OPTIMISATION OF RESOURCE ALLOCATION IN HIGH USER-DENSITY WIRELESS NETWORKS

A thesis submitted to the University of Manchester
for the degree of Doctor of Philosophy
in the Faculty of Science and Engineering

2018

By
Hanifa Nabuuma
School of Electrical and Electronic Engineering
Faculty of Science and Engineering

Contents

List of Figures

List of Abbreviations

3G	Third Generation
3GPP	Third Generation Partnership Project
4G	Fourth Generation
ABS	Almost Blank Subframes
ACK	Acknowledgement
AID	Association Identifier
AMC	Adaptive Modulation and Coding
AP	Access Point
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
BS	Base Station
BSS	Basic Service Set
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CF	Contention Free
CP	Cyclic Prefix

CSMA Carrier Sense Multiple Access

CQI Channel Quality Indicator

CRC Cyclic Redundancy Check

DAS Distributed Antenna System

DCF Distributed Coordination Function

DCG Dynamic Coloured Graph

DFT Discrete Fourier Transform

DIFS Distributed Coordination Function Inter-Frame Space

DS Distribution System

ECR Energy Consumption Ratio

e-ICIC Enhanced Inter-Cell Interference Coordination

EIRP Effective Isotropic Radiated Power

ESS Extended Service Set

FBS Femtocell Base Station

FCS Frame Check Sequence

FDD Frequency Division Duplexing

FDMA Frequency Division Multiple Access

FFR Fractional Frequency Reuse

FFT Fast Fourier Transform

FMS Femtocell Management System

FUE Femtocell User Equipment

GB-DCF Group Based Distribution Coordination Function

H2H Human-to-human

HARQ Hybrid Automatic Repeat Request

HOL Head of Line

ICI Inter-Cell Interference

ICIC Inter-Cell Interference Coordination

IDFT Inverse Discrete Fourier Transform

IE Information Element

IFFT Inverse Fast Fourier Transform

IM Interference Management

IoT Internet of Things

ISI Intersymbol interference

LOS Line of Sight

LTE Long Term Evolution

M2H Machine-to-Human

M2M Machine-to-Machine

MAC Medium Access Control

MBS Macrocell Base Station

MCS Modulation and Coding Scheme

MIMO Multiple Input Multiple Output

MoC Matrix of Conflict

MQAM M-ary Quadrature Amplitude Modulation

MR Measurement Report

MRC Maximal Ratio Combining

MTC Machine Type Communications

MUE Macrocell User Equipment

MPDU MAC Protocol Data Unit

NDP Null Data Packet

OBSS Orthogonal Basic Service Set

OFDM Orthogonal Frequency Division Multiplexing

OFDMA Orthogonal Frequency Division Multiple Access

OMS Operator Management System

PAPR Peak to Average Peak Ratio

PCF Point Coordination Function

PCI Physical Cell Identity

PDF Probability Distribution Function

PER Packet Error Rate

PIFS Point Coordination Function Inter-Frame Space

PLCP Physical Layer Converegence Procedure

PPDU PLCP Protocol Data Unit

PRB Physical Resource Block

PSDU PLCP Service Data Unit

QoS	Quality of Service
RAW	Restricted Access Window
RB	Resource Block
RF	Radio Frequency
RID	Response Indication Deferral
RMS	Root Mean Square
RRC	Radio Resource Control
RRM	Radio Resource Management
RS	Reference Signal
RSRP	Reference Signal Received Power
RSRQ	Reference Signal Received Quality
RSSI	Received Signal Strength Indicator
SBS	Small cell Base Station
SC-OFDM	Single Carrier Orthogonal Frequency Division Multiplexing
SFR	Soft Frequency Reuse
SIFS	Short Inter-Frame Space
SINR	Signal to Interference plus Noise Ratio
SNR	Signal to Noise Ratio
SON	Self Organising Network
STA	Station
STBC	Space Time Block Codes

SUE Small cell base station UE

SUI Stanford University Interim

T-R Transmitter-Receiver

TDD Time Division Duplexing

TDM Time Division Multiplexing

TDMA Time Division Multiple Access

TTI Transmission Time Interval

UE User Equipment

WLAN Wireless Local Area Network

List of Mathematical Notations

$\lceil \cdot \rceil$	Largest integer greater than $(\cdot)$
$\lfloor \cdot \rfloor$	Largest integer less than $(\cdot)$
$Q(\cdot)$	Q function
$\log_2(\cdot)$	Base-2 logarithm
$\log_{10}(\cdot)$	Base-10 logarithm
$\exp(\cdot)$	Exponential function
$(\cdot)_{*i}$	Elements in column i of matrix $(\cdot)$
$(\cdot)_{i*}$	Elements in row i of matrix $(\cdot)$
$\mathbb{C}$	Matrix with complex numbers
$\mathbb{E}(\cdot)$	Expectation of $(\cdot)$
$\mathbb{R}$	Matrix with real numbers
$*$	Convolution
$\emptyset$	Empty set
$\wedge$	AND operator
$\vee$	OR operator
$\lvert \cdot \rvert$	The number of elements in set or vector $(\cdot)$

$(\cdot)^{N\times M}$	An $N \times M$ matrix
$(\cdot)_{n,m}$	Element in row n and column m of matrix $(\cdot)$
$R(\cdot)$	Rows in $(\cdot)$
$\odot$	Componentwise multiplication of vectors

List of Variables

α	Pathloss exponent
λ	Wavelength
ζ	Interference/Detection map
ζ^{MG}	Merged detection map
θ	Angle of Arrival
θ_l	Phase shift for *lth* multipath component
$\bar{\tau}$	Mean excess delay
$\mathfrak{b}$	Number of bits derived from SINR
τ_l	Excess delay spread of *lth* multipath component
τ_{max}	Maximum delay spread
τ_{RMS}	Root mean square delay spread
$\triangle f$	Subcarrier spacing
ϕ_s	Duration of successful data packet transmission
ϕ_f	Duration of failed data packet transmission
σ	Standard deviation of shadowing
ϵ	Polyphase code index

$\mathfrak{I}$	SIR difference matrix
Υ	SINR
Υ_{min}	Minimum SINR for given γ_{th}
Υ_l^L	The lowest SINR for MQAM level L
ℓ	Enterprise building length
γ	SIR
γ_{th}	SIR threshold
Λ	Potential allocations matrix
$\mathfrak{b}$	The number of bits derived from the SINR
ρb	Bit error rate
ρp	Packet error rate
ρp_a	Packet error rate for ACK frame
ρp_b	Packet error rate for beacon frame
ρp_d	Packet error rate for data frame
ρs	Symbol error rate
Φ_i	Duration, in time slots, of the *ith* transmission in a group slot
φ	Standard deviation of shadowing
μ_{pa}	Power amplifier efficiency
μ_{ps}	Power supply efficiency
ε	Normalised carrier frequency offset
$\wp$	Propagation delay

