for balanced network load.

Handoffs Index The transition of a network to a different BBU-RRH configuration could result in forced handovers. Associating an RRH to a different BBU leads to forced handovers of all the User Equipment (UE) associated with that RRH. Frequent forced handovers causes degradation in network performance. Therefore, minimising the handover index h($\mathbf{I}$) is considered as another objective function to enhance network performance and is given by eq. 3.12.

Proximity Index To avoid frequent handovers among small mmw cells, mmw RRHs clustering is considered as a third KPI. Hence, the mmw RRHs served by the same BBU forms a compact cluster to minimise unnecessary handovers and reduce inter-cell interference. In addition, this compact RRHs cluster enables other techniques for more reliable mmw link without blockage [106]. The proximity of the RRHs is represented by a binary variable A_{ij}, where A_{ij}= 1, if RRH_i and RRH_j are adjacent,

else A_{ij}= 0. In case that a cluster has multiple RRHs, then those RRHs in that cluster must be adjacent and connected. To formulate the proximity of the RRHs in a cluster, let S1 be any subset of the set of RRHs served by BBU_m (Z_m), such that $S1 \subset Z_m$, $S1 \neq \emptyset$, and $S1 \neq Z_m$. Let S2 be another subset of Z_m, such that, $S2 = Z_m - S1$, i.e., S2 is the complementary set of S1. The compactness of RRHs in Z_m can be proved by satisfying the following property.

$$\sum_{i \in S1} \sum_{j \in S2} A_{ij} \geq 1 \tag{4.11}$$

Therefore, to ensure the compactness of an RRH group served by a BBU, a KPI for proximity Index factor (PI) is defined, which is the ratio of the total number of boundary edges in the network to the number of common boundary edges between different RRH clusters i.e.,

$$PI(\mathbf{I},\mathbf{S}) = \frac{\sum_i^N \sum_{j \neq i}^N A_{ij}(\mathbf{I},\mathbf{S})}{\sum_i^N \sum_{j \neq i}^N A_{ij}(\mathbf{I},\mathbf{S}) Z_{ij}(\mathbf{S})} \tag{4.12}$$

The aim is to achieve minimum PI in the network for RRH clusters. Therefore, the main QoS function is defined as a weighted combination of three KPIs, discussed earlier, and is given as

$$\begin{aligned} &\text{Max } QoS(\mathbf{I},\mathbf{S}) = w_1 \psi(\mathbf{I},\mathbf{S}) - w_2 h(\mathbf{I}) - w_3 PI(\mathbf{I},\mathbf{S}) \\ &\text{s.t.} \sum_{k=1}^{K} I_{m,k} N_{RB,\kappa}(\mathbf{S}) \leq P_{RB}, \forall m \in \{1,2,...,M\} \\ &\sum_{m=1}^{M} I_{m,k} = 1, \forall k \in \{1,2,...,K\} \end{aligned} \tag{4.13}$$

where w_1, w_2 and w_3 are the weights assigned to each function or simply the priority levels selected for each KPI. The weights are selected based on Rank Order Centroid (ROC) method [152], where the functions are sorted according to their importance or priority. The weights or priority levels are then applied to each KPI for decision making. The corresponding weight of a KPI defines its preference

value and is set according to network operator's preferences. The priority of each function is taken as an input and converted into weight using the following formula:

$$\mathrm{w_i} = \left(\frac{1}{\mathrm{F}}\right) \sum_{\mathrm{n=i}}^{\mathrm{F}} \frac{1}{\mathrm{n}} \tag{4.14}$$

where F is the number of functions (KPIs) and w_i is the weight of the i^{th} function. The QoS maximisation is subjected to two constraints the first one is that the total PRBs demand of the users served by the same BBU must not exceed the total number of PRBs of each BBU (P_{RB}). The other constraint is each user must be served by one BBU only.

4.4 Self-Optimised mmw Based C-RAN Algorithm

In this section, a self-optimised mmw C-RAN algorithm is proposed based on the intuitive analysis above. The algorithm uses network gain information to identify proper network resource usage and optimum BBU-RRH configuration. The block diagram of the algorithm is shown in Fig. 4.3. Network information is collected in the first step and analysed to initiate a cell split or merge. The information includes user location indicator **u**; load contributed by each RRH, BBU-RRH mapping vector **r**, network load, and the required number of BBUs. The algorithm seeks to utilise the network resources efficiently by calculating the necessary amount of BBUs and the RRHs to serve capacity demands at the end of each cycle. Apart from a single BBU required to serve load requirements, proper BBU-RRH configuration is adjusted at the end of the optimisation step by comparing the analysed and optimised QoS values. For the optimisation part of the algorithm, a Discrete Particle Swarm Optimisation (DPSO) algorithm is developed as an Evolutionary Algorithms (EA) to solve the BBU-RRH configuration problem and is explained in the next section. The optimisation process continues until the cycle is completed. Note that, the algorithm shown in Fig. 4.3 is triggered at the beginning of each cycle.

The pseudo-codes for cell splitting and integration are given in Algorithm 6 and Algorithm 7, respectively.

Algorithm 6: Pseudo-code for Cell Split and initial BBU-RRH Mapping

```
Input : Current network load η (I,S) from (3.7)
        BBU-RRH mapping vector r
        Required number of BBUs from (3.6)
if No. of active BBUs =1 then
    if η_m(I,S) ≥ |P_RB| then
        -Activate required No.of BBUs
        -Split Macro-Cell to tier-2 RRH structure
        for m=1 to No. of active BBUs do
            -Compute η_m(I,S) from (3.8)
            if η_m(I,S) ≤ Upper limit then
                A← BBU_m {Add BBU_m to active BBU list A}
            end
            Sort A in increasing order of BBU loads.
            i=1;
            while Not the end of Set S do
                m=1;
                if m > |A| then
                    m=1
                end
                BBU_m ← RRH_i{Map RRH_i to BBU_m
                }i=i+1;
            end
        end
        -Update BBU-RRH mapping Vector r
    else
        -No Cell split required
        -Tier-1 structure remains
    end
else
    if No. of active BBUs ≤ No. of required BBUs then
        -Activate the required BBUs
        -Update BBU-RRH vector r
    end
end
```

Both algorithms execute the scaling of RRHs and BBUs with respect to network load distribution. However, an important consideration is the first association of RRHs to the required number of BBUs during cell differentiation and integration. The algorithms cover all possible cases of initial BBU-RRH assignment during cell split or merge along

Algorithm 7: Pseudo-code for Cell-merge and initial BBU-RRH mapping

Input : List A of No. active BBUs
No. of required BBUs
BBU-RRH mapping vector r

if *No. of required BBUs =1* **then**
  -Integrate all cells into tier-1 RRH structure, i.e., a high power BS should serve the geographical area.
  -Switch-off remaining BBUs.
  -Update BBU-RRH mapping vector **r**.
else
  if *No. of required BBUs* $< |A|$ **then**
    for *m=1 to* $|A|$ **do**
      for *i=1 to C* **do**
        -Select RRH_{ij} from BBU-RRH vector **r**
        if $RRH_i = m$ **then**
          $Z_m \leftarrow \text{RRH}_i$
          {Z_m is a List of RRHs handled by BBU_m}
        end
      end
    end
  end
  -Sort List A in decreasing order of BBU loads
  for *m=1 to end of List A* **do**
    if $m \neq |$*No. of required BBUs*$|$ **then**
      $\overline{A} \leftarrow \text{BBU}_m\{\overline{A}$ is a List of required BBUs$\}$
    else
      $\overline{B} \leftarrow \text{BBU}_m$
      $\{\overline{B}$ is a List of BBUs to be switched off$\}$
    end
  end
  -Sort List $\overline{A}$ in increasing order of BBU loads
  for *i=1 to end of List* $\overline{B}$ **do**
    -Select i^{th} BBU from List $\overline{B}$
    -Select List Z_m of the i^{th} BBU
    m=1;
    for *j=1 to end of* Z_m **do**
      if $m > |\overline{A}|$ **then**
        m=1;
      end
      -Select RRH at j^{th} index in List Z_m
      -Select BBU at m^{th} index of List $\overline{A}$
      $-\text{BBU}^m \leftarrow \text{RRH}^j$
      m++
    end
  end
  -Switch off all BBUs in List $\overline{B}$
end

with cases where the number of BBUs are increased, decreased or remain unchanged. The initial BBU-RRH mapping is important for efficiently utilising the available BBU resources so as to prevent high blocking rate before a proper BBU-RRH mapping is identified in the optimisation step.

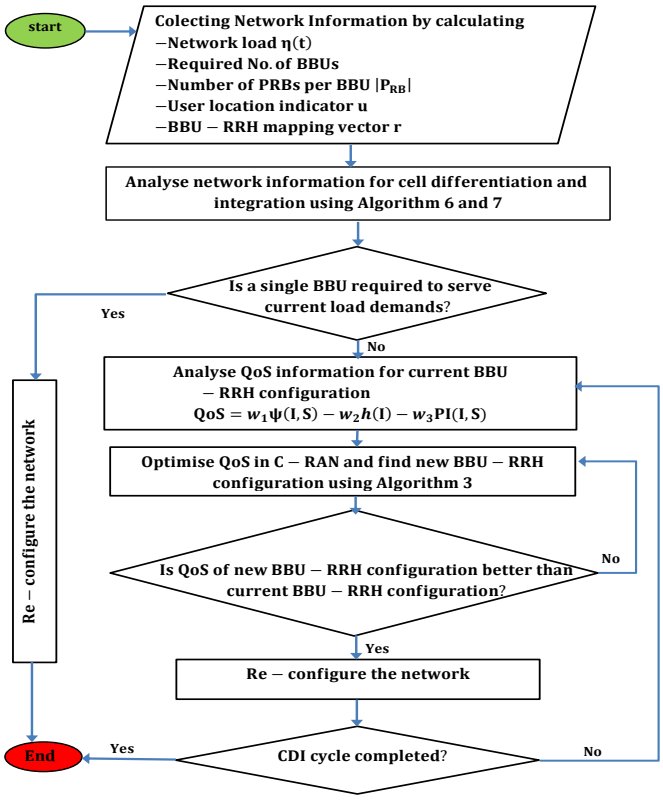


Fig. 4.3 Block diagram of CDI Algorithm for one CDI cycle.

4.5 Optimisation Algorithms

4.5.1 Discrete Particle Swarm Optimisation (DPSO)

Particle Swarm Optimisation (PSO) utilises a population (or swarm) of particles, where each individual particle represents a solution [140], namely BBU-RRH association vector **r** defining the BBU-RRH configuration. As the QoS represented in eq. (4.13) is considered as the main objective function, PSO seeks to maximise the QoS function by finding the best solution vector $\{r_1, r_2, ..., r_N\}$. PSO utilises a group of particles (or solutions) to probe the solution space in a random way with different velocities. To

direct the particles to their best fitness values, the velocity of each particle is changed stochastically at each iteration. The velocity update of each particle j depends on the historical best position experience (pbest) of the particle itself and the best location experience of neighbouring particles i.e., the global best position (gbest) [153] and is given as

$$\begin{aligned} &\mathrm{v}_j^{\mathrm{I}} = w\mathrm{v}_j^{\mathrm{I}-1} + c_1\varepsilon_1\left(\mathrm{pbest}_j^{\mathrm{I}} - x_j^{\mathrm{I}}\right) + c_2\varepsilon_2(\mathrm{gbest}_j^{\mathrm{I}} - x_j^{\mathrm{I}}) \\ &1 \leq j \leq |\Delta| \end{aligned} \tag{4.15}$$

where $|\Delta|$ represents the population (or swarm) of particles and I represents the Iteration number. x_j^{I} is the current position of particle j in iteration I and $\varepsilon_1, \varepsilon_2$ are random numbers between 0 and 1. Both c_1 and c_2 are acceleration constants that pulls the particle towards best position. Values in the range 0-5 are chosen for c_1 and c_2. The inertial weight w represents the effect of preceding velocity on the updated velocity. Choosing an optimum value for w can assist a balanced proportion between global and local exploration of the search space. Usually values between 0-1 are selected for w [141]. A value of 0.9 for w is selected in this work. The new position of particle j for the next iteration $\mathrm{I}+1$ will be:

$$x_j^{\mathrm{I}+1} = x_j^{\mathrm{I}} + \mathrm{v}_j^{\mathrm{I}} \tag{4.16}$$

PSO terminates if a stopping criterion is satisfied, e.g. after reaching a predefined number of iterations I_{max}. Since the solution vector, $\mathbf{r}$ (or particle) should be real-valued, the standard PSO algorithm can not be applied to solving this discrete optimisation problem. In this study, a DPSO is developed to solve the QoS maximisation problem defined in eq. (4.13). The pseudo code of DPSO is given in chapter three.

4.5.2 Genetic Algorithm

Genetic Algorithm (GA) is a biological population inspired algorithm, it is used for searching optimising complex problems solution [154]. It finds the solution to a specific problem in a competing population as individuals. GA first chooses a Population, which is random set of solutions. Each solution in the population is called a chromosome. For

measuring the solutions' fitness in the Population, each chromosome is evaluated for competing with the others. In this process, fitness values are assigned to each chromosome which gives the probability of each chromosome to be the possible solution to the problem. The chromosomes that have the highest fitness probabilities would be selected for mating. New solutions/chromosomes are generated by the process of crossover and mutation. the algorithm stops When it converges to the optimum solution, which means it produce new solutions similar to the previous generation or when a certain predefined criterion is met. The GA is illustrated in the following steps and block diagram[155].

Step 1: Produce an initial population R^0 with $|\Delta|$ chromosomes, N-bit (gene) chromosomes (RRH-BBU association). N is considered according to the number of RRHs in the network.

Step2: Calculate the fitness value of each chromosome (RRH-BBU association) in current population using the fitness function F (i.e., QoS in eq. 4.13).

Step 3: If the convergence criterion is qualified by the best candidate RRH-BBU association solution (chromosome) or the maximum number of generations have been reached out, then terminate, else proceed to next step.

Step 4: Generate a set of $|\beta|$ best chromosomes (RRH-BBU association) from the currently sorted population R^I. The selection probability P_s is used to select the best RRH-BBU association to form set β (i.e., $\beta = P_s|\Delta|$).

Step 5: produce new chromosomes η ($|\eta| = |R^I| - |\beta|$) by performing crossover and mutation operations on set β. The newly generated RRH-BBU solutions η then replaces the infeasible solutions (R^I-β) of the current population R^I in order to generate a new population.

Step 6: Go to step 2 and repeat all steps.