κ	Number of sectors
$\triangle f$	Subcarrier spacing
$\mathcal{A}_s$	Allocation Order
$\mathcal{B}_{sig}$	System bandwidth
$\mathcal{B}_c$	Coherence bandwidth
$\mathcal{B}_{rb}$	Resource block bandwidth
$\mathcal{C}$	Cluster
$\mathcal{D}_w$	Ward's inter-cluster distance metric
$\mathcal{F}$	Set of all femtocells
$\mathcal{F}_{in}$	Set of femtocells with inner UEs
$\mathcal{F}_{max}$	Set of all femtocells with the maximum number of detected UEs
$\mathcal{F}_{pc}$	Set of all femtocells reducing power
$\mathcal{G}_i$	Group i
$\mathcal{H}_m$	MAC header size
$\mathcal{H}_p$	PLCP header size
$\mathcal{K}_c$	UEs interfered by power reducing BS
$\mathcal{K}_s$	UEs served by power reducing BS
$\mathcal{L}$	Payload size
$\mathcal{M}_1$	Maximum number of tranmissions in a group slot
$\mathcal{M}_2$	Minimum number of tranmissions in a group slot
$\mathcal{M}_s$	Maximum number of successful tranmissions in a group slot

$\mathcal{N}$	Zero mean Gaussian random variable with unit variance
$\mathcal{R}$	User data rate
$\overline{\mathcal{R}}$	Discrete capacity
$\mathcal{R}_{min}$	Minimum datat rate
$\mathcal{R}_{tot}$	System sum rate
$\mathcal{S}_{ie}$	Sectorisation information element
$\mathcal{Q}_b$	Beacon frame size in bytes
$\mathcal{Q}_d$	Data packet size in bytes
$\mathbb{G}$	Set groups with size less than g_t
$\mathbb{P}$	Transmit power of all femtocells
$\mathbb{S}$	Array with number of RBs allocated to each UE
$\mathbb{S}_t$	Total RB utilisation
$\mathcal{U}$	Number of detected UEs by all SBSs
$\overline{\mathcal{U}}$	Number of detected UEs by all SBSs after power reduction
$\mathcal{X}_j$	UEs served by femtocell j
$\tilde{a}_k$	Modulation symbol at subcarrier k
a_l	Amplitude of the *lth* multipath component
b_{in}	Total information required for the algorithm
$b_{\mathfrak{I}}$	Bits required to transmit an element in the SIR difference matrix
$b_{\triangle p}$	Bits required to transmit power reduction instruction to a base station
b_{ζ}	Bits required to transmit an element in the matrix of conflict

c	Speed of light
d	Transmitter receiver separation distance
d_0	Reference distance in pathloss calculations
d_{max}	Maximum distance between UE and serving femtocell
e_i	*ith* filter coefficient
f	Carrier frequency
f_d	Doppler shift
g_t	Target group size
h	Channel gain
$h(t)$	Channel impulse response
h_{eq}	Equaliser impulse response
n	zero mean Gaussian noise
$n_{d,i}$	Number of STAs that did not received the beacon
$n_{o,i}$	Number of STAs that received the beacon
t_a	Duration of an ACK frame
t_b	Duration of an beacon frame
t_d	Duration of a data frame
t_{difs}	Duration of a DIFS
t_{sifs}	Duration of a SIFS
u_j	Detected UEs by BS j
u_j^{in}	Number of inner UEs served by BS j

v	Velocity
x	Transmitted signal
y	Received signal
$A(\triangle\tau, \triangle t)$	Channel auto correlation function
A	Peak amplitude of LOS component
$\widetilde{A}_{j,r}$	Allocation indicator to show if base station j allocated a UE at RB r
$\overline{A}$	Allocation matrix
$\overline{F}$	Forbidden matrix
$\breve{A}$	Unique allocations in $\overline{A}$
B	Random variable denoting number of backoff slots in a group slot
CW_{min}	Minimum contention window size
CW_{max}	Maximum contention window size
D_r	Distance ratio
E	Symbol equivocation
E_b	Bit energy
$E_{S\|G}$	Expectation of successful transmissions given group size
H	Channel frequency response
I	Indicator function
$\overline{N}$	The number of unique columns in $\overline{A}$
N_0	Noise density
N_{FFT}	FFT size

N_g Number of groups

N_{it} Number of required iterations

N_p Number of resource blocks in partition

N_{rb} Number of resource blocks in an OFDM symbol

N_{sub} Number of subcarriers

N_T Number of resource blocks in a frame

N_u Number of users in the network

S Random variable denoting number of successful transmissions in a group slot

T Duration of data symbol

T_0 Duration of OFDM symbol

T_c Coherence Time

T_{cp} Cyclic prefix length

T_F LTE frame period

Th Throughput

T_s Group slot duration

$T_{b,i}$ Duration, in time slots, of the backoff for the *ith* transmission

$T_{t,m}$ Duration, in time slots, of m transmissions in a group slot

$\triangle p_i$ Maximum possible power reduction by BS i to meet γ_{th}

$\triangle \overline{p}^i_j$ Required power reduction by BS j to share RB with BS i

P_{DL} Power consumption for downlink transmission

P_{RF} Radio frequency component of power consumption

P_{SP} Signal processing component of power consumption

P_{th} Received power threshold

P_t Transmit power

P_{max} Maximum transmit power

P_{sen} Minimum receiver sensitivity

PL Path Loss

$\mathbf{V}_m^g$ Viable patterns for group size g and m transmissions in group slot

X_σ Shadowing with standard deviation of σ

Y Received signal envelope

Abstract

Combined with technological advancements, resource allocation is a key tool used to improve the performance of wireless networks. Today's heterogeneous networks are mainly interference limited making resource allocation an important tool in interference management. Of particular interest in this thesis is the use of resource allocation for interference avoidance.

In this thesis, a power dimension is added to established radio resource management (RRM) techniques to improve the throughput attained in small cell networks. Two power allocation techniques are proposed to this end. The first technique uses power allocation to modify the interference map generated in order to increase radio resource reuse and throughput. The second technique increases reuse after initial allocation where all base stations transmit the same power to their users. In the second allocation, some users are selectively granted access to specific resources if they can reduce their power to meet the interference constraints of incumbent users, allocated in the initial allocation. Simulation results for both power allocation techniques show that they attain increased throughput while maintaining the minimum data rate performance of the network.

In most 802.11 networks, resource allocation is distributed and tends to be limited to channel selection and in some cases power allocation. However in 802.11ah networks, which are designed for wide coverage and thousands of sensor type devices, there is some centralisation in the form of restriction of which stations (STAs) are permitted to contend for access to the wireless medium at a given time. The contention works fairly well in low traffic scenarios without hidden nodes. However, in saturated scenarios with hidden nodes, its performance is poor. The research presented in this thesis proposes a grouping algorithm to solve the hidden node problem. Furthermore it proposes that backoff timers set during the contention process are set based on the unique

association identifiers (AID) of the group members. This ensures that collisions due to devices choosing the same backoff timer value are eliminated. An analytical model is developed and is verified through simulations. Further simulations show an improvement in throughput, delay and energy efficiency when AID-based backoff timers are used compared to the standard random backoff timers in saturated scenarios.

In a network using AID-based backoff and the grouping technique that manages hidden nodes, this research proposes a solution to the outage problem in 802.11ah networks which involves the use of connected stations (STAs) to relay packets for neighbouring STAs in outage. The results obtained through simulations show that for dense networks, using connected STAs as relays can help solve the outage problem, in the absence of relays.

Declaration

No portion of the work referred to in this dissertation has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://www.campus.manchester.ac.uk/medialibrary/policies/intellectual-property.pdf), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see http://www.manchester.ac.uk/library/

aboutus/regulations) and in The University's policy on presentation of Theses

Acknowledgements

First and foremost, I'd like to thank the Almighty God, without whom none of this would have been possible. I also thank Dr. Emad Alsusa for the support guidance he has given me from the start to the end of my PhD.

Special thanks to Wahyu, Aysha, Edwin, Abdul-Hameed, and Makram for your support. I also thank all my other colleagues in the MACS group for the camaraderie.

I also extend my gratitude to my siblings, with special thanks to Maama Sophie, for the support and encouragement throughout this PhD. Thanks to my friends, especially Grace and Stephen, for the encouraging words and support. Thanks to my family in U.K., including, the Nsubuga's, the Bbossa's and Zam, for the support.

Special thanks to Hajat Fatuma Lubega, it started with you. Thank you for the prayers and encouragement. Thanks to Hajji Nasser Lubega for the support.

Special thanks to Hajat Sarah Lutale, for the support and prayers from start to finish. Thanks to Dr. Sentongo for the encouraging words.

I thank my children, Zura, Thobait, Nyla, Shasmeen, and Sara. I know this has been tough but thank you hanging in there. I pray that one day you can read this work and understand why I spent so much time in Manchester.

Finally, I thank my beloved husband. Words cannot describe how grateful I am for the moral, financial, and emotional support that you have given me throughout this PhD. Thank you for standing in the gap. You're one in a billion and may Allah reward you abundantly.

Dedication

To the Mukasas: Faisal, Zurah, Thobait, Nyla, Shasmeen and Sara.

Chapter 1

Introduction

RECENT years have seen an explosive growth in the number of wireless devices and the applications of wireless devices [4, 5]. With a projection of over 10 billion mobile wireless devices by 2020, more than the world's population, there's a clear challenge in ensuring that all these devices get the connectivity they need on demand [6]. Moreover, the heterogeneity of devices along with their quality of service (QoS) requirements complicates matters even further [7]. As a result, there has been an exponential growth in user generated data traffic [8, 9] which may broadly be categorised as human-to-machine (H2M) and human-to-human (H2H) and machine-to-machine (M2M) communication traffic [10–12]. The most prevalent wireless networks today are wireless local area networks (WLANs) and cellular networks [13]. WLANs were initially built to provide broadband services to fixed or slow moving users in a limited geographical area while cellular networks were designed to provide voice and broadband services to both fixed and mobile users in a wide geographical area [13, 14]. Both WLANs and cellular networks have had to adapt to deal with the exponential growth in data traffic. In fact, one of the ways that cellular networks are dealing with the problem is by offloading part of their traffic to WLANs [15]. Other approaches to meeting the rising demand can broadly be described as technological advancements, which include: improvement of signal processing capabilities of devices [16], use of multiple input, multiple output (MIMO) technologies [17, 18] and use of high-order modulations [19]. These techniques are being used in both WLANs [20] and cellular networks to increase spectral efficiency [16, 19]. Other techniques are specific to the type of network because of the differences in the network architectures. For instance,

traditional cellular networks provide wide geographical coverage using licensed bands with centralised resource allocation while WLANs provide small area coverage using unlicensed bands and decentralised resource allocation [5, 10, 16, 21]. These key differences in the way these networks operate and how they are designed imply that some of the solutions to address the challenges they face differ.

In order to deal with the challenge of demand for higher data rates, the traditional approach for cellular networks is cell splitting. However, in already dense deployments, gains from cell splitting would be minimal because of inter-cell interference (ICI) [22]. Further, costs associated with cell splitting are also prohibitive in the long run [9, 22]. However, reducing the distance between the base station (BS) and the user equipment (UE) is still a logical solution because it reduces the pathloss between the UE and BS hence reducing the required transmission power and interference caused. This has led to the rise of the small cell base station (SBS). SBSs are low cost, low power and low coverage base stations (BSs) that are deployed to underlay the traditional macrocell network and boost capacity by reusing the same frequency while causing minimal interference to the macrocell users. SBSs provide an energy efficient means of boosting cell capacity because they use much less power than traditional BSs. The five major types of SBSs are microcells, picocells, femtocells, relays and distributed antenna systems (DAS) [8]. Microcells are regular base stations with inter-base station distance less than 500 m and are usually deployed in urban areas [23]. Picocells are regular base stations however they differ from macro base stations and microcells in the following ways: they have low transmit power, have omni-directional antennas and can also be deployed indoors [22]. Femtocells are indoor consumer deployed base stations that use the consumer's digital subscriber line for backhaul and also have omni-directional antennas [22]. Relay nodes are base stations without a wired backhaul i.e. they have a wireless backhaul [22, 23]. The wireless backhaul is either in-band (uses the same resources as the UE and base station) or out-of-band [22]. Relay nodes are either full duplex or half duplex and are used to provide coverage extension and throughput enhancement by transmitting an enhanced signal to/from the macro base station from/to the UEs [22, 23]. DAS involves spatially separating the antennas of a conventional BS and connecting them via a common processing unit [23]. This enables the reduction of transmit power by each antenna as it covers a smaller area [23]. Picocells, DAS and relay nodes may be deployed both indoors and outdoors. Their transmit power ranges from 250 mW to 2 W for outdoor deployments, while it is usually 100 mW or less for indoor deployments [8, 22]. Femtocells are only deployed indoors and their transmit

power is 100 mW or less [22]. A network with a mix of macro BSs that are under-laid with SBSs is referred to as a heterogeneous network (HetNet) [8, 9, 22]. In addition to boosting capacity, SBSs are also used to alleviate coverage dead zones for example femtocells are used to provide coverage indoors where the coverage for the macrocell may be poor and relays are used in outdoor areas with coverage holes [8]. For cellular network operators to achieve the projected gains in capacity from HetNets, a number of technical challenges need to be addressed first. One of the biggest challenges is interference management.

On the other hand, traditional WLANs may be categorised as small cell networks because they do not have macrocell base stations (MBSs). However, the rapid rise in machine type communications (MTC) including M2M communications, especially due to the rapid growth in sensor applications, has created a new problem for WLANs [4, 12, 24]. Many sensor applications utilise thousands of sensors spread over a wide geographical area. In order to meet the requirements of these applications, a new standard, 802.11ah has been introduced [10,21]. Among the key modifications made to the existing standards is the support for macro deployment in order to widen the coverage of the access point (AP) [25]. An access point is the equivalent of a base station in a cellular network. Other modifications include introduction of slotted access to reduce contention and a longer association identifier (AID) in order to enable association of over 6000 devices to one AP [21, 26]. Despite the modifications, 802.11ah networks suffer from collisions that are inherent to WLANs and exacerbated by the hidden node problem caused by the wider separation between contending devices [25, 27].

1.1 Motivations

As stated in the previous section, interference management is needed in order to enjoy the benefits of heterogeneous networks. Of particular interest is interference avoidance, where restrictions are put on the resources used by different BSs [28]. The restrictions may be in form of the time-frequency resources available to a BS or restrictions on the transmit power used by a BS on particular time-frequency resources [28]. In some cases it may involve both types of restrictions [29, 30]. An interference avoidance technique was introduced in [31] which used centralised allocation to enforce time-frequency restrictions to meet a required signal to interference (SIR) threshold.