4.6 Computational Results and Analysis

4.6.1 Selecting Initial Swarm/Population for DPSO

Generally, the swarm size is randomly selected from a set $\{1, 2, ..., N\}$ with a uniform distribution. However, the possibility of achieving a large number of solutions with small

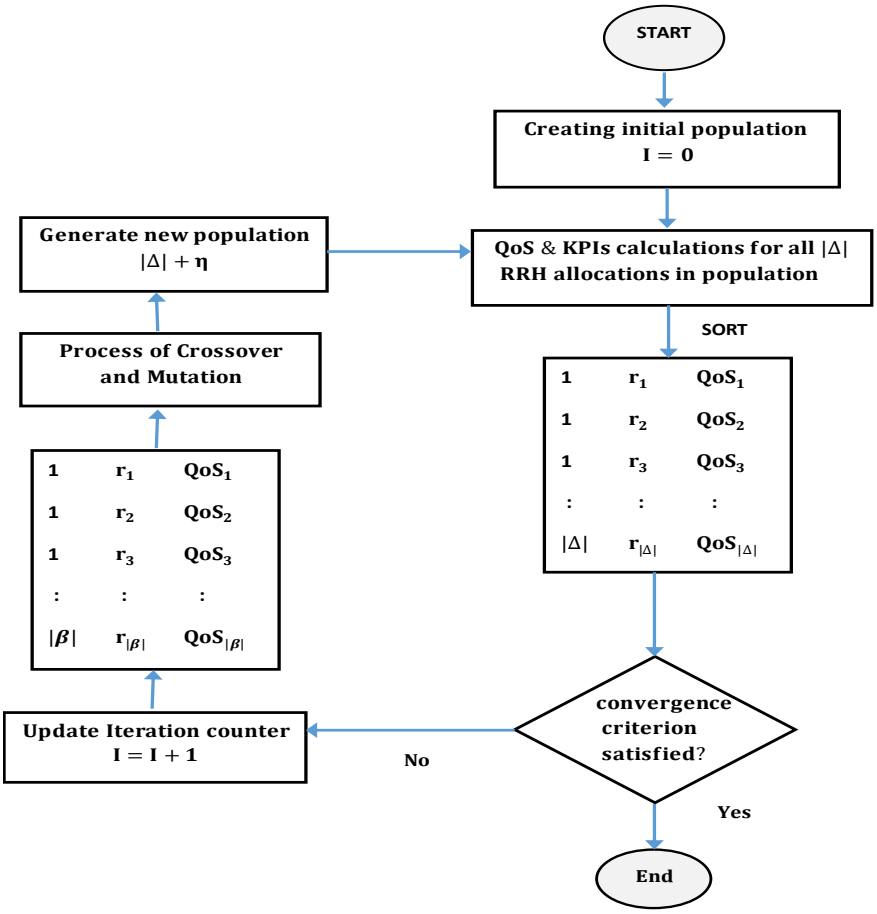


Fig. 4.4 Genetic Algorithm block diagram.

fitness values can be expected if the swarm size is much smaller than the actual size of the search space. To avoid this issue, 30 per cent of initial swarm size is selected to be precise in the form of BBU-RRH clusters desired (As explained earlier).

In this work, a benchmark problem is considered to demonstrate and verify the performance of DPSO algorithm and compared to the known GA and ES. The benchmark problem consists of a network scenario with 5 BBUs and 7 RRHs. A Monte Carlo analysis is performed, where 3 RRHs are randomly chosen to have high user density. RRHs with high user density have 4 times more users than a non-dense RRHs. The users are uniformly distributed under each RRH coverage area. For simulation, this study chooses 6 and 24 users in non-dense and highly dense RRH coverage areas, respectively.

A Monte Carlo analysis is performed by repeating the DPSO, GA and ES algorithms 50 times, and the results obtained by each algorithm is averaged. Note that, a different

initial population/swarm is selected for each Monte Carlo iteration. The swarm size ($|\Delta|$) selected for simulation is 2×10^3, which is less than 1% of the entire solution space. The algorithms are terminated after 200 iterations or generations. The ES algorithm, however, does not depend on number of iterations or generations. In fact, the ES algorithm is utilised to search for all possible BBU-RRH configurations or solutions ($|\mathrm{M}|^{|\mathrm{N}|}$). The best solution identified by ES helps to demonstrate the improvement at each iteration or generation of DPSO and GA, respectively. The optimum values in each graph help in demonstrating the performance of GA and DPSO algorithm over each iteration or generation.

4.6.2 Performance of DPSO Algorithm

This work defines Convergence Rate (CR) as the ratio of a number of times an optimum solution is achieved to the total number of iterations or generations performed.

Figs. 4.5, 4.6, 4.7 and 4.8 present the QoS function, the load fairness index, the handover index, and the proximity index values over 200 iterations for the benchmark problem defined. The optimum values in the figures are achieved by ES algorithm after finding all possible BBU-RRH configurations ($|\mathrm{M}|^{|\mathrm{N}|}$). The ES does not depend on the number of iterations or generations. The optimum values help in demonstrating the improvement at each iteration/generation of GA and DPSO algorithms.

In Fig. 4.5, both GA and DPSO are converging to the optimal value. The convergence rate of GA and DPSO (to the optimal solution) is 0.8 and 0.815, respectively. Where, the optimal solution is achieved 160 and 163 times over 200 iterations/generations by GA and DPSO, respectively. GA finds the optimal solution after $40 \times |\Delta|$ fitness evaluations whereas DPSO finds the optimum solution after $37 \times |\Delta|$ fitness evaluations. However, the ES algorithm finds the optimum value after 5^7 ($|M^{|N|}$) fitness evaluations, which are too large.

Figs. 4.6 and 4.7 demonstrate that the algorithms converge to the optimal load balancing index and handover index values after 12^{th} and 40^{th} generations for GA and after 5^{th} and 37^{th} iterations for DPSO.

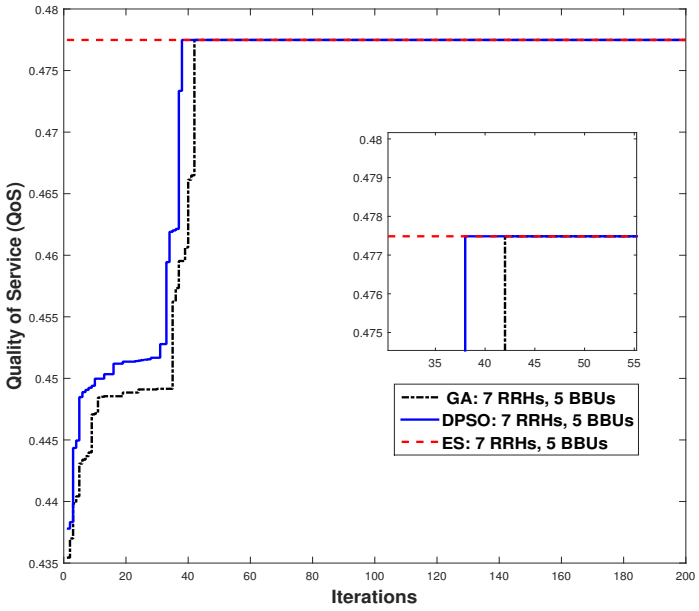


Fig. 4.5 QoS values for DPSO, GA, and ES.

Fig. 4.8 demonstrates that the DPSO algorithm converges faster to the optimal solution compared to GA. The number of times DPSO finds the optimal solution is 164 over 200 iterations. Whereas, GA finds 159 optimal solutions over 200 iterations. Both GA and DPSO converge to the optimal solutions faster than the ES algorithm. However, DPSO outperforms GA in terms of convergence.

4.6.3 Complexity Comparison

The computation complexity of ES algorithm is $\mathrm{O}\left(|\mathrm{M}|^{|\mathrm{N}|}\right)$, where M and N are the number of active BBUs and RRHs in the network, respectively. This shows that number of fitness evaluations required by ES algorithm is $|\mathrm{M}|^{|\mathrm{N}|}$. However the DPSO and GA complexity in the I^{th} iteration/generation is $\mathrm{O}\left(|\Delta|\mathrm{I}\right)$ where the number of fitness evaluations performed by both algorithms is $|\Delta|$. Therefore, the required fitness evaluations for both GA and DPSO are $|\Delta|\mathrm{I}$. The ES, however, requires an enormous amount of fitness evaluations which is even $|\mathrm{M}|^{|\mathrm{N}|} - |\Delta|\mathrm{I}$ more than the number of fitness evaluations performed after the I^{th} iteration/generation of DPSO and GA.

In the benchmark problem defined, the optimal solution for QoS is achieved after 40 and 37 fitness evaluations for GA and DPSO, respectively. Whereas, the optimal value of for QoS by ES is found after 5^7 fitness evaluations. Note that, the optimum QoS value

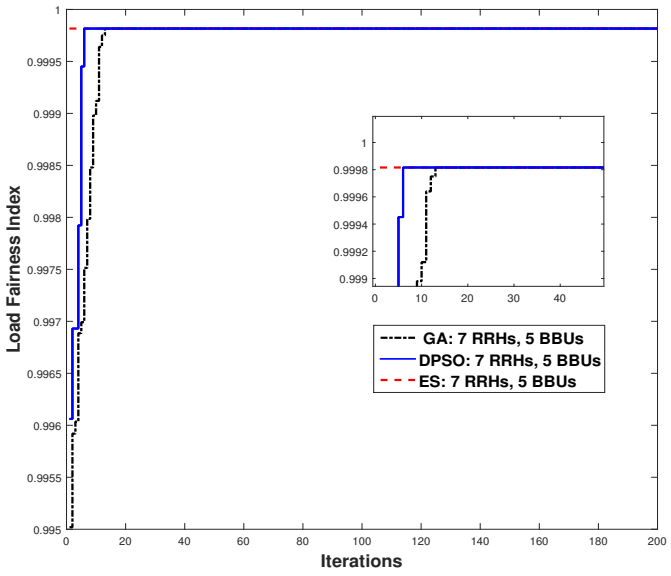


Fig. 4.6 Load fairness index values for DPSO, GA, and ES.

found by ES required 78.125×10^3 fitness evaluations which are more than the number of fitness evaluations at the 37^{th} and 40^{th} iteration/generation of DPSO and GA algorithm, respectively.

4.6.4 Capacity and Blocked Users Calculations

The CDI algorithm performance for a self-organised C-RAN is investigated using NR wireless access technology and compared to fixed C-RAN and CDI enabled C-RAN for more extensive analysis. The CDI algorithm is installed on a SON server in the BBU pool, which is responsible for cell differentiation and integration, according to the capacity demand variation. The end users' different data rates are proposed based on 3GPP standard simulation parameters, which are 4-25 kbps for audio, 32-384 kbps for video, 28.8 kbps for data, and 60 kbps for real-time gaming services[142]. After each CDI cycle, the RRH-BBU association vector $\mathbf{r} = \{r_{11}, r_{12}, ..., r_{in}\}$ is preserved and updated. Newly activated RRHs and BBUs in the network are mapped according to the cell differentiation and integration algorithm proposed in [124]. In all the compared scenarios in this work, the BBU cloud holds five BBUs, which can be activated and deactivated based on traffic demand. However, the fixed C-RAN scenario does not support cell differentiation or integration, and seven RRHs serve the entire macrocells coverage area using the Long

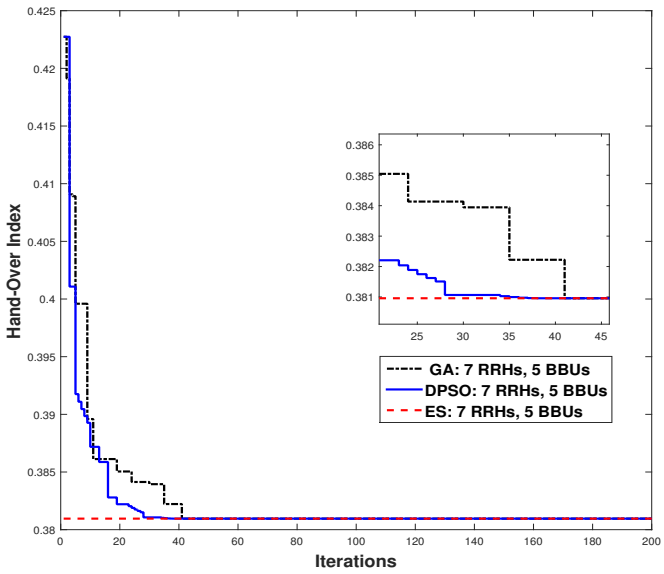


Fig. 4.7 Hand-over index values for DPSO, GA, and ES.

Term Evolution-Advanced (LTE-A) band constantly, regardless the traffic condition. Whereas, The CDI enabled C-RAN scenario supports cell differentiation and integration using LTE-A band only. Furthermore, an increasing user arrival is considered in the network with a random data rates requirement, as elaborated upon earlier. Moreover, a Monte-Carlo analysis is performed, where 100 uniformly distributed users are envisaged for each instance, and the average of all distributions are taken into account regarding network load, throughput, blocked users, and resource shortage analysis. Figs. 4.9 and 4.10 show the relative performances regarding average blocked users and average network throughput with the Proportional Fair (PF) and Round Robin (RR) scheduling techniques. The performance of network adopting mmw-CDI algorithm is investigated by differentiating the overloaded macrocell to multiple small cells, which use NR FR2 for cellular access. The results of simulation validate the advantage of using NR wireless access in CDI-enabled C-RAN compared to LTE-A and a fixed C-RAN setting. The average blocked users in the network are much higher with lower average throughput when fixed C-RAN is considered adopting any scheduling technique, as shown in Fig. 4.9 and Fig. 4.10 compared to mmw-CDI. The average blocked users significantly decreases and the average network throughput tremendously increases when using NR-CDI compared to CDI-enabled C-RAN and fixed C-RAN. The results from Fig. 4.9

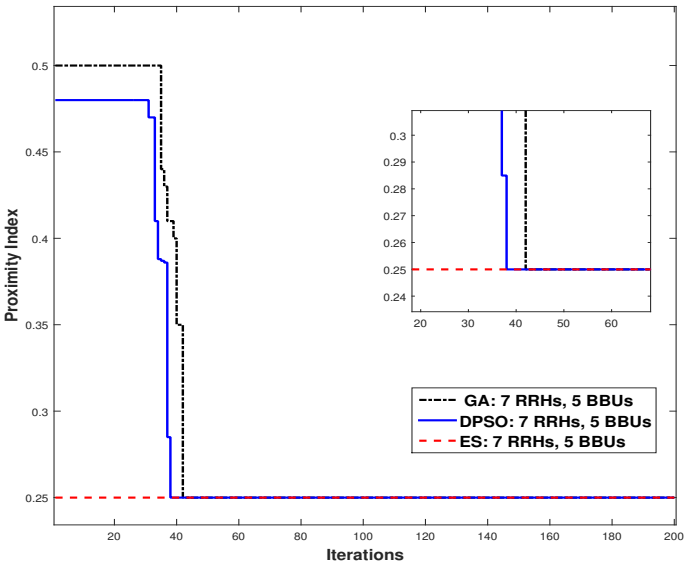


Fig. 4.8 Proximity index values for DPSO, GA, and ES.

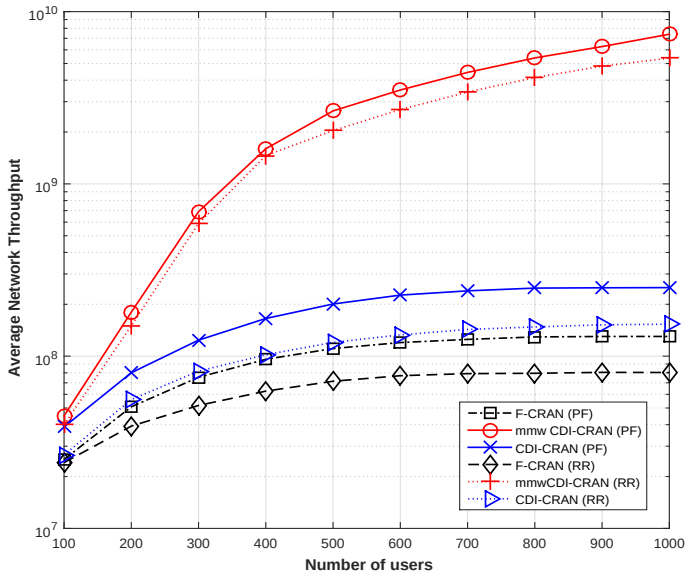


Fig. 4.9 Average network throughput comparison for different scenarios.

shows average network throughput increases by 224,06% and 138.83% in the NR-CDI network compared to fixed C-RAN and CDI-enabled C-RAN, respectively, all being enabled with RR schedulers. Whereas with PF schedulers, an increase of 282.9% and 121.7% is observed respectively.