Another similar technique was introduced in [32] however with distributed allocation. Both techniques assume equal pilot power allocation by all SBSs which through measurement reports sent by UEs to their serving BSs influenced the interference map generated by the SBSs to guide the radio resource allocation. In this thesis, we investigate the impact of pilot power allocation on the throughput of SBSs using these interference avoidance techniques. The pilot power of base station determines the coverage of the base station. A number of papers address the issue of determining the appropriate pilot power for a femtocell in a heterogeneous network. Most of the prior work on pilot power, focuses on either load balancing [33, 34] or maintaining a specified coverage radius such that an STA at that radius will receive at least the same power from the macrocell as it receives from the SBSs [35, 36]. However, in this thesis, a technique that modifies the pilot power (and consequently the transmit power on the resources) in order to modify the interference map to reduce restrictions and improve network throughput is proposed. The key constraint is that the minimum data rate of the network should not be reduced at the cost of improved throughput.

A closer look at the interference avoidance technique in [31] shows that when the SBSs transmit the equal pilot power, there are opportunites to increase resource reuse and network throughput if some of the SBSs reduce the power allocated on particular resources in order to meet the required SIR threshold on the resources where they were previously barred. In this thesis two algorithms are proposed to apply power control to ensure that these opportunities are exploited with the help of a central controller. Yet again, the target of these algorithms is to increase throughput while maintaining the minimum data rate of the network.

The uncoordinated nature of the access mechanism used in WLANs or 802.11 networks means that they suffer from collisions [10, 37]. The access mechanism presents an even bigger problem in 802.11ah networks, which have a much wider coverage (up to 1 km) than traditional 802.11 networks and also support low power devices like sensors [21, 26, 27]. One of the key causes of collisions is the random backoff timer used to determine when stations (STAs) may access the medium [37]. If two or more devices choose the same backoff timer value, a collision is inevitable. In this thesis, an approach to setting backoff timers, using the AIDs of the STAs in a group is proposed. It eliminates collisions by ensuring each STA has a unique backoff timer value.

Another issue that plagues 802.11ah networks is outage of cell-edge STAs. In the absence of relays, a few techniques have been proposed to reduce the outage in 802.11ah networks [25]. This thesis investigates the possibility of using connected STAs as relays to forward packets for STAs in outage. The work assumes that connected STAs set their backoff timers using the AIDs of UEs in their group.

1.2 Objectives

The main objective of this research is to improve the performance of wireless networks in terms of key performance indicators including throughput, energy efficiency, minimum data rate, radio frequency (RF) power consumption, outage and delay. The performance indicators used to evaluate the performance of a given scenario may vary depending on the network.

1.3 Contributions

The major contributions of this thesis are summarised as follows:

- Design of a power allocation algorithm that modifies the interference map generated by the SBSs in an indoor small cell network in order to improve the throughput while at least maintaining the minimum data rate of the network. The performance of the technique is shown in both a homogeneous (small-cell only) and heterogeneous network.

- Design of a power control technique to maximise the throughput of a small cell network. The technique uses two power control algorithms to achieve this with one maximising resource block reuse resulting in higher throughput while the other minimises interference for low data rate UEs and in the process improves the minimum data rate of the network.

- Proposal of a novel approach to setting backoff timers of STAs by using their AIDs in sectorised 802.11ah networks to eliminate collisions and reduce idle time. First, the STAs are grouped using a grouping technique that was devised for uniform sized hidden node-free groups in 802.11ah networks. Then the STAs

use the novel AID based approach to set their backoff timers. Finally, a mathematical model that accurately captures the normalised throughput for the distributed coordination function with AID-based backoff counters is presented.

- Proposal of an outage reducing technique using connected STAs as relays for STAs in outage in 802.11ah networks. The connected STAs associate with outage STAs and forward their packets to the access point. The performance of the technique is evaluated for varying number of STAs in the basic service set.

1.4 Thesis Organisation

The remainder of this thesis is organised as follows. Chapter 2 presents the relevant wireless communications theory required for this thesis. The theory includes, among other things, channel models, pathloss models and equalisation, channel access mechanisms, heterogeneous networks and interference characterisation in heterogeneous networks.

Chapter 3 presents a pilot power allocation technique used to improve both the minimum data rate and throughput for an indoor small cell network. The aim of the technique is to modify the interference map generated for radio resource management in order to improve reuse and consequently improve throughput but without degrading the minimum data rate of the network. The technique is shown to improve the minimum data rate, throughput and energy efficiency of two radio resource management techniques.

Chapter 4 presents a power control technique to maximise the throughput of a small cell network. The technique combines two power control algorithms: the first one increases reuse and throughput by allocating UEs to previously forbidden resource blocks as long as they can reduce their power to meet the set threshold; the second algorithm improves the throughput by reducing the excessive power allocated to high data rate UEs in order to improve data rates of lower rate UEs. The results show that the technique is able to improve the throughput albeit the minimum data rate of the network is slightly reduced for higher base station densities.

In Chapter 5, a new approach to setting backoff timers in saturated 802.11ah networks

in order to eliminate collisions is presented. A grouping technique based on clustering and the Welsh-Powell algorithm is used to group the STAs. Performance analysis for the proposed AID-based backoff timers is presented. Through simulations, the performance of the technique in terms of throughput, delay and energy consumption is compared to existing techniques. Further, the performance of a technique to improve the outage in 802.11ah networks is evaluated. The technique uses AID-based backoff timers and assumes all STAs have network virtualisation capabilities that enable them to function as both an STA and a relay at different times.

Finally Chapter 6 presents the conclusions drawn from this thesis and future work to be done.

1.5 List of Publications

Published and Submitted Papers

1. H. Nabuuma, E. Alsusa, W. Pramudito and M. W. Baidas, "A Power Allocation technique for Fairness and Enhanced Energy Efficiency in Future Networks," *in Proc. IEEE International Wireless Communications and Mobile Computing Conference (IWCMC), Paphos, Cyprus 2016.*

2. H. Nabuuma and E. Alsusa, "Enhancing the Throughput of 802.11ah Sectorized Networks using AID-based Backoff Counters," *in Proc. IEEE International Wireless Communications and Mobile Computing Conference (IWCMC), Valencia, Spain 2017.*

3. H. Nabuuma, E. Alsusa and W. Pramudito, "A Load-Aware Base Station Switch-Off Technique for Enhanced Energy Efficiency and Relatively Identical Outage Probability," *in Proc. IEEE Vehicular Technology Conference (VTC Spring), Glasgow, 2015, pp. 1-5.*

4. H. Nabuuma, E. Alsusa and and M. W. Baidas, "Enhancing the Throughput of 802.11ah Sectorised Networks Using AID-based Backoff Timers ," *IEEE Trans. Internet of Things (major corrections).*

5. H. Nabuuma and E. Alsusa, "Throughput Maximisation in Small Cell Networks

using Power Control," *in Proc. IEEE Wireless Communications and Networking Conference (WCNC), Barcelona, Spain 2018 (accepted).*

Under Preparation

1. H. Nabuuma and E. Alsusa, "A comparative study of the performance of polyphase codes and OFDM in the asynchronous uplink of a dense small cell network," *in Proc. IEEE Vehicular Technology Conference (VTC Spring), Porto, Portugal 2018 (under preparation).*

2. H. Nabuuma and E. Alsusa, "Reducing Outage in 802.11ah Using STAs as Relays," *in Proc. IEEE Vehicular Technology Conference (VTC Spring), Porto, Portugal 2018 (under preparation).*

Chapter 2

Background and Overview

THIS chapter presents the background needed for this thesis. This includes an overview of the following: multipath channel fading, pathloss models, equalisation, multiple input multiple output (MIMO) techniques, heterogeneous networks, WLANs and different access mechanisms in wireless networks.

2.1 Radio Wave Propagation

The wireless channel causes fluctuations in propagating signal and this phenomenon is referred to as fading [38–40]. Multiple reflections of the signal cause the radio wave to travel along different paths with different path lengths to get to the receiver. The interaction of these waves causes multipath fading at the receiver which results in rapid fluctuations of the received signal strength over short distances (a couple of wavelengths) [39]. Generally, the average strengths of the waves decrease with increase in separation between the transmitter and the receiver and this is referred to as pathloss [41]. The variation in average signal strength tends to be more gradual and over longer separation distances (hundreds of meters) when compared to multipath fading [39]. For this reason, multipath fading is referred to as small scale fading while the pathloss is referred to as large scale fading.

2.1.1 Large Scale Fading

Large scale fading refers to variation in the average received signal strength over large distances between the transmitter and the receiver. These variations are attributed to pathloss and shadowing.

2.1.1.1 Pathloss

Pathloss is a result of dissipation of the power transmitted by the transmitter. A number of propagation models have been developed to predict the pathloss at a given distance from the transmitter. Most models generally assume that the pathloss at a given transmitter-receiver (T-R) separation is the same [41]. The simplest model for signal propagation is the free space pathloss model which is used to predict received signal strength when there is an unobstructed LOS path between the receiver and the transmitter [39]. The received signal power at a distance d, from the transmitter in free space is given by the Friis free space equation [39, 41],

$$P_r\,[dBm] = P_t\,[dBm] + 10\log_{10}(G_t) + 10\log_{10}(G_r) - PL\,[dB]\,, \tag{2.1}$$

where P_r is the received power in dBm, P_t is the transmit power in dBm, G_t and G_r are the transmitter and receiver gains respectively and PL is the pathloss in dB which is given by [39, 41]:

$$PL\,[dB] = 20\log_{10}\left(\frac{4\pi d_0}{\lambda}\right) + 10\alpha\log_{10}\left(\frac{d}{d_0}\right) \tag{2.2}$$

where α is the pathloss exponent whose value depends on the propagation environment [39, 41], λ is the wavelength, d_0 is the reference distance in the antenna far-field and it is usually assumed to be 1 m in indoor scenarios and 10-100 m in outdoor scenarios [39, 41]. Equation 2.1 is only valid for the far-field of the transmitting antenna.

2.1.1.2 Shadowing

Equation (2.2) implies that the average signal strength at a specific T-R separation is always the same irrespective of the differences in the environment which is not true [39, 41]. The difference in environment causes the received signal strength to significantly differ from average signal strength predicted by (2.2). Measurements have shown that $PL(d)$ is in fact a log normally distributed random variable given by

$$PL(d) = 20\log_{10}\left(\frac{4\pi d_0}{\lambda}\right) + 10\alpha\log_{10}\left(\frac{d}{d_0}\right) + X_\sigma, \tag{2.3}$$

where X_σ is a zero-mean log normally distributed random variable with a standard deviation σ and is referred to as shadowing [38, 39]. Shadowing is caused by clutter or obstacles in the propagation path and it causes receivers at the same separation distance from the transmitter to have markedly different signal strengths [38].

2.1.2 Small Scale Fading

Small scale fading refers to the rapid fluctuations in the received signal phase and amplitude over a short period of time or travel distance [38, 39]. Small scale fading results from either time spreading of the signal, caused by multipath propagation to the receiver, or time-variant behaviour of the channel caused by motion between the transmitter and the receiver [40]. Small scale fading is called Rayleigh fading, when there is no line of sight (LOS) component in the received signal and Ricean fading when there is a dominant nonfading LOS component in the received signal [40].

2.1.2.1 Multipath Channel Model

The multipath channel is modelled as a linear filter with a time varying impulse response where the time variation is due to motion of the receiver and the filtering aspect is due to the summation of amplitudes, phases and delays of the different waves arriving at the receiver [39]. The impulse response $h(t, \tau)$ completely characterises the channel with t representing time variation due to motion and τ representing time spreading due to multipath propagation for a fixed value of t. The received signal $y(t)$

is therefore represented as a convolution of the transmitted signal $x(t)$ and the channel $h(t, \tau)$

$$y(t) = \int_{-\infty}^{\infty} x(t)h(t, \tau)d\tau. \tag{2.4}$$

The baseband channel impulse response is modelled by (2.5) to capture the fact that the received signal in a multipath channel is a summation of attenuated, delayed and phase shifted copies of the transmitted signal.

$$h(t, \tau) = \sum_{l=0}^{L-1} a_l(t, \tau) \exp(j\theta_l)\delta(\tau - \tau_l(t)), \tag{2.5}$$

where $a_l(t, \tau)$, θ_l and $\tau_l(t)$ are the amplitude, phase shift and excess delay (the relative delay of a multipath component to the first arriving multipath component) respectively of the lth multipath component at time t. L is the total number of equally spaced multipath components. If the channel is assumed to be wide sense stationary (time invariant) over a small scale of time or distance, then it is modelled by

$$h(\tau) = \sum_{l=0}^{L-1} a_l \exp(j\theta_l)\delta(\tau - \tau_l). \tag{2.6}$$

2.1.2.2 Fading due to Time Spreading

Time spreading results in different time of arrival for the different multipath signals. One of the problems associated with time spreading is intersymbol interference (ISI) which causes errors during detection at the receiver [39]. A number of parameters are used to quantify the amount of time dispersion in the channel and these include the mean excess delay, rms delay spread and excess delay spread [39]. All these delays are measured relative to the time of arrival of the first multipath component [39]. The mean excess delay is given by

$$\bar{\tau} = \frac{\sum_l a_l^2 \tau_l}{\sum_l a_l^2}. \tag{2.7}$$

The rms delay spread is given by

$$\tau_{RMS} = \sqrt{\bar{\tau^2} - (\bar{\tau})^2}, \tag{2.8}$$

where

$$\bar{\tau^2} = \frac{\sum_i a_i^2 \tau_i^2}{\sum_i a_i^2}. \tag{2.9}$$

The maximum excess delay τ_{max} is a measure of the time it takes for the multipath energy to fall below X_{th} dB below the maximum multipath energy [39]. X_{th} dB is a threshold that relates the multipath noise floor to the power in the strongest multipath component [39]. The delay spread is a measure of the difference in time of arrival of the first signal path and the last significant signal path [38, 40]. A large τ_{RMS} signifies a highly dispersive channel while a small one signifies a channel that is not very dispersive. Generally, $\tau_{max} = 5\tau_{RMS}$ [38]. The dual of delay spread in the frequency domain is the channel coherence bandwidth, $\mathcal{B}_c$ which is the bandwidth over which the channel frequency response is correlated [38, 40]. The coherence bandwidth and the rms delay spread are related by

$$\mathcal{B}_c \approx \frac{1}{5\tau_{RMS}}. \tag{2.10}$$

Fading due to the time spreading is either frequency selective or flat. Flat fading occurs when the coherence bandwidth is greater than the signal bandwidth, $\mathcal{B}_{sig}$, i.e. $\mathcal{B}_c > \mathcal{B}_{sig}$ while frequency selective fading occurs when $\mathcal{B}_c < \mathcal{B}_{sig}$. In the time domain, for a given symbol duration T, this implies that $\tau_{max} < T$ for flat fading and vice versa for frequency selective fading. When $\tau_{max} > T$, the result is deep fades at certain frequencies hence the term 'frequency selective'. In the past, frequency selective fading was a major limitation to the performance of wireless communications however advances in technology now show that it can be manipulated to improve system capacity through schemes like multi-user diversity [38].