This is because of the features of wide bandwidth and directional transmission when using NR FR2 for mmw CDI-CRAN, which leads to higher throughput. This also indicates efficient resource utilisation during cell differentiation and integration.

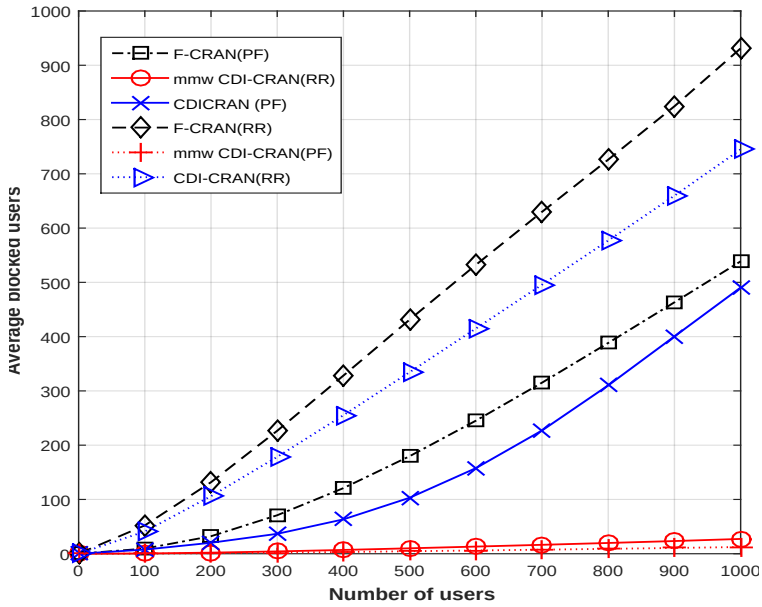


Fig. 4.10 Average blocked users comparison for different scenarios.

Moreover, the average number of blocked users dramatically decreases in the mmw-CDI compared to fixed C-RAN and CDI-enabled C-RAN. That is, about 97.9% and 96.4% reduction in the average number of blocked users with PF scheduler while 88.19% and 84.23% with RR is observed as shown in Fig. 4.10. provided that in both the CDI-CRAN and fixed C-RAN scenarios there is 5 BBUs, 5×100 PRBs available, whilst for the NR-CDI there is 5 BBUs, and 5×264 PRBs [3].

4.6.5 Energy Efficiency

The energy efficiency (EE) is the power amount required by the network to transmit data. It is one of the important performance metrics in emerging 5G networks and the key point in the sense of green communications due to the increasing circuit power consumption especially with densely deployed BSs (higher EE denotes lower energy consumption). Therefore, this section focuses on the energy consumption aspect in the C-RAN, where the energy efficiency for the overall system can be defined as the ratio of the network throughput to the total consumed power, which is given by:

$$EE = \frac{Throughput}{P_{total}} \tag{4.17}$$

The power model is adopted from [124, 156], which depends on the load to estimate network power consumption. The estimated power consumed by the active BBUs is given by:

$$P_t = \sum_{m=1}^{M} (P_{BBU} + \sum_{r \in Z_m} P_{RRH}) \tag{4.18}$$

where P_{BBU} is power consumed by that particular BBU, P_{RRH} is the sum of the power consumed by all the RRHs associated with that particular BBU and Zm represents the list of RRHs handled by BBU.

From the graphs in Fig. 4.11, the average network energy efficiency increases by 272.43% and 60% in the mmw CDI-CRAN compared to fixed C-RAN and CDI-CRAN, respectively, all being enabled with PF schedulers. Whereas with RR schedulers, an increase of 235% and 57.3% is observed respectively. This rise in energy efficiency comes from the noticeable increase in the average network throughput and power gain that coming from adopting CDI algorithm and mmw band.

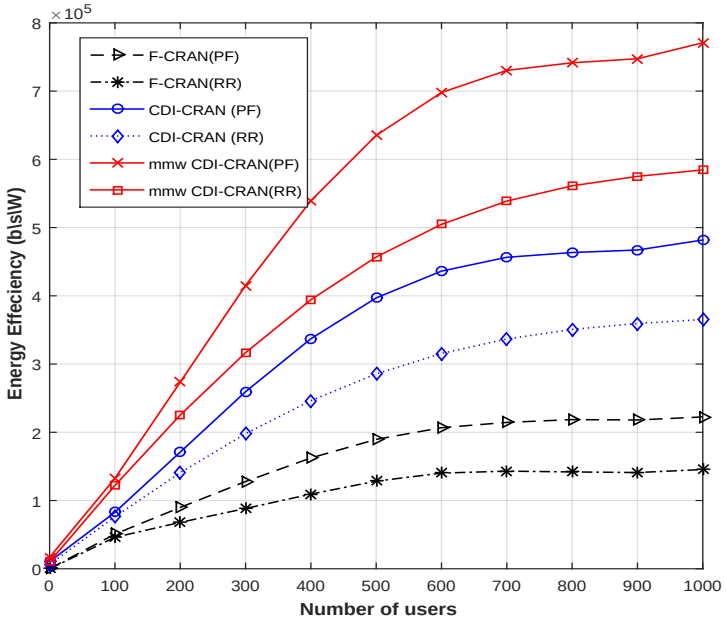


Fig. 4.11 Average network energy efficiency for different scenarios.

4.7 Summary

In this study, a cell differentiation and integration scheme is proposed for the future 5G networks. This scheme is aimed at exploiting the NR NSA feature of coexistence with LTE-A and inter-working between high and low frequencies so as to adapt to the varying capacity demand and thus, utilise network resources more efficiently. A CDI algorithm installed on a SON server in the BBU pool cloud monitors and responds to the capacity demand variation. In this scheme, at high traffic demand the macrocell service area is divided into small cells by activating the deployed RRHs, which use wide band NR FR2 for cellular access to increase the capacity. At low traffic demand conditions, the small cells are reintegrated again into a macrocell served by a high power MBS. The CDI-enabled network performance has been tested for an mmw small cells and μw small cells scenarios and compared to fixed C-RAN configuration. The BBUs-RRHs association is formulated as an optimisation problem with constraint to maximise the QoS. DPSO and GA are used as EA to solve the optimisation problem. Both algorithms converged to the optimum solution that is found by the ES. DPSO and GA reach the optimal solution in less fitness evaluation than ES and are less complex. The computational results show an order of magnitude increase in average network throughput for the mmw CDI-CRAN network compared to a fixed C-RAN and CDI-CRAN using the same scheduling technique. Furthermore, the average number of blocked users is extremely reduced for the mmw CDI-CRAN compared to the other scenarios.

Chapter 5

An Interference Mitigation Scheme for Millimetre Wave Heterogeneous Cloud Radio Access Network with Dynamic RRH Clustering[1]

5.1 Introduction

In its early deployment, the Fifth-Generation (5G) is expected to be integrated with conventional Micro Wave (μw) base stations (BSs) in small cells using the Millimetre Wave (mmw) band for access to enhance the traditional system. The main goal of 5G is universal coverage and at a higher rate. In this chapter, a scheme for increasing the rate coverage of the conventional macro cell by exploiting the wide mmw band spectrum is proposed. This scheme is inspired by the Soft Frequency Reuse (SFR) technique. The cell edge areas have low coverage rates of one tenth of the peak data rate due to high interference and less received power. For this reason, it is served by a mmw band, where small mmw Remote Radio Heads (RRHs) are overlaid onto conventional macro cell edges in a Heterogeneous Cloud Radio Access Network (HC-RAN) architecture. The neighbouring RRHs are clustered together based on the suggested clustering algorithm, which is formulated as a bin packing problem. Simulation results show that the HC-

[1] A part of this chapter has been published in *IEEE International Symposium on Networks, Computers and Communications (ISNCC): Wireless and Mobile Networks*, Istanbul, 2019, pp. 169-174, 978-1-7281-1244-2/19/©2019 IEEE [157]

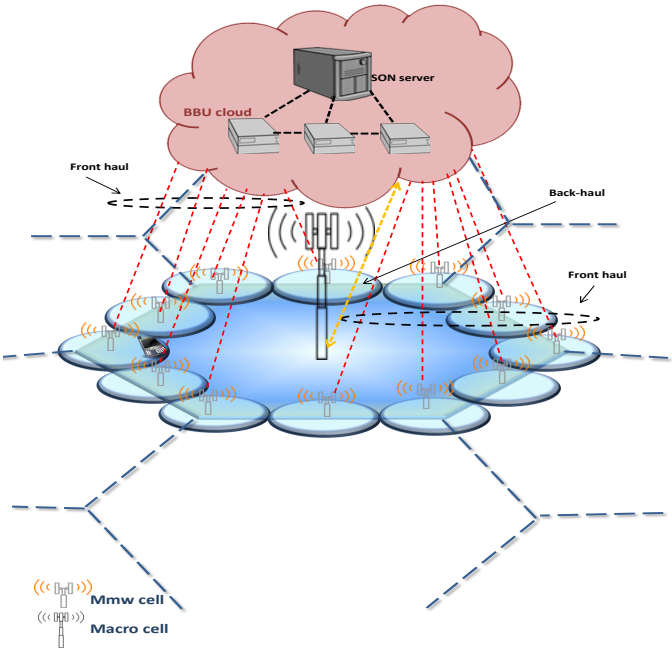


Fig. 5.1 Structure of millimetre wave heterogeneous network

RAN system adopting the suggested scheme has significantly higher coverage and rate compared to an HC-RAN system with small cells only deployed in hot spots. This enhancement in rate and coverage is due to a significant reduction in the interference level resulting from the adopting the suggested scheme, which benefits from wide band properties of an mmw channel. Furthermore, the impact of the mmw cell size and density on the system performance is investigated.

5.2 System Model

5.2.1 Proposed Architecture

The proposed architecture is shown in Fig. 5.1, with the system being adopted from the 3rd Generation Partnership Project (3GPP) Heterogeneous Networks (HetNet) wrap around model [158]. The two tier model consists of a hexagonal grid of high power macrocells and low power RRHs are overlaid at centre points on the boundaries between the adjacent macrocells to cover the cell edge areas. The Base Band Units (BBUs) are decoupled from the low power RRHs and aggregated together in a centralised BBU pool, where the BBUs are coordinated and cooperate in the cloud for efficient resource

utilisation. This also is aimed at overcoming the aforementioned issues of blockage and frequent handover of the mmw channel. The RRHs are left on the cells sites with a simple Radio Frequency (RF) and connected to the BBU pool via front haul. There are several potential realisations of the front haul link between the BBU pool and RRHs depending on the cost and the required QoS, which can be wired, wireless or a combination of the two. Furthermore, the fronthaul should offer high capacity, low latency to support any substantial escalation of data traffic. HC-RANs that use mmw for small cell access favour the use of mmw wireless fronthaul for low latency, high capacity, cost effective and easy deployment when compared to fibre optic, which is also more expensive and less scalable. High power Macro Base Stations (MBSs) are also connected to the BBU pool through the backhaul using the data and control interfaces, S1 and X2, respectively, which are inherited from the 3GPP standards.

The proposed system performs a modified SFR technique for 5G. The traditional SFR is an inter cell interference mitigation technique in which the cell service area is partitioned into spatial subregions (centre region and edge region), with each being assigned to different μw frequency sub bands [159]. The suggested architecture follows the same scheme, yet it assigns the whole Long Term Evolution-Advanced (LTE-A) band to the inner area and allocates the mmw band to the cell edge area. Traditional MBS are used to serve the inner region Users' Equipment's (UEs) and provide control signalling for the whole cell area, while RRHs used to serve edge area. In this work, dynamic small cell clustering is proposed, where the BBU resources are shared and scheduled dynamically between the cluster of RRHs associated to it depending on the users' demand and according to the suggested clustering model.

5.2.2 Channel Model

In this work, a downlink two tiers heterogeneous wireless network is considered, which contains two classes of cells, as shown in Fig. 5.1. The first class is a tier of hexagonal grid of **M**w BSs with an antenna in the middle of the cell. The second class is a mmw small cell, which represents the second overlaid tier of **R** RRHs in each macrocell that

covers the cell edge area. There are **K** users uniformly distributed across the cell area. The Signal-to-Interference-and-Noise-Ratio (SINR) received by a user k located in the cell centre region, which is served by the MBS , is represented as:

$$\gamma_{k_{Mi}} = \frac{\mathrm{H}_{k_{m_i}} \mathrm{P}_{m_i}}{\sum_{j \in M, j \neq i} \mathrm{H}_{k_{m_j}} \mathrm{P}_{m_j} + \mathrm{N}_0} \tag{5.1}$$

where, $\mathrm{P}_{k_{mi}}$ and $\mathrm{H}_{k_{mi}}$ are the transmit power and channel gain between the serving MBS m_i and user k. N_0 is the power of Additive White Gaussian Noise and the channel gain is a composite fading channel, which involves path-loss and both small and large scale fading, given as:

$$\mathrm{H}_k = \mathrm{h}^*_{k_{m_i}} l_{k_{mi}} \left[\mathrm{AD}^{-\delta_M}_{k_{m_i}} \right] \tag{5.2}$$

where, $\mathrm{h}^*_{k_{m_i}}$ and $l_{k_{m_i}}$ represent the small and large scale fading channel, respectively, between the user k and MBS m_i. The small scale fading is assumed to be Rayleigh random variables with a distribution envelope of zero-mean and a unity-variance Gaussian process. The large scale fading is assumed to be a lognormal random variable with a standard deviation of 10dB [131], $\mathrm{AD}^{-\delta_M}_{k_{m_0}}$ reflects the path-loss between a MBS m_i and user k, where A is a constant that depends on the carrier frequency f_c and can be calculated from:

$$\mathrm{A}_j(f_c) = \left(\frac{C}{4\Pi d_0 f_c}\right)^{\delta_j} \tag{5.3}$$

where, d_0 is a reference distance of 1 metre [149]. $\mathrm{D}_{k_{m_0}}$ is the distance between user k and MBS m in cell i, whilst a path-loss exponent of δ_M = 3 is considered for an urban area cellular radio environment [130]. $\sum_{j \in M, j \neq i} \mathrm{H}_{k_{m_j}} \mathrm{P}_{m_j}$

represents the inter-cell interference power received from all other MBS cells , except for the serving cell.

In this chapter, it is assumed that after the macrocell is differentiated into C small cells, the RRHs use New Radio (NR) mmw access (FR2) to serve the users. Therefore, for a user k served by an RRH r, a blockage probability must be considered. The blockage model of the mmw channel is adopted from reference [150]. In this model the blockage

is expressed as the probability of the link (D) between the user k and the serving RRH r is blocked. The longer the link, the higher the probability that it will be blocked by one or more obstacles. The Line Of Sight (LOS) probability function $P(D)$ in the network is derived from stochastic blockage models, where the blockage parameters are characterised by random distributions. In this paper, the blockages are modelled as a rectangle Boolean scheme as in [150]. It is shown that the LOS probability function $P(D)$ exponentially decays with the link length. Consequently, the probability that RRH is LOS with user k is represented by $p(D) = e^{-\zeta D}$. ζ is parameter that is calculated from the density and average size of the blockages. $1/\zeta$ is called the LOS range and $1/\zeta = 141.4$ metres[151]. The Non-Line Of Sight (NLOS) or (blockage) probability is expressed as $1 - P(D)$. Since the mmw channel is affected by blockage, there are different values of path-loss parameters for LOS and NLOS users in the mmw channel model [150].