Fading due to Time Variant Channel

The relative motion of the transmitter and the receiver causes the channel to change with time t. The time during which the channel's response is highly correlated is called the coherence time, T_c. Relative motion causes Doppler shifts in the received signal in the frequency domain whose magnitude is given by:

$$f_d = \frac{v}{\lambda} cos\hat{\theta} = \frac{fv}{c} cos\hat{\theta}, \tag{2.11}$$

where λ is the wavelength, c is the speed of light, v is the velocity of motion, f is the carrier frequency and $\hat{\theta}$ is the angle of arrival of the signal at the receiver. These shifts cause spectral broadening of the signal [40]. The coherence time and Doppler spread are reciprocally related by [39]:

$$T_c \approx \frac{1}{f_d}. \tag{2.12}$$

Fading due to the time variant nature of the channel is either fast or slow fading. Fast fading occurs when $T_c \leq T$ while slow fading occurs when $T_c \gg T$ from which it can be seen that fast fading usually occurs with low data rate applications [38, 39]. A number of techniques have been developed to mitigate the effects of fading including the use of diversity techniques, equalisation and the use of multicarrier modulation.

2.2 Equalisation

Equalisation is a term used to describe all signal processing done to the received signal to eliminate or reduce the impact of ISI on the received signal [39]. Owing to the random and time variant nature of the wireless channel, effective equalisers should be able to track the channel [39, 42]. Such filters are referred to adaptive equalisers. If $x(t)$ is the transmitted signal and $h(t)$ is the combined baseband impulse response of the channel, transmitter and RF/IF sections of the receiver, then the received signal is

$$y(t) = x(t) * h(t) + n(t), \tag{2.13}$$

where $n(t)$ is the noise at the input of the equaliser and $*$ denotes convolution. Denoting the impulse response of the equaliser by $h_{eq}(t)$, the output of the equaliser is given by

$$\hat{x}(t) = x(t) * h(t) * h_{eq}(t) + n(t) * h_{eq}(t). \tag{2.14}$$

The baseband impulse response of the equaliser is given by

$$h_{eq}(t) = \sum_i e_i \delta(t - T), \tag{2.15}$$

when e_i are the filter coefficients of the equaliser. Ignoring the noise in (2.14), in order to get the transmitted signal at the output of the equaliser then

$$h(t) * h_{eq}(t) = \delta(t). \tag{2.16}$$

In the frequency domain, (2.16) is given by

$$H_{eq}(f) H(f) = 1, \quad |f| < \tfrac{1}{2T} \tag{2.17}$$

where $H_{eq}(f)$ and $H(f)$ are the Fourier transforms of $h_{eq}(t)$ and $h(t)$, respectively. The most common equaliser structure is the linear transversal equaliser which consists of tapped delay lines spaced a symbol duration apart. In this filter, the past, current and delayed values of the received signal are linearly weighted by the filter coefficients and summed to generate the output.

2.3 Multiple Access Techniques

Cellular networks have an uplink and downlink. The uplink, also commonly referred to as the reverse link has many transmitters sending signals to one receiver which is usually referred to as the base station (BS) [41]. On the other hand, the downlink, also commonly referred to as the forward link, has one transmitter sending signals to several receivers [41]. In current network implementations, it is generally impossible to simultaneously receive and transmit on the same frequency because of the resulting

interference. Therefore, the uplink and the downlink are orthogonal to each other in either the time domain or the frequency domain [41]. This process of separating the uplink and downlink is referred to as duplexing. Time division duplexing (TDD) involves assigning orthogonal time slots for the uplink and the downlink while frequency division duplexing (FDD) involves assigning different frequency bands for the uplink and downlink. One key benefit of TDD over FDD is that bidirectional channels tend to have the same channel gains hence channel estimation on say the downlink can be used to estimate the channel in the uplink [41].

Both the uplink and the downlink are multi-user channels however they need the use of multiple access techniques to enable multiple users to share the spectrum in an efficient manner. In order for users to share the spectrum, they need to be orthogonal to each other in the time domain, frequency domain or code domain. For orthogonality in the time domain, users use the same spectrum but each user accesses the spectrum in their allocated time slot [40, 41]. This is referred to as time division multiple access (TDMA) [40]. In the frequency domain, there's frequency division multiple access (FDMA) where the spectrum is divided into orthogonal channels and each channel is allocated to a particular user [39,40]. In systems that use orthogonal frequency division multiplexing (OFDM), multiple access is referred to as orthogonal frequency division multiple access (OFDMA), where FDMA is implemented by assigning different users to different subcarriers [28]. Code division multiple access (CDMA) allows all users to use the whole spectrum simultaneously and achieves orthogonality in the code domain [39, 40]. To attain orthogonality in the code domain, each user's signal is spread using a code that is orthogonal to other users' codes [40]. Once the receiver correlates the received signal with user's code, the other users interfering signals are suppressed [40].

In WLANs, packet radio access techniques are designed to enable several users to access the channel in an uncoordinated manner. Because the traffic is bursty, dedicated channel assignment is inefficient for such scenarios [39, 41]. The lack of coordination means that collisions occur at the receiver when two or more users simultaneously transmit data [39]. One of the major protocols used is carrier sense multiple access (CSMA) which employs a *listen-before-talk* policy that requires transmitters to first sense the channel before starting their own transmissions.

The remainder of this section presents some more information on OFDMA and CSMA which are relevant to the work presented in later chapters of this thesis.