SINR received by user k served by mmw RRH r_i is represented by the following equation:

$$\gamma_{k_{R_i}} = \frac{\mathrm{G}_{r_i}\beta_{k_{r_i}}\mathrm{P}_{r_i}}{\sum_{j\in R, j\neq i}\mathrm{G}_{r_j}\beta_{k_{r_j}}\mathrm{P}_{r_j} + \mathrm{N}_0} \tag{5.4}$$

Where, P_{r_i} is the transmit power between serving RRH r_i and the user k. G is the directivity gain of the transmission uniform linear array antenna and its array pattern is approximated using a sectored antenna model. The transmit antenna pattern is assumed to be 10 dB, -10 dB and 30^o, which represent main lobe directivity gain, back lobe gain, and main lobe beamwidth, respectively [151]. $\beta_{k_{r_i}}$ is the channel gain between user k in the cell edge area, with its serving RRH r_i, which is a composite fading channel that involves small scale fading, large scale fading and path loss, given by:

$$\beta_{k_{r_i}} = \mathrm{g}^*_{k_{r_i}} l_{k_{ri}} \left[\mathrm{AD}^{-\delta_{mm}}_{k_{r_i}}\right] \tag{5.5}$$

where, $\mathrm{g}^*_{k_{r_i}}$ and $l_{k_{r_i}}$ represent the small and large scale fading channel, respectively, between the user k and its serving RRH r_i. The small scale fading is assumed to be a Nakagami with normalised gamma distribution[131] and its parameters are $\mathrm{N}_L = 3$ and $\mathrm{N}_N = 2$. The large scale fading is assumed to be a lognormal random variable with a standard deviation of 3.1 dB for LOS and 8.2 dB for NLOS [104]. $\mathrm{AD}^{-\delta_{mm}}_{k_{r_i}}$ reflects the

path-loss between the serving RRH r_i and user k, where A is a constant, which depends on the carrier frequency f_c and can be calculated from eq. 5.3 The LOS and NLOS path loss exponents are $\delta_{mm_L} = 2$ and $\delta_{mm_N} = 4$ respectively. $\sum_{j\in R, j\neq i} \mathrm{G}_{r_j} \beta_{k_{r_j}} \mathrm{P}_{r_j}$ represents the inter-cell interference power received from the RRHs in different clusters.

5.2.3 Users Association Scheme

The users in the cell centre area are associated with an MBS based on the maximum received power. The remaining MBSs operating at the μw frequency band are considered as interferes. At the cell edge area where the MBS received power is poor, the users are associated with the mmw base station based on the closest BS association. Open access is assumed in this work, which means that users can be associated with a mmw RRH or μw BS, depending on the maximum received downlink power. However, as the macro BS transmits higher power than the mmw RRHs, positive power biasing is adopted to offload more edge users from the former to the latter and a user associates itself to the nearest BS according to:

$$j = \arg \max_{j\in\{R,M\}} P_j \, B_j \, |D_{K_{ji}}|^{-\delta_j} \tag{5.6}$$

where, B_j is a biasing factor for user association, B_M=0 dB is for the users associated with the macro BS in the cell centre area based on maximum received power and $B_R > 0$ is for users associated with the mmw RRHs in the cell edge area based on maximum received biased power. In this work, a B_R=10 dB biasing factor is adopted for small cell users.

5.2.4 Small Cells Clustering Algorithm

In order to avoid inter-cluster interference and frequent hand overs between mmw RRHs, it is assumed that the neighbouring RRHs are assigned to the same BBU in a compact cluster so that there is no interference between neighbouring small cells[124]. Also, the neighbouring cells that belong to a different cluster are assumed to use different sub bands. A dynamic load aware small cell clustering algorithm is proposed to share the mmw BBU

resources with minimum inter-cluster interference. There are $\mathscr{M}$ BBUs in the BBU pool each has D Physical Resource Blocks (PRBs), and a set of R RRHs $r_1, ...r_R$ each with PRB demand $d_1, ...d_R$. The RRHs proximity is defined by introducing a binary variable A_{ij}, where A_{ij}= 1, if r_i and r_j are adjacent else A_{ij}= 0 when they share no border. If a cluster has multiple RRHs, then the RRHs in that cluster must be adjacent and connected. To formulate the connectedness of a cluster and proximity of the RRHs, let S1 be any proper subset of the set of RRHs served by BBU_m (Z_m), such that S1 $\subset Z_m$, S1 $\neq \emptyset$, and S1 $\neq Z_m$. Let S2 be another subset of Z_m such that, S2 $= Z_m -$ S1, i.e., S2 is the complementary set of S1.

The problem consists of finding a number of required BBUs to be activated $\mathscr{M}$ based on the demand of edge users and the number of cluster C. The optimal RRH clustering solution is found when it has minimum $\mathscr{M}$.

$$\text{Minimum} \quad \mathscr{M} = \sum_{r=i}^{R} y_i, \quad subject \quad to:$$

$$C_1 : \sum_{j=S1} \sum_{j \in S2} A_{ij} \geq 1, \quad \forall S1,S2 \in Z_m, \forall m \in \{1,2,...M\}$$

$$C_2 : \sum_{j=1}^{R} d_j x_{ij} \leq D_{y_i}, \qquad \forall i \in \{1,2,...R\}$$

$$y_i \in \{0,1\}, \quad \forall i \in \{1,2,...R\}$$

$$x_{ij} \in \{0,1\}, \quad \forall i \in \{1,2,...R\},$$

$$\forall j \in \{1,2,...R\} \quad (5.7)$$

where $y_i = 1$ if BBU i is used, $x_{ij} = 1$ if r_j is associated to BBU i. To find the optimal solution, first the bin packing problem is divided into subproblems and find the optimal solution for each. Then, the optimum solution is found by taking the best among all [160].

5.3 Coverage Probability

In this section, the coverage probability is presented, which is the probability that a user located in a defined cell area has an SINR greater than a predefined threshold SINR ($\mathscr{T}_C$). It is also equivalent to the complementary cumulative distribution function (CCDF) of the SINR.

$$\mathscr{C}(\mathscr{T}_c) = \mathbb{P}(SINR > \mathscr{T}_C) \tag{5.8}$$

Coverage is a critical metric to consider as it is a measure of link reliability and it can give insight into the cell edge user and overall cell QoS. The coverage probability of the suggested mmw HC-RAN system can be given by the following equation:

$$\mathscr{C}_{HC-RAN}(\mathscr{T}_c) = \mathscr{A}_R \mathscr{C}_{R_{mmw}}(\mathscr{T}_c) + \mathscr{A}_M \mathscr{C}_{M_{\mu\mathrm{w}}}(\mathscr{T}_c) \tag{5.9}$$

or

$$\mathscr{C}_{HC-RAN} = \left(\bigcup_{j \in \{R,M\}} \mathscr{A}_j \, \mathbb{P}\left(\gamma_{k_{j_i}} > \mathscr{T}_C\right) \right) \tag{5.10}$$

where, $\mathscr{A}_{mmw}$ and $\mathscr{A}_{\mu\mathrm{w}}$ are the association probabilities of mmw and μw users, respectively. $\mathscr{C}_{mmw}$ and $\mathscr{C}_{\mu\mathrm{w}}$ are the coverage probabilities of mmw and μw users, respectively, which are conditioned on users' association to the specific tier.

5.4 Rate Coverage Probability

The above mentioned coverage probability reflects the proximity effect, i.e. the ($UE - MBS \setminus RRHs$) association, but it does not capture the rate achieved by the associated user. In an open access scheme, a UE is considered within rate coverage, if its downlink rate is greater than a target threshold. If the predefined threshold rate is denoted as ($\mathscr{T}_r$), then the rate coverage is given by the following:

$$\mathscr{R}(\mathscr{T}_r) = \mathbb{P}(\mathscr{R} > \mathscr{T}_r) \tag{5.11}$$

The rate achieved by user k is given by:

$$\mathscr{R}(k_{j_i}) = log_2(1+\gamma_{k_{ji}}), \quad \text{j} \in \{R,M\} \tag{5.12}$$

The rate coverage for the suggested system according to the total probability law is given by:

$$\mathscr{R}_{HC-RAN} = \mathscr{A}_R \mathscr{R}_{R_{mmw}} + \mathscr{A}_M \mathscr{R}_{M_{\mu\text{w}}} \tag{5.13}$$

where, $\mathscr{R}_{R_{mmw}}$ and $\mathscr{R}_{M_{\mu\text{w}}}$ are rate coverages that are conditioned on users association to the specific tier. Open access is assumed where users are allowed to access μw or mmw BSs. The rate coverage can also be given as follows:

$$\mathscr{R}_{HC-RAN} = \left(\bigcup_{j\in\{R,M\}} \mathscr{A}_j \, \mathbb{P}(log_2(1+\gamma_{k_{j_i}}) > \mathscr{T}_r) \right) = \left(\bigcup_{j\in\{R,M\}} \mathscr{A}_j \, \mathbb{P}((\gamma_{k_{j_i}}) > (2^{\mathscr{T}_r}-1) \right) \tag{5.14}$$

5.5 Optimising the Small Cells Size

For further investigation into the impact of the suggested scheme on the system performance, an optimisation problem is formulated to find the optimum small cells radii that give the maximum average cell rate and the best rate coverage probability as an indicator for a good QoS. The optimisation results are compared to the suggested system to consolidate it.

The number of small mmw cells is assumed to be R and their radii are denoted as r_{mm}, which defines the size of edge area. Now, a problem for finding the optimum mmw cells' radii is presented, i.e. the size of mmw edge cell to achieve the highest overall average cell rate and cell rate coverage, which is defined as:

$$\text{r}_{mm} = \arg\max \left(E[\mathscr{R}_{HC-RAN}(\text{r}_{mm})] \right), \;\; \text{r}_{mm} > 0 \tag{5.15}$$

where:

$$\mathrm{E}[\mathscr{R}_{HC-RAN}(\mathrm{r}_{mm})] = \frac{1}{K}\left(\sum_{n=1}^{R}\sum_{i=1}^{k_{mmr}} B_{mm}\, log_2(1+\gamma_{k_{ni}}) + \sum_{i=1}^{k_M} B_{\mu\mathrm{w}}\, log_2(1+\gamma_{k_{mi}})\right) \quad (5.16)$$

where, k_{mmr} is the number of users served by mmw RRH r and k_m is the number in the cell centre area served by MBS m. For each given value of r_{mmn}, the value of the average macrocell rate is computed as the edge area served by the mmw cell increase with larger r_{mm} value. Finding the optimum value of mmw cell size that gives the maximum average rate is achieved by maximising eq. 5.16. As mentioned earlier, the rate coverage is an important metric that gives insight about the edge users' QoS. The problem to find the mmw small cell radii to achieve the highest rate coverage is defined in eq. 5.15

where, $\mathscr{R}_{HC-RAN}$ is presented in eq. 5.14. For each value of r_{mm} in this problem there is a different value of the rate coverage, these values are sorted and the r_{mm} gives highest value, which is adopted as the optimum radius for the mmw small cells.

5.6 Computational Results and Performance Analysis

Consider a two-tier mmw HC-RAN system comprising nineteen hexagonal MBSs as the first tier, and overlaid by mmw small BSs placed at centre points on the macrocell edge area as the second tier. The cell area is divided into two spaces: inner region and an outer region. The inner area is served by a traditional BS using standard μw LTE-A band and the edge area is served using mmw RRHs with a 28 GHz frequency band for cellular access. The suggested scheme is based on the legacy SFR technique for interference mitigation and improving the cell edge users' rate and coverage. The proposed scheme enhances the over all cell rate and coverage rather than just the edge area. The system architecture is simulated using the MATLAB programme and system performance is validated using Monte Carlo analysis. To examine the proposed system performance, different metrics have been calculated with the simulation parameters listed in Table.5.1.

Table 5.1 Simulation parameters

Parameter	Value
Macro cell's radius (r_M) and small cells' radius (r_{mm})	500 m, 125m
Macro cell and small cells' carrier frequency	2 GHz, 28 GHz
P_M, P_R	46dBm, 30dBm
$\delta_M, \delta_{mm_L}, \delta_{mm_n}$.	3, 4, 2
Noise power density	-174 dBm/Hz

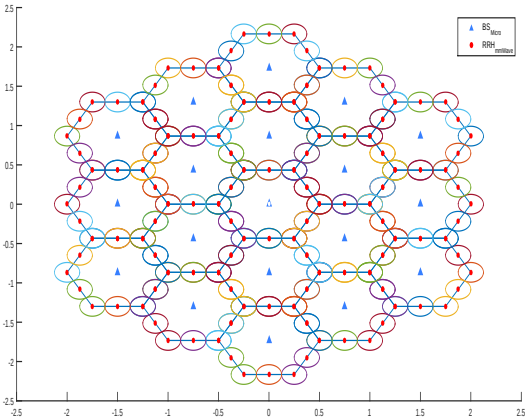
5.6.1 Coverage Probability Comparison

First, the coverage probability of an mmw HC-RAN network applying the suggested scheme is analysed. In order to verify the performance improvement realised by adopting this scheme, it is compared to the coverage probability of the mmw HC-RAN network with the same of number randomly deployed mmw RRHs and same network set up. Furthermore, it is compared to the traditional system applying the universal Frequency Reuse (UFR) and SFR schemes with reuse factor of 3. The coverage probability is affected by various system parameters, such as the SINR threshold and the number of the mmw RRHs. Fig. 5.3 presents the obtained numerical simulation results for coverage probability (i.e. the CCDF of $\mathscr{T}_C$) of a macrocell for different configuration schemes with the same number of users. The coverage probability changes with different values of $\mathscr{T}_C$, and converges at higher values of $\mathscr{T}_C$ for all the scenarios. It is observed from Fig. 5.3 that there is a clear improvement in the coverage probability over the whole cell area for HC-RAN when applying the suggested scheme compared to the other schemes.

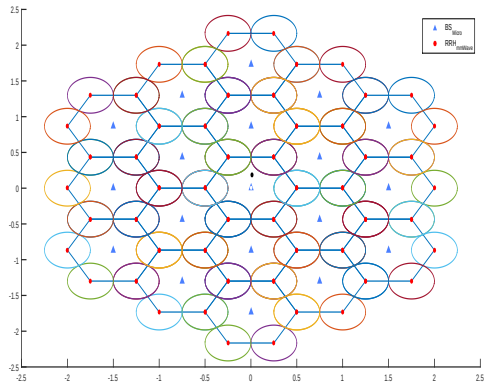
For example, at $\mathscr{T}_C$=-10db, when the suggested scheme is applied for HC-RAN, 99% of users are covered compared to 97%, 95% and 87% when applying the randomly located mmw RRHs, SFR and UFR schemes, respectively. At higher $\mathscr{T}_C$ values, for example 10dB, the average fraction of covered users drops to 60% for the proposed scheme compared to 40%, 35% and 30% for SFR, randomly distributed RRHs and UFR schemes, respectively, which gives a coverage gain of (20%-30%) for the proposed scheme over the other schemes.

The density of the mmw RRHs also affects the total coverage probability. Fig. 5.3 also reveals that a higher coverage probability is obtained when a higher number and

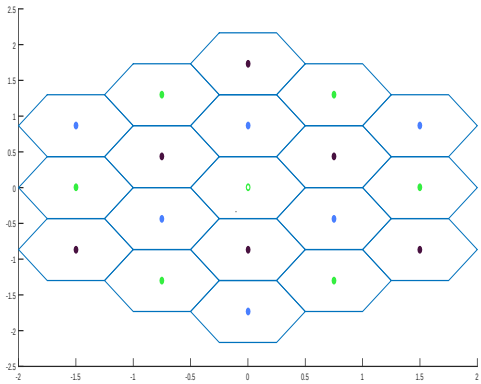
smaller size of mmw cells are used to serve the cell edge area compared to a lesser number and bigger size mmw cells (20% average coverage gain for the proposed scheme with 12 mmw RRHs and a small cell radii of 125m compared to 6 mmw RRHs and a 250m small cells radii). This is because at lower RRH density the average distance between the UE and the serving RRH is larger, which means considerably higher path losses and a lower received signal for the mmw links. Consequently, there is lower received SNR for users served by mmw RRHs and lower total coverage probability. However, even with less density of mmw RRHs, the proposed scheme still delivers higher coverage probability than the other schemes, as shown in Fig. 5.3.