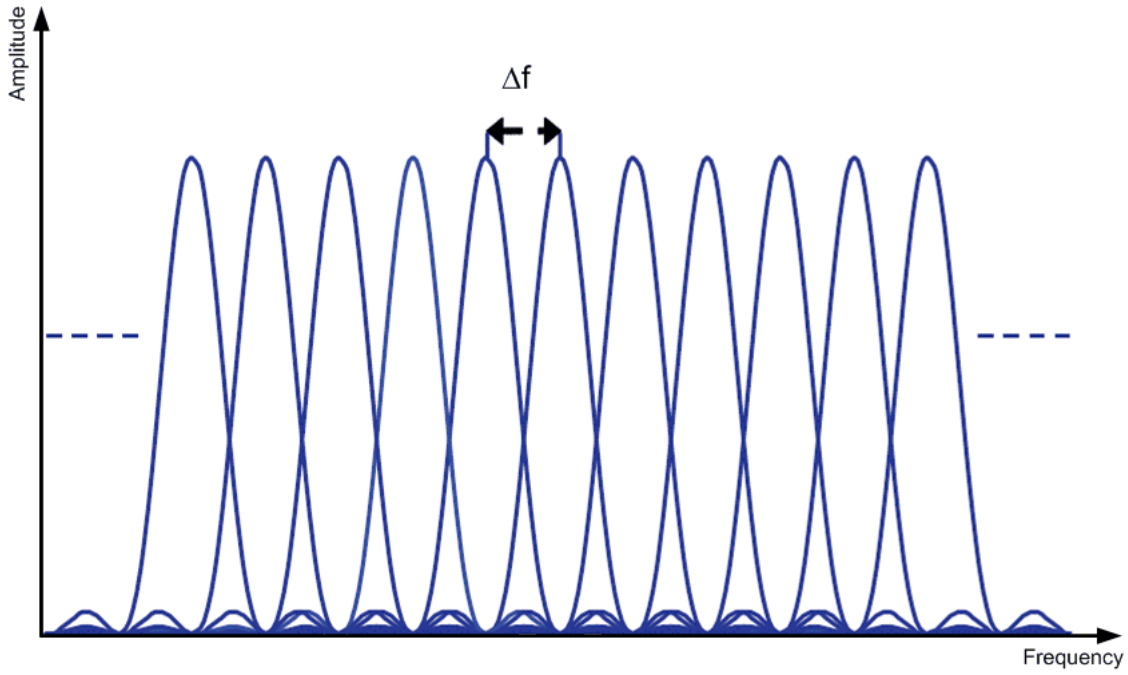


Figure 2.1: OFDM Subcarrier Overlap [1]

2.3.1 OFDMA

Orthogonal Frequency Division Multiple Access (OFDMA) is a scheme that uses OFDM to enable the simultaneous access to bandwidth by different users. The benefits of OFDM include: simple equalisation, scalability to different bandwidths, and efficient handling of different data rates simultaneously [43]. OFDM uses a multitude of orthogonal subcarriers to transmit data where the orthogonality is not achieved by separation of the subcarriers but rather by the shape and spacing of the subcarriers [43]. Figure 2.1 shows how the OFDM subcarriers overlap and from it, it can be seen that with properly timed sampling, orthogonality between subcarriers is achieved.

An OFDM symbol over the duration $mT_0 \leq t < (m+1)T_0$ is represented in complex baseband notation by 2.18

$$x(t) = \sum_{k=0}^{N_{sub}-1} \tilde{a}_k^{(m)} e^{j2\pi k \triangle f t}, \tag{2.18}$$

where $\triangle f$ is the subcarrier spacing, T_0 is the OFDM symbol duration, $\tilde{a}_k^{(m)}$ is the, generally complex, modulation symbol applied to the *kth* subcarrier during the *mth* OFDM symbol interval [1] and N_{sub} is the number of subcarriers.

In the time domain, each subcarrier is a pulse with duration T_0 and $\Delta f = 1/T_0$ to ensure orthogonality. OFDM is efficiently implemented using Inverse Fast Fourier Transform/ Fast Fourier Transform (IFFT/FFT) processing. The nature of the OFDM symbol makes it possible to allocate the resource in both time and frequency domain. This is why the resource in OFDM is represented as a grid.

2.3.1.1 Cyclic Prefix Insertion

The time dispersive nature of the wireless channel discussed earlier results in ISI as well as inter-carrier interference which results in loss of subcarrier orthogonality at sampling time. This is solved by the insertion of a cyclic prefix (CP) that is at least as long as the channel's delay spread. The CP is added by copying the last part of the OFDM symbol to the beginning of the symbol. The CP is dropped at the receiver and if it is at least as long as the time dispersion, the received symbol will have no ISI [1, 28]. When the CP is dropped, the linear convolution of the time dispersive channel becomes a circular convolution at the receiver during the interval T_0 and after FFT processing, the received signal is simply the transmitted symbols multiplied by channel taps. The received signal is easily equalised by multiplication with the complex conjugate of the channel taps [1, 28]. The CP reduces the spectral efficiency of OFDM because it is dropped however, its use simplifies equalisation. Therefore, the use of CP is a trade-off between the power, bandwidth loss and the simplification of equalisation [1].

2.3.1.2 OFDMA in LTE

In Long Term Evolution (LTE) networks, the OFDM symbol duration, T_0, is $66.7\mu s$ derived from a subcarrier spacing of $15\ kHz = 1/T_0$. The subcarrier spacing was selected to provide a balance between sensitivity to channel time selectivity and the CP overhead [1]. Further, this spacing enables easier implementation of dual-mode LTE/HSPA handsets as the LTE sampling rate is either a multiple or sub-multiple of the HSPA sampling rate of $3.84\ Mchip/s$ [1]. A reduced subcarrier spacing of $7.5\,kHz$ is supported for multicast transmissions [1, 28]. LTE transmissions are organised into $10\ ms$ radio frames that are further divided into ten subframes of length $1\ ms$ each. A subframe contains two slots of length $0.5\,ms$, with each slot consisting a number of OFDM symbols plus a cyclic prefix. Normal subframes contain 7 OFDM symbols

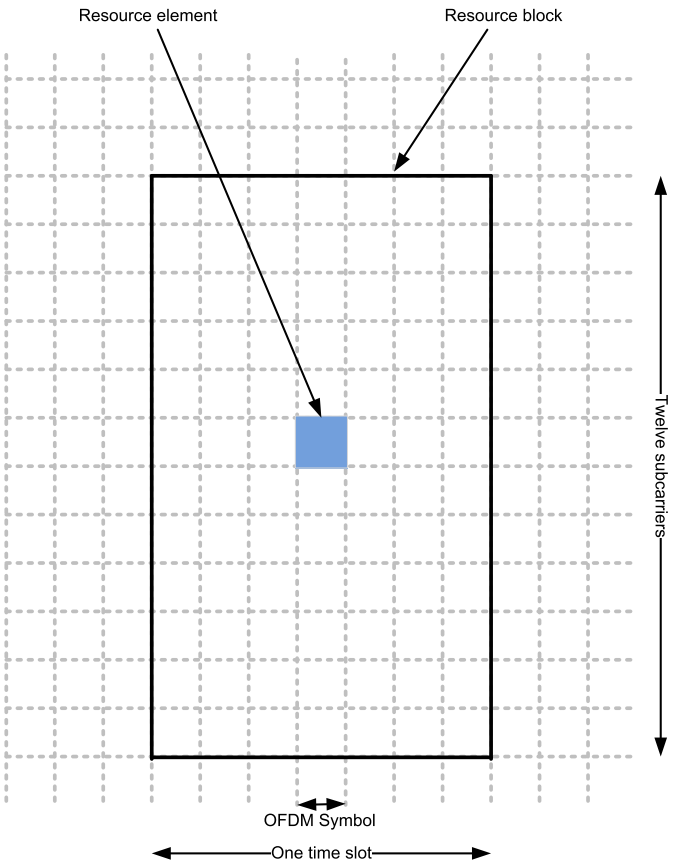


Figure 2.2: LTE time-frequency resource

[1]. The smallest physical resource in LTE is a resource element (RE) which is one subcarrier in one symbol duration as shown in Figure 2.2.