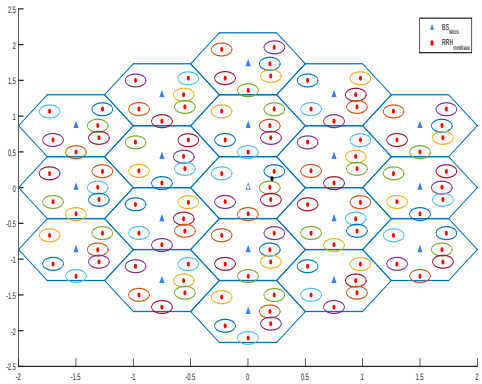
(a) Proposed scheme 12 RRH



(b) Proposed scheme 6 RRH



(c) SFR



(d) Randomly distributed mmw-RRHs

Fig. 5.2 Comparison scenarios

5.6.2 Rate Coverage Probability Comparison and Throughput Calculation

As discussed above, the rate coverage is the probability that the achieved data rate is higher than a threshold value $\mathscr{T}_r$. Rate coverage is considered as a QoS indicator that means how many users are satisfied with their QoS. Its probability varies with $\mathscr{T}_r$ values and it decreases as the $\mathscr{T}_r$ value increases. It is observed from Fig. 5.4 that with HC-RAN applying the proposed scheme there is considerably higher rate coverage than for the other schemes for the same network set up. Specifically, the average gain in rate coverage is 48%, 37% and 33% compared to SFR, UFR, HC-RAN with randomly distributed mmw RRHs, respectively. The best total rate coverage is experienced when adopting the proposed scheme, even for lower RRH density , which comes from the higher available bandwidth and higher received SINR values as the interference levels are reduced, as shown in Fig. 5.8.

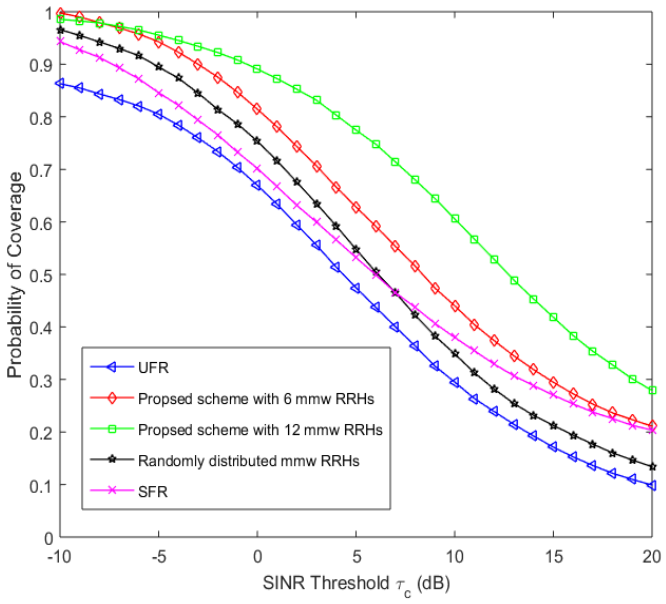


Fig. 5.3 Coverage probability comparison of the proposed system with others scenarios

For instance, an average rate coverage probability gain of 10% is obtained when 12 mmw RRHs are used to serve the cell edge area compared to 6 mmw RRHs for the same macrocell area and network set up. Furthermore, for $\mathscr{T}_r$ values lower than 0.3 bits/sec/Hz, the total rate coverage is higher when using SFR than UFR, but it drops for $\mathscr{T}_r$ values

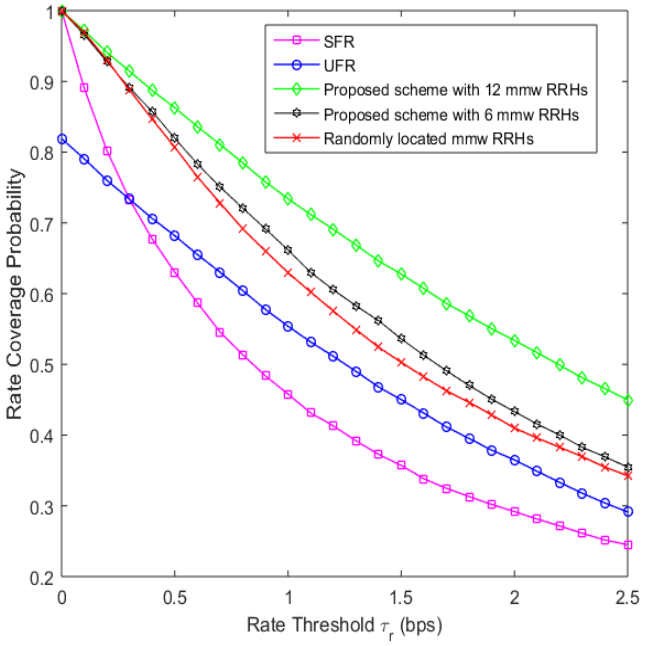


Fig. 5.4 Proposed system rate coverage probability comparison with other scenarios

higher than 0.3 bits/sec/Hz for the same network set up. This is because when using SFR, lower achievable rate is obtained, as available bandwidth is divided between the cell edge and centre areas.

Furthermore, to have more insight about the network performance, the average users throughput has been calculated for different users densities. The demand d_i of RRH_i is the sum of users demand associated with RRH_i. The end users' different data rates are proposed based on 3GPP standard simulation parameters, which are 4-25 kbps for audio, 32-384 kbps for video, 28.8 kbps for data, and 60 kbps for real-time gaming services[142].

5.6.3 Optimum Small Cell Radius

As illustrated earlier in section 5.5, the size of the small cells used to cover the cell edge also has an impact on the system performance, as the radii of the small cells define the size of the edge area served by the mmw band. Fig. 5.5 shows the small cells size impact on the total coverage probability for $\mathscr{T}_c$ values between-10dB and 30dB. It is clear from the figure, that as the small cells' radii increase the average probability of coverage also increases. Furthermore, the curve reveals that the optimum small cells radii that gives the

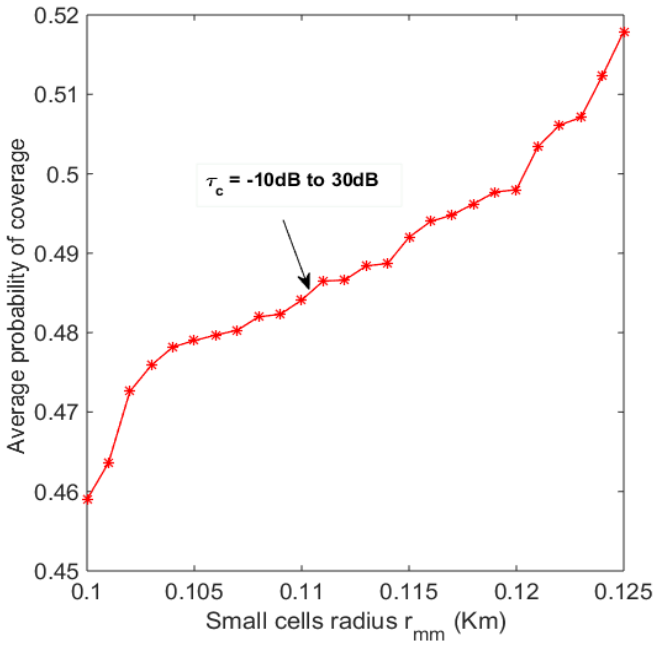


Fig. 5.5 Small cells' radii versus average coverage probability for $\mathscr{T}_c$ values (from-10dB to 20dB)

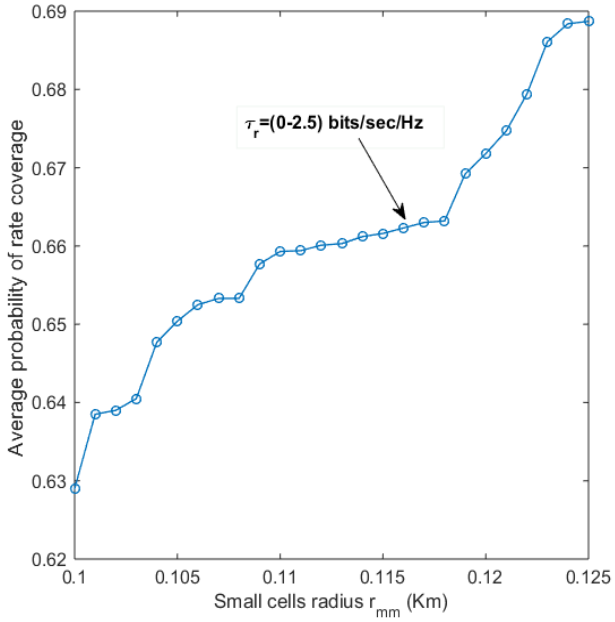


Fig. 5.6 Small cells' radii effect on the average rate coverage probability for $\mathscr{T}_r$ values of (0-2.5) bps/Hz

highest average probability of coverage is 0.125Km, which validates the assumption made for the system configuration of the proposed scheme. Fig. 5.6 demonstrates the impact of small cell size on the rate coverage probability. The figure also illustrates that the rate coverage increases proportionally with an increase in small cells' radii. Furthermore, the

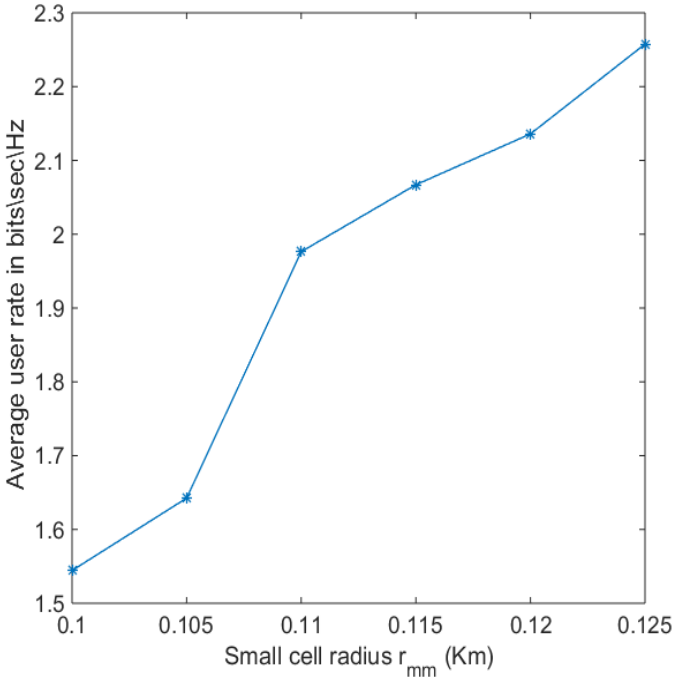


Fig. 5.7 Average user rate bps/Hz

maximum average rate coverage is obtained at small cells' radii of 0.125Km. Moreover, the average rate per user trend as a function of mmw cells' radii is presented in Fig. 5.7, which shows that as the radii increase the average rate increases. Also, as shown in Fig. 5.5 and Fig. 5.6 the SINR coverage and the rate coverage increases with the increase in the small cells' radii.

Finally, the effect of interference on the HC-RAN network when applying the proposed scheme is investigated as a function of the edge area size, which is represented by small cells' radii. When observing Fig. 5.8, it is clear that the total interference level drastically decrease when adopting the proposed scheme. Furthermore, It is still evident from the figure that the optimum small cells radii that gives minimum aggregated interference for the macrocell users is 0.125Km, which is the same small cells' radii that deliver the best coverage and rate.

Furthermore, to have more insight about the network performance, the average users throughput has been calculated for different users densities. The demand d_i of RRH_i is the sum of users demand associated with RRH_i. The end users' different data rates are proposed based on 3GPP standard simulation parameters, which are 4-25 kbps for audio,

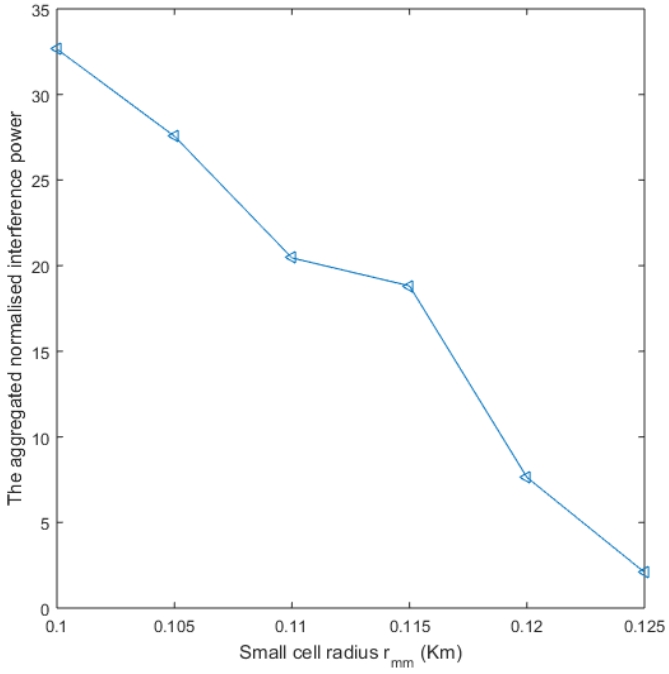


Fig. 5.8 Aggregated users' interference

32-384 kbps for video, 28.8 kbps for data, and 60 kbps for real-time gaming services[142]. The users are normally distributed and each user's physical resource block demand is equal to the ratio of requested data rate to the achievable data rate as shown previously. The average network throughput has been compared for two schedulers (Proportional Fair (PF) and Round Robin (RR)). As shown in Fig. 5.9, there is a noticeable increase in the average network throughput using the proposed scheme compared to the conventional network with UFR. For both scenarios, using PF scheduler improved the average network throughput compared to RR scheduler as PF guarantees minimum service to each user regardless the SINR value.

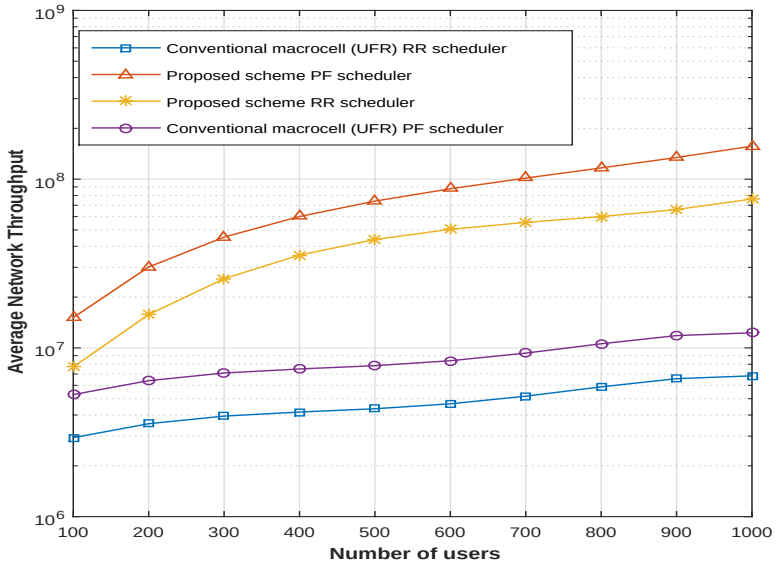


Fig. 5.9 Average network throughput *bps*

5.7 Summary

The 5G is going to first be deployed in Non-Stand Alone mode (NSA) working together with the existing LTE-A. In this study, a scheme for mmw HC-RAN system has been proposed for the early stages of 5G deployment. This scheme is aimed at exploiting the mmw frequency properties of wide band and short range to alleviate the cell edge problems of a poor received signal and high interference. In this scheme, the cell area is divided into two spaces as in the traditional SFR technique, a centre area served by a conventional μW base station and an edge area served by a mmw band by deploying RRHs in a wrap around model. The coverage and rate trend for the proposed scheme have been investigated. The impact of small cells size has also been probed, as it defines the size of edge area. Also, the average network for the suggested scheme using both PF and RR schedulers has been inspected and compared the traditional system. The simulation results show that the proposed scheme achieves higher coverage probability and rate coverage compared to the SFR and UFR schemes. Moreover, it dramatically reduces the interference level at the cell, whilst creating an edge area with high user rate and no interference by suggesting RRHs clustering algorithm.

Chapter 6

Conclusion and Future work

This chapter completes the thesis by summarising the main contributions in section 6.1. Also, some ideas about open research issues for future work are presented in section 6.2.

The tremendous growth of wireless communication systems plus the emerging applications that have high capacity requirements, have brought forward the need for a larger spectrum. This requirement leads to the call for research on increasing band width efficiency and exploiting higher frequency bands for cellular access. The future aims have been outlined by the international telecommunication union (ITU) for international mobile telecommunications in the next generation 5G usage scenarios, which are eMBB, URLLC and mMTC.

To this end, this work combines some of the significant 5G enabling technologies to achieve high capacity system without sacrificing the network energy efficiency. This is by a exploiting self-optimised feature in C-RAN architecture to introduce a network that adapt to the variable capacity demand efficiently.

6.1 Conclusion

For this thesis, several 5G enabling technologies have been combined, which are C-RAN flexible architecture, operation in the mmwave band and the self-optimising feature of SON to enhance the QoS. The main contributions of this research are as follows.

In Chapter 2, there was an overview about the fundamentals of the technologies that has been used in this thesis, such as C-RAN, self-organising networks (SON) and the

fifth generation wireless access technology New Radio (NR).

In Chapter 3, an algorithm that exploits the capacity routing characteristic of C-RAN was proposed to scale the capacity of the network based on the traffic demand. The algorithm is based on the concept of cell differentiation and integration. The network load is monitored after a time cycle and at low traffic conditions the macrocell area is served by high power base station. When the traffic load increases to past the macrocells capacity resource limit, the macrocell geographical area is differentiated into small cells to increase the bandwidth efficiency. In addition, these small cells can be further differentiated into smaller cells to satisfy any capacity shortage. The BBU-RRH association is dynamically optimised to enhance the overall QoS. RRH clustering is formulated as optimisation problem, whereas the QoS fitness function is the linear sum of multi-objective functions, which target maximisation of average network load and load fairness-index, while at the same time minimising the number of handovers during the BBUs-RRHs re-association. A SON algorithm installed in the BBU cloud is responsible for the cell differentiation and integration as well as performing the dynamic RRH clustering. A tool from AI known as an evolutionary algorithm, specifically Particle Swarm Optimisation (PSO), which has been widely used to solve self-optimising problems in LTE, is used to solve the optimisation problem. The CDI enabled C-RAN shows notable improvement in throughput and a significant reduction in the number of blocked users. Furthermore, a power model has been developed to estimate the power consumption by the C-RAN architecture. The obtained results show that the CDI-enabled C-RAN delivers considerable power saving compared to fixed C-RAN.

In Chapter 4, mm wave operation for small cells access has been tested for the CDI enabled C-RAN. The 5G NR mmw band frequency range2 FR2 operation was investigated. Furthermore, two evolutionary algorithms, namely a GA and DPSO, were used to solve the BBU-RRH association optimisation problem. A complexity comparison was made between the two algorithms. An order of magnitude increase in the average network

throughput was observed without sacrificing the overall energy efficiency of the network. Moreover, the average number of blocked users is notably decreased. This network performance showing significant improvement without energy efficiency deterioration is owing to the adoption of s self-optimising feature plus the inter operation between the μW band and the wide band mm wave frequency.

Finally, in Chapter 5, a scheme for a 5G early deployment NSA scenario, which also benefits from the self-optimisation feature of SON, was introduced for interference mitigation in a wireless network. 5G in its preliminary stages is going to coexist with LTE-A using the same infrastructure in a NSA scenario. However, since LTE-A uses unity frequency reuse, the problem of interference is affecting the edge area throughput, which reaches one tenth of the peak throughput. The proposed scheme to address this was inspired by soft frequency reuse SFR scheme, where the cell area is divided into two spaces: inner region and outer region. Each region is served by a different frequency band, with the 5G mm wave small cell for cellular access being deployed in the edge area to reduce interference and increase throughput. There is a considerable enhancement in coverage and rate of the network when applying proposed scheme compared with the traditional SFR scheme and a randomly deployed mm wave small cells in HC-RAN system.

To sum up, the added SON features to the C-RAN architecture can achieve efficient resource utilisation and also increases the system capacity by increasing available spectrum efficiency while maintaining high QoS. Furthermore, the proposed algorithm is shown to achieve optimum solution when using NR FR2 and small LTE-A networks. It reaches near optimum for the LTE-A spectrum in the case of large networks. This consolidates the proposed notion of the benefits of SON in C-RAN.

6.2 Future Research Directions

This section presents some future directions to extend the work in this thesis.

The proposed CDI SON algorithm must be executed in a few milliseconds to adapt to load and channel variations, which is particularly important for proactive performance of the network. Regarding which, the CDI execution time involves collecting KPIs, executing the SON algorithm and the time for realising the new configuration by switching and remapping the BBUs and RRHs.For this reason, to save the execution a mobility prediction algorithm, such as a Markov chain for mobility prediction, is necessary for predicting the next time interval load to find BBU-RRH allocation and gainful resource optimisation. Furthermore, a tool from stochastic geometry could be used to calculate the average SINR and throughput of the network to give more insight into the lower bound system performance instead of time consuming Monte-Carlo analysis.
The high capacity requirements of 5G use cases, which are eMBB, URLLC and mMTC, will also impose high capacity demands on the fronthaul. For this reason, the front haul capacity could be one of the constraints in designing the SON algorithms, especially the centralised ones. Moreover, to alleviate the load the fronthaul, functional split between the BBUs and RRHs could be optimised based on the use case scenario as well as capacity and time delay requirements.
Similarly, massive traffic is expected on the backhaul between BBU pools and core network due to the 5G requirement. This load could be beyond the BBU or backhaul capacity. Consequently, load distribution between the BBU pools could be considered as an optimisation problem with backhaul capacity and delay constraints.

CoMP is an appealing technique for increasing capacity and reducing interference through joint transmission or coordinated beamforming/scheduling. That is, these CoMP features could be utilised to extend the work carries out for the current study. Since the CoMP requires multi RRHs communicating with UE, the RRH clustering proximity constraint could pave the way for integrating CoMP communication in future work.

For the H-CRAN in Chapter 5, a transmission power control for each small cell could be taken as an optimisation problem. In addition, the study of optimal cell biasing for different parameter settings of the network is another possible research direction.

Bibliography

[1] Cisco Visual Networking Index. "Global mobile data traffic forecast update, 2017–2022". *Cisco White paper*, 2019.

[2] A Minokuchi, S Isobe, H Takahashi, and S Nagata. "5G Standardisation Trends at 3GPP". *NTT DOCOMO Technical Journal*, 19(3):5–12, 2018.

[3] 3GPP. "5G; NR; Base Station (BS) radio transmission and reception (Release 15) ". TS V15.2.0, 3rd Generation Partnership Project (3GPP), July 2018.

[4] Aleksandra Checko, Henrik L Christiansen, Ying Yan, Lara Scolari, Georgios Kardaras, Michael S Berger, and Lars Dittmann. "Cloud RAN for mobile networks—A technology overview". *IEEE Communications surveys & tutorials*, 17 (1):405–426, 2015.

[5] Andreas F Molisch, Vishnu V Ratnam, Shengqian Han, Zheda Li, Sinh Le Hong Nguyen, Linsheng Li, and Katsuyuki Haneda. "Hybrid Beamforming for Massive MIMO: A Survey". *IEEE Communications Magazine*, 55(9):134–141, 2017.

[6] ETSI LTE. "Evolved universal terrestrial radio access (e-utra) and evolved universal terrestrial radio access network (e-utran)(3gpp ts 36.300, version 8.11. 0 release 8), december 2009". *ETSI TS*, 136(300):V8, 2011.

[7] Mugen Peng, Chonggang Wang, Vincent Lau, and H Vincent Poor. "Fronthaul-constrained cloud radio access networks: Insights and challenges". *IEEE Wireless Communications*, 22(2):152–160, 2015.

[8] Ning Zhang, Nan Cheng, Amila Tharaperiya Gamage, Kuan Zhang, Jon W Mark, and Xuemin Shen. "Cloud assisted HetNets toward 5G wireless networks". *IEEE communications magazine*, 53(6):59–65, 2015.

[9] Mugen Peng, Yuan Li, Jiamo Jiang, Jian Li, and Chonggang Wang. "Heterogeneous cloud radio access networks: A new perspective for enhancing spectral and energy efficiencies". *IEEE Wireless Communications*, 21(6):126–135, 2014.

[10] Mugen Peng, Yong Li, Zhongyuan Zhao, and Chonggang Wang. "System architecture and key technologies for 5G heterogeneous cloud radio access networks". *IEEE network*, 29(2):6–14, 2015.

[11] Mugen Peng, Kecheng Zhang, Jiamo Jiang, Jiaheng Wang, and Wenbo Wang. "Energy-efficient resource assignment and power allocation in heterogeneous cloud radio access networks". *IEEE Transactions on Vehicular Technology*, 64 (11):5275–5287, 2015.

[12] Stefania Sesia, Matthew Baker, and Issam Toufik. *"LTE-the UMTS long term evolution: from theory to practice"*. John Wiley & Sons, 2011.

[13] Supratim Deb, Pantelis Monogioudis, Jerzy Miernik, and James P Seymour. "Algorithms for enhanced inter-cell interference coordination (eICIC) in LTE HetNets". *IEEE/ACM transactions on networking*, 22(1):137–150, 2014.

[14] Juho Lee, Younsun Kim, Hyojin Lee, Boon Loong Ng, David Mazzarese, Jianghua Liu, Weimin Xiao, and Yongxing Zhou. "Coordinated multipoint transmission and reception in LTE-advanced systems". *IEEE Communications Magazine*, 50(11), 2012.

[15] Daewon Lee, Hanbyul Seo, Bruno Clerckx, Eric Hardouin, David Mazzarese, Satoshi Nagata, and Krishna Sayana. "Coordinated multipoint transmission and reception in LTE-advanced: deployment scenarios and operational challenges". *IEEE Communications Magazine*, 50(2), 2012.

[16] Aleksandra Checko, Henrik L Christiansen, Ying Yan, Lara Scolari, Georgios Kardaras, Michael S Berger, and Lars Dittmann. "Cloud RAN for mobile networks—A technology overview". *IEEE Communications surveys & tutorials*, 17 (1):405–426, 2015.

[17] Mugen Peng, Chonggang Wang, Vincent Lau, and H Vincent Poor. "Fronthaul-constrained cloud radio access networks: Insights and challenges". *IEEE Wireless Communications*, 22(2):152–160, 2015.

[18] Mamta Agiwal, Abhishek Roy, and Navrati Saxena. "Next generation 5G wireless networks: A comprehensive survey". *IEEE Communications Surveys & Tutorials*, 18(3):1617–1655, 2016.

[19] "On global electricity usage of communication technology: trends to 2030", author=Andrae, Anders and Edler, Tomas. *Challenges*, 6(1):117–157, 2015.

[20] Evolved Universal Terrestrial Radio Access Network. "Self-configuring and self-optimizing network (SON) use cases and solutions". TR 36.902 V9.3.1, 2012.

[21] NGMN Alliance. "NGMN Informative List of SON Use Cases". *Annex deliverable*, April 2007.

[22] "Challenges in 5G: how to empower SON with big data for enabling 5G", author=Imran, Ali and Zoha, Ahmed and Abu-Dayya, Adnan. *IEEE network*, 28(6): 27–33, 2014.

[23] Jessica Moysen and Lorenza Giupponi. "From 4G to 5G: Self-organized network management meets machine learning". *Computer Communications*, 2018.

[24] European Commission. "FP7 Project SOCRATES. Self-Optimisation and Self-Configuration in Wireless Networks." URL http://www.fp7-socrates.eu/.

[25] TELUS and Huawei. "Next Generation of SON for 5G". *white paper*, 2016.

[26] Jun Xu, Junmei Yao, Lu Wang, Kaishun Wu, Lei Chen, and Wei Lou. "Revolution of self-organizing network for 5G mmwave small cell management: from reactive to proactive". *IEEE Wireless Communications*, 25(4):66–73, 2018.

[27] Olav Østerbø and Ole Grøndalen. "Benefits of self-organizing networks (SON) for mobile operators". *Journal of Computer Networks and Communications*, 2012, 2012.

[28] 3GPP. "5G; Study on Scenarios and Requirements for Next Generation Access Technologies (Release 14)". TS 38.913 v14.2.0 , 3rd Generation Partnership Project (3GPP), May 2017.

[29] 3GPP. "5G; Management and orchestration; 5G end to end Key Performance Indicators (KPI) (Release 15)". TS 28.554 V 15.0.1 , 3rd Generation Partnership Project (3GPP), 2018.

[30] Yonghua Lin, Ling Shao, Zhenbo Zhu, Qing Wang, and Ravie K Sabhikhi. "Wireless network cloud: Architecture and system requirements". *IBM Journal of Research and Development*, 54(1):4–1, 2010.

[31] K Chen and R Duan. "C-RAN: The Road Towards Green RAN", White Paper Version 3.0, China Mobile Research Institute, Dec. 2013.

[32] Mugen Peng, Yaohua Sun, Xuelong Li, Zhendong Mao, and Chonggang Wang. "Recent advances in cloud radio access networks: System architectures, key techniques, and open issues". *IEEE Communications Surveys & Tutorials*, 18(3): 2282–2308, 2016.

[33] Dimitrios Pliatsios, Panagiotis Sarigiannidis, Sotirios Goudos, and George K Karagiannidis. "Realizing 5G vision through Cloud RAN: technologies, challenges, and trends". *EURASIP Journal on Wireless Communications and Networking*, 2018(1):136, 2018.

[34] 3GPP. "Technical Specification Group Radio Access Network; Evolved universal terrestrial radio access (E-UTRA); Medium Access Control (MAC) protocol specification (Release 14)". TS 36.321 v14.2.1, 3rd Generation Partnership Project, March 2017.

[35] 3GPP. "Technical Specification Group Radio Access Network; Evolved universal terrestrial radio access (E-UTRA); Radio Link Control (RLC) protocol specification (Release 14)". TS 36.322 v14.0.0, 3rd Generation Partnership Project, March 2017.

[36] M. Denker. "Remote Radio Head systems - Requirements and concept of lightning and over voltage protection". In *Intelec 2013; 35th International Telecommunications Energy Conference, SMART POWER AND EFFICIENCY*, pages 1–5, Oct 2013.

[37] Mohammad Asif Habibi, Meysam Nasimi, Bin Han, and Hans D Schotten. "A Comprehensive Survey of RAN Architectures towards 5G Mobile Communication System". *IEEE Access*, 2019.

[38] Osvaldo Simeone, Andreas Maeder, Mugen Peng, Onur Sahin, and Wei Yu. "Cloud radio access network: Virtualizing wireless access for dense heterogeneous systems". *Journal of Communications and Networks*, 18(2):135–149, 2016.

[39] Line MP Larsen, Aleksandra Checko, and Henrik L Christiansen. "A survey of the functional splits proposed for 5G mobile crosshaul networks". *IEEE Communications Surveys & Tutorials*, 21(1):146–172, 2019.

[40] Md Farhad Hossain, Ayman Uddin Mahin, Topojit Debnath, Farjana Binte Mosharrof, and Khondoker Ziaul Islam. "Recent research in cloud radio access network (C-RAN) for 5G cellular systems-A survey". *Journal of Network and Computer Applications*, 2019.

[41] Vinay Suryaprakash, Peter Rost, and Gerhard Fettweis. "Are heterogeneous cloud-based radio access networks cost effective?" *IEEE Journal on Selected Areas in Communications*, 33(10):2239–2251, 2015.

[42] Jun Wu, Zhifeng Zhang, Yu Hong, and Yonggang Wen. "Cloud radio access network (C-RAN): a primer". *IEEE,Network*, 29(1):35–41, Jan 2015.

[43] Raad S Alhumaima, Muhammad Khan, and Hamid S Al-Raweshidy. "Component and parameterised power model for cloud radio access network". *IET Communications*, 10(7):745–752, 2016.

[44] Riccardo Bassoli, Marco Di Renzo, and Fabrizio Granelli. "Analytical energy-efficient planning of 5G cloud radio access network". In *2017 IEEE International Conference on Communications (ICC)*, pages 1–4. IEEE, 2017.

[45] ZTE. "Green Technology Innovations". *white paper*, 2016.

[46] 3GPP. "Study on new radio access technology: Radio access architecture and interfaces (Release 14)" . TS 38.801 V14.0.0 , 3rd Generation Partnership Project (3GPP), 2017.

[47] Kan Zheng, Long Zhao, Jie Mei, Mischa Dohler, Wei Xiang, and Yuexing Peng. "10 Gb/s hetsnets with millimeter-wave communications: access and networking-challenges and protocols". *IEEE Communications Magazine*, 53(1):222–231, 2015.

[48] Shinobu Namba, Takashi Matsunaka, Takayuki Warabino, Shoji Kaneko, and Yoji Kishi. "Colony-RAN architecture for future cellular network". In *Future Network & Mobile Summit (FutureNetw), 2012*, pages 1–8. IEEE, 2012.

[49] S. Namba, T. Warabino, and S. Kaneko. "BBU-RRH switching schemes for centralized RAN". In *7th International Conference on Communications and Networking in China*, pages 762–766, Aug 2012. doi: 10.1109/ChinaCom.2012.6417586.

[50] Y. S. Chen, W. L. Chiang, and M. C. Shih. "A Dynamic BBU-RRH Mapping Scheme Using Borrow-and-Lend Approach in Cloud Radio Access Networks". *IEEE Systems Journal*, PP(99):1–12, 2017. ISSN 1932-8184. doi: 10.1109/JSYST.2017.2666539.

[51] K. Boulos, M. El Helou, and S. Lahoud. "RRH clustering in cloud radio access networks". In *2015 International Conference on Applied Research in Computer Science and Engineering (ICAR)*, pages 1–6, Oct 2015. doi: 10.1109/ARCSE.2015.7338135.

[52] Doug Brake. "5G and next generation wireless: Implications for policy and competition". *Information Technology & Innovation Foundation*, pages 1–22, 2016.

[53] Quoc-Tuan Vien, Tuan Anh Le, Balbir Barn, and Ca V Phan. "Optimising energy efficiency of non-orthogonal multiple access for wireless backhaul in heterogeneous cloud radio access network". *Iet Communications*, 10(18):2516–2524, 2016.

[54] Karthikeyan Sundaresan, Mustafa Y Arslan, Shailendra Singh, Sampath Rangarajan, and Srikanth V Krishnamurthy. "FluidNet: a flexible cloud-based radio access network for small cells". *IEEE/ACM Transactions on Networking*, 24(2):915–928, 2016.

[55] Kmar Thaalbi, Mohamed Taher Missaoui, and Nabil Tabbane. "Short Survey on Clustering Techniques for RRH in 5G networks". In *2018 Seventh International Conference on Communications and Networking (ComNet)*, pages 1–5. IEEE, 2018.

[56] Xi Chen, Na Li, Jing Wang, Chengwen Xing, Liang Sun, and Ming Lei. "A dynamic clustering algorithm design for C-RAN based on multi-objective optimization theory". In *2014 IEEE 79th Vehicular Technology Conference (VTC Spring)*, pages 1–5. IEEE, 2014.

[57] Karen Boulos, Melhem El Helou, Marc Ibrahim, Kinda Khawam, Hadi Sawaya, and Steven Martin. "Interference-aware clustering in cloud radio access networks". In *2017 IEEE 6th International Conference on Cloud Networking (CloudNet)*, pages 1–6. IEEE, 2017.

[58] Stefan Parkvall, Erik Dahlman, Anders Furuskar, and Mattias Frenne. "NR:The new 5G radio access technology". *IEEE Communications Standards Magazine*, 1 (4):24–30, 2017.

[59] A Benjebbour, K Kitao, et al. "3GPP Defined 5G requirements and evaluation conditions". *NTT DOCOMO Tech. J*, 3:13–23, 2018.

[60] 3GPP. "LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception (Release 12)". TS 36.101 V 12.11.0 , 3rd Generation Partnership Project (3GPP), May 2016.

[61] 3GPP. "LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception (Release 12)" . TS 36.101 V12.14.1 , 3rd Generation Partnership Project (3GPP), 2017.

[62] T Uchino et al. "Carrier Aggregation Enhancement and Dual Connectivity Promising Higher Throughput and Capacity". *NTT DOCOMO Technical Journal*, 17(2): 36–46, 2015.

[63] Xingqin Lin, Jingya Li, Robert Baldemair, Thomas Cheng, Stefan Parkvall, Daniel Larsson, Havish Koorapaty, Mattias Frenne, Sorour Falahati, Asbjörn Grövlen, et al. "5G new radio: Unveiling the essentials of the next generation wireless access technology". *arXiv preprint arXiv:1806.06898*, 2018.

[64] Toru Uchino Teruaki Toeda Anil umish, Wuri A Hapsari and Hedeaki Takahashi. "5G Radio access standardization trends". *NTT DOCOMO Tech. J*, 19:36–47, 2018.

[65] C. Jiang, H. Zhang, Y. Ren, Z. Han, K. C. Chen, and L. Hanzo. "Machine Learning Paradigms for Next-Generation Wireless Networks". *IEEE Wireless Communications*, 24(2):98–105, April 2017. ISSN 1536-1284. doi: 10.1109/MWC.2016.1500356WC.

[66] Mandavilli Srinivas and Lalit M Patnaik. "Genetic algorithms: A survey". *computer*, 27(6):17–26, 1994.

[67] Aimin Zhou, Bo-Yang Qu, Hui Li, Shi-Zheng Zhao, Ponnuthurai Nagaratnam Suganthan, and Qingfu Zhang. "Multiobjective evolutionary algorithms: A survey of the state of the art". *Swarm and Evolutionary Computation*, 1(1):32 – 49, 2011. ISSN 2210-6502. doi: http://dx.doi.org/10.1016/j.swevo.2011.03.001. URL http://www.sciencedirect.com/science/article/pii/S2210650211000058.

[68] Z. Zhang, K. Long, J. Wang, and F. Dressler. "On Swarm Intelligence Inspired Self-Organized Networking: Its Bionic Mechanisms, Designing Principles and Optimization Approaches". *IEEE Communications Surveys Tutorials*, 16(1):513–537, January 2014.

[69] James Kennedy. "Particle Swarm Optimization". In *Encyclopedia of machine learning*, pages 760–766. Springer, 2011.

[70] A. Vasilakos, M. P. Saltouros, A. F. Atlassis, and W. Pedrycz. "Optimizing QoS routing in hierarchical ATM networks using computational intelligence techniques". *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 33(3):297–312, Aug 2003. ISSN 1094-6977. doi: 10.1109/TSMCC.2003.817354.

[71] NhuQuan Phan, ThiOanh Bui, Huilin Jiang, Pei Li, Zhiwen Pan, and Nan Liu. "Coverage optimization of LTE networks based on antenna tilt adjusting considering network load". *China Communications*, 14(5):48–58, May 2017.

[72] YA Adediran, H Lasisi, and OB Okedere. "Interference management techniques in cellular networks: A review". *Cogent Engineering*, 4(1):1294133, 2017.

[73] Ruchi Sachan, Tae Jong Choi, and Chang Wook Ahn. "A Genetic Algorithm with Location Intelligence Method for Energy Optimization in 5G Wireless Networks". *Discrete Dynamics in Nature and Society*, 2016, 2016.

[74] X. Huang, G. Xue, R. Yu, and S. Leng. "Joint Scheduling and Beamforming Coordination in Cloud Radio Access Networks With QoS Guarantees". *IEEE Transactions on Vehicular Technology*, 65(7):5449–5460, July 2016. ISSN 0018-9545. doi: 10.1109/TVT.2015.2464278.

[75] K. Wang, W. Zhou, and S. Mao. "On Joint BBU/RRH Resource Allocation in Heterogeneous Cloud-RANs". *IEEE Internet of Things Journal*, PP(99):1–1, 2017. ISSN 2327-4662. doi: 10.1109/JIOT.2017.2665550.

[76] Y. Shi, J. Zhang, and K. B. Letaief. "Group Sparse Beamforming for Green Cloud-RAN". *IEEE Transactions on Wireless Communications*, 13(5):2809–2823, May 2014. ISSN 1536-1276. doi: 10.1109/TWC.2014.040214.131770.

[77] Shinobu Namba, Takayuki Warabino, and Shoji Kaneko. "BBU-RRH switching schemes for centralized RAN". In *Communications and Networking in China (CHINACOM), 2012 7th International ICST Conference on*, pages 762–766. IEEE, 2012.

[78] Karthikeyan Sundaresan, Mustafa Y Arslan, Shailendra Singh, Sampath Rangarajan, and Srikanth V Krishnamurthy. FluidNet: a flexible cloud-based radio access network for small cells. *IEEE/ACM Transactions on Networking*, 24(2):915–928, 2016.

[79] Y. S. Chen, W. L. Chiang, and M. C. Shih. "A Dynamic BBU-RRH Mapping Scheme Using Borrow-and-Lend Approach in Cloud Radio Access Networks". *IEEE Systems Journal*, PP(99):1–12, 2017. ISSN 1932-8184. doi: 10.1109/JSYST.2017.2666539.

[80] C. Ran, S. Wang, and C. Wang. "Optimal load balancing in cloud radio access networks". In *2015 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1006–1011, March 2015. doi: 10.1109/WCNC.2015.7127607.

[81] M. Awais, A. Ahmed, M. Naeem, M. Iqbal, W. Ejaz, A. Anpalagan, and H. S. Kim. "Efficient Joint User Association and Resource Allocation for Cloud Radio Access Networks". *IEEE Access*, 5:1439–1448, 2017. ISSN 2169-3536. doi: 10.1109/ACCESS.2017.2663758.

[82] R. S. Alhumaima M. Khan and H. S. Al-Raweshidy. "QoS-Aware Dynamic RRH Allocation in a Self-Optimized Cloud Radio Access Network With RRH Proximity Constraint". *IEEE Transactions on Network and Service Management*, 14(3): 730–744, 2017.

[83] Wencheng He, Jinjin Gong, Xin Su, Jie Zeng, Xibin Xu, and Limin Xiao. "SDN-Enabled C-RAN? An Intelligent Radio Access Network Architecture". In *New Advances in Information Systems and Technologies*, pages 311–316. Springer, 2016.

[84] Massimo Condoluci, Toktam Mahmoodi, and Giuseppe Araniti. "Software-Defined networking and network Function Virtualization for C-RAN Systems". *5G Radio Access Networks: Centralized RAN, Cloud-RAN and Virtualization of Small Cells*, page 117, 2017.

[85] S. Gu, Z. Li, C. Wu, and H. Zhang. "Virtualized Resource Sharing in Cloud Radio Access Networks Through Truthful Mechanisms". *IEEE Transactions on Communications*, 65(3):1105–1118, March 2017. ISSN 0090-6778. doi: 10.1109/TCOMM.2016.2637900.

[86] H. Dahrouj, A. Douik, O. Dhifallah, T. Y. Al-Naffouri, and M. S. Alouini. "Resource allocation in heterogeneous cloud radio access networks: advances and challenges". *IEEE Wireless Communications*, 22(3):66–73, June 2015. ISSN 1536-1284. doi: 10.1109/MWC.2015.7143328.

[87] Tony QS Quek, Mugen Peng, Osvaldo Simeone, and Wei Yu. *"Cloud radio access networks: Principles, technologies, and applications"*. Cambridge University Press, 2017.

[88] Sascha Berger, Albrecht Fehske, Paolo Zanier, Ingo Viering, and Gerhard Fettweis. "Comparing online and offline SON solutions for concurrent capacity and coverage optimization". In *Vehicular Technology Conference (VTC Fall), 2014 IEEE 80th*, pages 1–6. IEEE, 2014.

[89] Yi-Wei Ma, Jiann-Liang Chen, and Chia-Ju Lin. "Automated Network Load Balancing and Capacity Enhancing Mechanism in Future Network". *IEEE Access*, 6:19407–19418, 2018.

[90] J. M. Ruiz-Avilés, M. Toril, S. Luna-Ramírez, V. Buenestado, and M. A. Regueira. "Analysis of Limitations of Mobility Load Balancing in a Live LTE System". *IEEE Wireless Communications Letters*, 4(4):417–420, Aug 2015. ISSN 2162-2337. doi: 10.1109/LWC.2015.2430345.

[91] Nikolay Dandanov, Hussein Al-Shatri, Anja Klein, and Vladimir Poulkov. "Dynamic Self-Optimization of the Antenna Tilt for Best Trade-off Between Coverage and Capacity in Mobile Networks". *Wireless Personal Communications*, 92(1): 251–278, 2017.

[92] John M Graybeal and Kamakshi Sridhar. "The evolution of SON to extended SON". *Bell Labs Technical Journal*, 15(3):5–18, 2010.

[93] Osianoh Glenn Aliu, Ali Imran, Muhammad Ali Imran, and Barry Evans. "A survey of self organisation in future cellular networks". *IEEE Communications Surveys & Tutorials*, 15(1):336–361, 2013.

[94] Oluwatayo Y Kolawole, Satyanarayana Vuppala, and Tharmalingam Ratnarajah. "Multiuser Millimeter Wave Cloud Radio Access Networks with Hybrid Precoding". *IEEE Systems Journal*, (99):1–12, 2017.

[95] Xianfu Chen, Pei Liu, Hang Liu, Celimuge Wu, and Yusheng Ji. "Multipath transmission scheduling in millimeter wave cloud radio access networks". In *2018 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2018.

[96] Katla Satyanarayana, Mohammed El-Hajjar, Ping-Heng Kuo, Alain AM Mourad, and Lajos Hanzo. "Adaptive Transceiver Design for C-RAN in mmWave Communications". *IEEE Access*, 6:16770–16782, 2018.

[97] H He, Jiang Xue, Tharmalingam Ratnarajah, and Mathini Sellathurai. "Performance analysis of millimeter wave cloud radio access networks". In *2016 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6. IEEE, 2016.

[98] Mustafa Y Arslan, Karthikeyan Sundaresan, and Sampath Rangarajan. Software-defined networking in cellular radio access networks: potential and challenges. *IEEE Communications Magazine*, 53(1):150–156, 2015.

[99] L. Liu and W. Yu. "Cross-Layer Design for Downlink Multihop Cloud Radio Access Networks With Network Coding". *IEEE Transactions on Signal Processing*, 65(7):1728–1740, April 2017. ISSN 1053-587X. doi: 10.1109/TSP.2017.2649486.

[100] C. PAN, H. Zhu, N. Gomes, and J. Wang. "Joint Precoding and RRH selection for User-centric Green MIMO C-RAN". *IEEE Transactions on Wireless Communications*, PP(99):1–1, 2017. ISSN 1536-1276. doi: 10.1109/TWC.2017.2671358.

[101] Z. Yu, K. Wang, H. Ji, X. Li, and H. Zhang. "Dynamic resource allocation in TDD-based heterogeneous cloud radio access networks". *China Communications*, 13(6):1–11, June 2016. ISSN 1673-5447. doi: 10.1109/CC.2016.7513198.

[102] K. Lin, W. Wang, Y. Zhang, and L. Peng. "Green Spectrum Assignment in Secure Cloud Radio Network with Cluster Formation". *IEEE Transactions on Sustainable Computing*, PP(99):1–1, 2017. ISSN 2377-3782. doi: 10.1109/TSUSC.2017.2678159.

[103] Cisco Visual Networking Index. "Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper". *link: http://goo. gl/ylTuVx*, 2017.

[104] Jeffrey G Andrews, Tianyang Bai, Mandar N Kulkarni, Ahmed Alkhateeb, Abhishek K Gupta, and Robert W Heath. "Modeling and analyzing millimeter wave cellular systems". *IEEE Transactions on Communications*, 65(1):403–430, 2017.

[105] Anil Umesh Wuri A Hapsari and Toru Uchino Teruaki Toeda Hideaki Takahashi. "5G Radio Access Network Standardization Trends". 2018.

[106] Hossein Shokri-Ghadikolaei, Carlo Fischione, Gabor Fodor, Petar Popovski, and Michele Zorzi. "Millimeter wave cellular networks: A MAC layer perspective". *IEEE Transactions on Communications*, 63(10):3437–3458, 2015.

[107] Kei Sakaguchi, Gia Khanh Tran, Hidekazu Shimodaira, Shinobu Nanba, Toshiaki Sakurai, Koji Takinami, Isabelle Siaud, Emilio Calvanese Strinati, Antonio Capone, Ingolf Karls, et al. "Millimeter-wave evolution for 5G cellular networks". *IEICE Transactions on Communications*, 98(3):388–402, 2015.

[108] Hisham Elshaer, Mandar N Kulkarni, Federico Boccardi, Jeffrey G Andrews, and Mischa Dohler. "Downlink and uplink cell association with traditional macrocells and millimeter wave small cells". *IEEE Transactions on Wireless Communications*, 15(9):6244–6258, 2016.

[109] Zhouyue Pi and Farooq Khan. "An introduction to millimeter-wave mobile broadband systems". *IEEE communications magazine*, 49(6), 2011.

[110] Sherif Adeshina Busari, Kazi Mohammed Saidul Huq, Shahid Mumtaz, Linglong Dai, and Jonathan Rodriguez. "Millimeter-Wave Massive MIMO Communication for Future Wireless Systems: A Survey". *IEEE Communications Surveys & Tutorials*, 2017.

[111] Cristian Tatino, Ilaria Malanchini, Nikolaos Pappas, and Di Yuan. "Maximum throughput scheduling for multi-connectivity in millimeter-wave networks". In *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2018 16th International Symposium on*, pages 1–6. IEEE, 2018.

[112] Hani Mehrpouyan, Michail Matthaiou, Rui Wang, George K Karagiannidis, and Yingbo Hua. "Hybrid millimeter-wave systems: A novel paradigm for HetNets". *IEEE Communications Magazine*, 53(1):216–221, 2015.

[113] Sarabjot Singh, Mandar N Kulkarni, Amitava Ghosh, and Jeffrey G Andrews. "Tractable model for rate in self-backhauled millimeter wave cellular networks". *IEEE Journal on Selected Areas in Communications*, 33(10):2196–2211, 2015.

[114] Marco Di Renzo. "Stochastic geometry modeling and analysis of multi-tier millimeter wave cellular networks". *IEEE Transactions on Wireless Communications*, 14(9):5038–5057, 2015.

[115] Diana Maamari, Natasha Devroye, and Daniela Tuninetti. "Coverage in mmWave cellular networks with base station co-operation". *IEEE Transactions on Wireless Communications*, 15(4):2981–2994, 2016.

[116] Muhammad Shahmeer Omar, Muhammad Ali Anjum, Syed Ali Hassan, Haris Pervaiz, and Qiang Niv. "Performance analysis of hybrid 5G cellular networks exploiting mmWave capabilities in suburban areas". In *Communications (ICC), 2016 IEEE International Conference on*, pages 1–6. IEEE, 2016.

[117] Tadilo Endeshaw Bogale and Long Bao Le. "Massive MIMO and mmWave for 5G wireless HetNet: Potential benefits and challenges". *IEEE Vehicular Technology Magazine*, 11(1):64–75, 2016.

[118] Omid Semiari, Walid Saad, and Mehdi Bennis. "Joint Millimeter Wave and Microwave Resources Allocation in Cellular Networks with Dual-Mode Base Stations". *IEEE Transactions on Wireless Communications*, 2017.

[119] Omid Semiari, Walid Saad, Mehdi Bennis, and Behrouz Maham. "Caching Meets Millimeter Wave Communications for Enhanced Mobility Management in 5G Networks". *IEEE Transactions on Wireless Communications*, 2017.

[120] Solmaz Niknam, Ali Arshad Nasir, Hani Mehrpouyan, and Balasubramaniam Natarajan. "A Multiband OFDMA Heterogeneous Network for Millimeter Wave 5G Wireless Applications". *IEEE Access*, 4:5640–5648, 2016.

[121] Simin Xu, Nan Yang, and Shihao Yan. "Optimal Load Balancing in Millimeter Wave Cellular Heterogeneous Networks". *arXiv preprint arXiv:1711.05917*, 2017.

[122] Esma Turgut and M Cenk Gursoy. "Coverage in heterogeneous downlink millimeter wave cellular networks". *IEEE Transactions on Communications*, 65(10): 4463–4477, 2017.

[123] Fazal Muhammad, Ziaul Haq Abbas, and Frank Y Li. "Cell association with load balancing in nonuniform heterogeneous cellular networks: Coverage probability and rate analysis". *IEEE Transactions on Vehicular Technology*, 66(6):5241–5255, 2017.

[124] Zainab H Fakhri, Muhammad Khan, Firas Sabir, and Hamed S Al-Raweshidy. "A resource allocation mechanism for cloud radio access network based on cell differentiation and integration concept". *IEEE Transactions on Network Science and Engineering*, 5(4):261–275, 2017.

[125] M Khan, Zainab H Fakhri, and Hamed S Al-Raweshidy. "Semistatic cell differentiation and integration with dynamic BBU-RRH mapping in cloud radio access network". *IEEE Transactions on Network and Service Management*, 15(1): 289–303, 2017.

[126] M. Peng, Y. Sun, X. Li, Z. Mao, and C. Wang. "Recent Advances in Cloud Radio Access Networks: System Architectures, Key Techniques, and Open Issues". *IEEE Communications Surveys Tutorials*, 18(3):2282–2308, thirdquarter 2016. ISSN 1553-877X. doi: 10.1109/COMST.2016.2548658.

[127] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann. "Cloud RAN for Mobile Networks;A Technology Overview". *IEEE Communications Surveys Tutorials*, 17(1):405–426, Firstquarter 2015. ISSN 1553-877X. doi: 10.1109/COMST.2014.2355255.

[128] M Khan, RS Alhumaima, and HS Al-Raweshidy. "Reducing energy consumption by dynamic resource allocation in C-RAN". In *Networks and Communications (EuCNC), 2015 European Conference on*, pages 169–174. IEEE, 2015.

[129] Yegui Cai, F Richard Yu, and Shengrong Bu. "A decision theoretic approach for clustering and rate allocation in coordinated multi-point (CoMP) networks with delayed channel state information". *Transactions on Emerging Telecommunications Technologies*, 28(1), 2017.

[130] J. Zhang, R. Zhang, G. Li, and L. Hanzo. "Distributed Antenna Systems in Fractional-Frequency-Reuse-Aided Cellular Networks". *IEEE Transactions on Vehicular Technology*, 62(3):1340–1349, March 2013. ISSN 0018-9545. doi: 10.1109/TVT.2012.2230282.

[131] Marvin K Simon and Mohamed-Slim Alouini. *"Digital communication over fading channels"*, volume 95. John Wiley & Sons, 2005.

[132] Christos Bouras, Georgios Diles, Vasileios Kokkinos, Konstantinos Kontodimas, and Andreas Papazois. "A simulation framework for evaluating interference mitigation techniques in heterogeneous cellular environments". *Wireless personal communications*, 77(2):1213–1237, 2014.

[133] H. Zhang, C. Jiang, J. Cheng, and V. C. M. Leung. "Cooperative interference mitigation and handover management for heterogeneous cloud small cell networks". *IEEE Wireless Communications*, 22(3):92–99, June 2015. ISSN 1536-1284. doi: 10.1109/MWC.2015.7143331.

[134] Claude Desset, Björn Debaillie, Vito Giannini, Albrecht Fehske, Gunther Auer, Hauke Holtkamp, Wieslawa Wajda, Dario Sabella, Fred Richter, Manuel J Gonzalez, et al. "Flexible power modeling of LTE base stations". In *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*, pages 2858–2862. IEEE, 2012.

[135] Hauke Andreas Holtkamp. *"Enhancing the energy efficiency of radio base stations"*. PhD thesis, University of Edinburgh, 2014.

[136] Rodney S Tucker and Kerry Hinton. "Energy consumption and energy density in optical and electronic signal processing". *IEEE Photonics Journal*, 3(5):821–833, 2011.

[137] Ka-Lun Lee, Behnam Sedighi, Rodney S Tucker, Hungkei Keith Chow, and Peter Vetter. "Energy efficiency of optical transceivers in fiber access networks [invited]". *Journal of Optical Communications and Networking*, 4(9):A59–A68, 2012.

[138] Angelos Chatzipapas, Sara Alouf, and Vincenzo Mancuso. "On the minimization of power consumption in base stations using on/off power amplifiers". In *Online conference on green communications (GreenCom), 2011 IEEE*, pages 18–23. IEEE, 2011.

[139] Gunther Auer, Vito Giannini, Claude Desset, Istvan Godor, Per Skillermark, Magnus Olsson, Muhammad Ali Imran, Dario Sabella, Manuel J Gonzalez, Oliver Blume, et al. "How much energy is needed to run a wireless network?" *IEEE Wireless Communications*, 18(5), 2011.

[140] A Kaveh. "Particle swarm optimization". In *Advances in Metaheuristic Algorithms for Optimal Design of Structures*, pages 11–43. Springer, 2017.

[141] Jakob Vesterstrom and Rene Thomsen. "A comparative study of differential evolution, particle swarm optimization, and evolutionary algorithms on numerical benchmark problems". In *Evolutionary Computation, 2004. CEC2004. Congress on*, volume 2, pages 1980–1987. IEEE, 2004.

[142] 3GPP. "Technical Specification Group Services and System Aspects;Services and service capabilities (Release 13) ". TS V13.0.0, 3rd Generation Partnership Project (3GPP), December 2015.

[143] Cisco Visual Networking Index. "Global mobile data traffic forecast update, 2016-2021", Feb 2017.

[144] W. Hong, K. Baek, and S. Ko. "Millimeter-Wave 5G Antennas for Smartphones: Overview and Experimental Demonstration". *IEEE Transactions on Antennas and Propagation*, 65(12):6250–6261, Dec 2017. ISSN 0018-926X. doi: 10.1109/TAP.2017.2740963.

[145] M. Khan, Z. H. Fakhri, and H. S. Al-Raweshidy. "Semistatic Cell Differentiation and Integration With Dynamic BBU-RRH Mapping in Cloud Radio Access Network". *IEEE Transactions on Network and Service Management*, 15(1):289–303, March 2018. ISSN 1932-4537. doi: 10.1109/TNSM.2017.2771622.

[146] J. Xu, J. Yao, L. Wang, K. Wu, L. Chen, and W. Lou. "Revolution of Self-Organizing Network for 5G MmWave Small Cell Management: From Reactive to Proactive". *IEEE Wireless Communications*, 25(4):66–73, AUGUST 2018. ISSN 1536-1284. doi: 10.1109/MWC.2018.1700420.

[147] Seppo Hämäläinen, Henning Sanneck, and Cinzia Sartori. *"LTE self-organising networks (SON): network management automation for operational efficiency"*. John Wiley & Sons, 2012.

[148] Ljupco Jorguseski, Adrian Pais, Fredrik Gunnarsson, Angelo Centonza, and Colin Willcock. "Self-organizing networks in 3GPP: standardization and future trends". *IEEE Communications Magazine*, 52(12):28–34, 2014.

[149] Yaniv Azar, George N Wong, Kevin Wang, Rimma Mayzus, Jocelyn K Schulz, Hang Zhao, Felix Gutierrez, DuckDong Hwang, and Theodore S Rappaport. "28 GHz propagation measurements for outdoor cellular communications using steerable beam antennas in New York City". In *Communications (ICC), 2013 IEEE International Conference on*, pages 5143–5147. IEEE, 2013.

[150] Tianyang Bai, Rahul Vaze, and Robert W Heath. "Analysis of blockage effects on urban cellular networks". *IEEE Transactions on Wireless Communications*, 13(9): 5070–5083, 2014.

[151] Tianyang Bai and Robert W Heath. "Coverage and rate analysis for millimeter-wave cellular networks". *IEEE Transactions on Wireless Communications*, 14(2): 1100–1114, 2015.

[152] Byeong Seok Ahn. "Compatible weighting method with rank order centroid: Maximum entropy ordered weighted averaging approach". *European Journal of Operational Research*, 212(3):552–559, 2011.

[153] Alessandro Passaro and Antonina Starita. "Particle swarm optimization for multimodal functions: a clustering approach". *Journal of Artificial Evolution and Applications*, 2008, 2008.

[154] David B Fogel. *"Evolutionary computation: toward a new philosophy of machine intelligence"*, volume 1. John Wiley & Sons, 2006.

[155] Denny Hermawanto. "Genetic algorithm for solving simple mathematical equality problem". *arXiv preprint arXiv:1308.4675*, 2013.

[156] Mohamad Omar Al Kalaa, Ali Imran, and Hazem H Refai. "mmWave based vs 2 GHz networks: What is more energy efficient?" In *2016 International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 67–71. IEEE, 2016.

[157] Zainab H Fakhri, Firas Sabir, and HS Al-Raweshidy. "An Interference Mitigation Scheme for Millimetre Wave Heterogeneous Cloud Radio Access Network with Dynamic RRH Clustering". In *2019 International Symposium on Networks, Computers and Communications (ISNCC)*, pages 1–8. IEEE, 2019.

[158] Evolved Universal Terrestrial Radio Access. "mobility enhancements in heterogeneous networks (release 11), 3rd Generation Partnership Project (3GPP), Sophia Antipolis, France". Technical report, TR 36.839.

[159] Thomas David Novlan, Radha Krishna Ganti, Arunabha Ghosh, and Jeffrey G Andrews. "Analytical evaluation of fractional frequency reuse for OFDMA cellular networks". *IEEE Transactions on wireless communications*, 10(12):4294–4305, 2011.

[160] Karen Boulos, Melhem El Helou, and Samer Lahoud. "RRH clustering in cloud radio access networks". In *Applied Research in Computer Science and Engineering (ICAR), 2015 International Conference on*, pages 1–6. IEEE, 2015.