REs are grouped into resource blocks (RBs) i.e. 12 consecutive subcarriers in the frequency domain and 1 slot in the time domain [1]. For normal subframes this implies 84 resource elements per resource block. The basic scheduling unit in LTE is a resource-block pair and transmissions are scheduled every 1 ms which is known as the transmission time interval (TTI) [1]. The use of RBs as the basic allocation unit means that LTE can benefit from the multi-user diversity [1]. LTE supports transmission bandwidths from around 1 MHz to 20 MHz and in LTE release 10 the bandwidth can increase to 100 MHz when carrier aggregation is used [1].

2.3.1.3 OFDM in the Uplink

One problem with OFDM is that it has large peak to average peak ratio (PAPR) variations which require inefficient amplifiers at the transmitter. Because the uplink transmitter is power constrained, OFDM cannot be used. However a variation of OFDM called single carrier OFDM (SC-OFDM) is used to rectify this problem. It involves first applying a block of M modulation symbols to a size M Discrete Fourier Transform (DFT) and then applying its output to inputs of the IFFT processing for normal OFDM transmitter. M should be less than N_{FFT}. At the receiver, after the FFT, a size M inverse DFT (IDFT) should be applied in order to recover the modulation symbols [1, 28]. This approach reduces the PAPR variations in OFDM and is used for uplink transmissions in LTE for this reason.

2.3.2 Carrier Sense Multiple Access (CSMA)

In CSMA, once the channel is detected to be busy, the transmitter waits until it detects that the channel is idle. It then waits a random time period, normally referred to as a *random backoff*, before transmitting. The random backoff is meant to reduce the probability of collisions by preventing multiple transmitters from transmitting as soon as the channel is idle. Further, CSMA requires the receiver to send an acknowledgement if the data is correctly received. If the transmitter does not receive this acknowledgement, it will assume that the packet was not correctly received and will schedule it for retransmission. In the wireless networks, the channel may prevent some users from detecting other users' signals and this is referred to as the hidden node problem [41]. Wireless local area networks (WLANs) based on the 802.11 standards use the Distributed Coordination Function (DCF), which is based on CSMA, for asynchronous data transmission.

2.3.2.1 Distributed Coordination Function

The Distributed Coordination Function (DCF) is used for uncoordinated asynchronous data transmission. Its operation is described next. In line with CSMA, before a device accesses the channel it senses the channel to ensure that it is not busy for a DCF interframe space (DIFS) period [37, 44]. After the DIFS period, the device enters a backoff

period where it selects a random backoff time uniformly between $[0 - CW_{min} - 1]$, where CW_{min} is the minimum contention window size. During the backoff time the device senses the channel further and if the channel is idle at the end of this period, the device starts to transmit. If it is busy, the device will freeze its backoff timer until it detects that the channel is idle again for a DIFS period [45]. Once the transmission is complete, the receiving device waits for a short inter-frame space (SIFS) period and then it transmits an ACK message to acknowledge successful reception of the frame. If more than one device has the same backoff time and a collision occurs or if the packet is received with errors, the receiver does not send an ACK message. When the transmitter doesn't receive an ACK message, it retains the packet for retransmission and doubles the contention window size. The contention window is doubled after each unacknowledged retransmission of a packet until the maximum contention window size, CW_{max}, is reached [37, 44]. For subsequent retransmissions, the contention window is frozen at CW_{max} until either the packet is dropped because it has reached the retransmission limit or an ACK is received from the receiver. The contention window is then reset to CW_{min} for the next packet in the queue [37, 44].

2.3.2.2 Point Coordination Function

802.11 standards also support a central controller in the form of the point coordination function (PCF) where the access point (AP) is the central controller and determines the access sequence of STAs in the WLAN [44, 46, 47]. Only a polled STA is allowed to access the channel. The AP takes control after a PCF inter-frame space (PIFS) duration and then sends a POLL frame [44, 46]. The polled STA sends a data frame after a SIFS period. The AP then sends a POLL+ACK frame after another SIFS period to simultaneously acknowledge the previously received message and poll the next STA. If the polled STA does not respond, the AP waits a PIFs duration and polls the next STA. At the beginning of the contention free (CF) period, when polling takes place, a beacon is sent by the AP while at the end of the contention free period, the CF-END frame is sent. The challenge with PCF is that it is not as widely implemented as DCF [44, 46, 47].

2.4 Multiple Antenna Techniques

The use of multiple antennas enables the creation of independent channels between a transmitter and receiver that can be exploited to improve capacity through diversity, eigen beamforming, which focuses energy on a specific receiver hence improving SINR, and creation of multiple parallel channels between the transmitter and receiver (spatial multiplexing) [38]. The use of multiple transmitters at the receiver and transmitter is commonly referred to as MIMO communication. Spatial diversity involves the use of two or more appropriately spaced antennas to eliminate deep fades in the received signal while frequency diversity involves the sending the same data on different frequencies and to increase the probability of receiving at least one signal without deep fades [38].

Spatial diversity is popular because it does not require extra power or bandwidth to implement unlike frequency diversity [38]. Two forms of spatial diversity are receive diversity and transmit diversity. Receive diversity is very common at BSs because most UEs do not have more than one antenna yet this is common at BSs [38]. The manner in which the received signal from the antennas is manipulated determines how much will be gained from the diversity. Two popular approaches are selection combining (SC) and maximal ratio combining (MRC). SC is a simple combiner that selects the signal with the highest power. It is not efficient because it wastes signal energy however its simplicity is its selling point [38]. Transmit diversity necessitates processing at both the transmitter and the receiver unlike receive diversity because the signal is sent from two different antennas. Transmit diversity is considered suitable for the downlink because the base station has the multiple antennas required. Transmit diversity can be either open loop or closed loop. Open loop transmit diversity does not require information about the channel at the transmitter while closed loop transmit diversity requires it [38]. The most popular form of open loop transmit-diversity scheme is space-time coding where a code that is known to the receiver is applied to the data before transmission [38].

Space-time block codes (STBC), a type of space-time codes, are the most popular means of achieving transmit diversity because of the ease of implementation [38]. STBC with two transmit antennas is implemented as follows: Two symbols, $s_1$and s_2 are transmitted as follows over two symbol periods: antenna 1 sends s_1 and antenna 2 send s_2 at time 0 then $-s_2^*$and s_1^* are sent by antenna 1 and 2 respectively at time

1. Assuming flat fading and that the channel stays constant over the two symbols, h_1 is the channel gain between antenna 1 and the receiver while h_2 is the channel gain between antenna 2 and the receiver. The received signal is given by [38]:

$$\begin{aligned} r(0) = &\quad h_1 s_1 + h_2 s_2 + n(0) \\ r(T) = &\quad -h_1 s_2^* + h_2 s_1^* + n(T) \end{aligned}, \tag{2.19}$$

where $n(\cdot)$ denotes white Gaussian noise. The following diversity scheme is used at the receiver

$$\begin{aligned} y_1 = &\quad h_1^* r(0) + h_2 r^*(T) \\ y_2 = &\quad h_2^* r(0) - h_1 r^*(T) \end{aligned}. \tag{2.20}$$

With substitution and simplification (2.20) becomes

$$\begin{aligned} y_1 = &\quad (|h_1|^2 + |h_2|^2)\, s_2 + h_1^* n(0) - h_2 n^*(T) \\ y_2 = &\quad (|h_1|^2 + |h_2|^2)\, s_2 + h_2^* n(0) - h_1 n^*(T) \end{aligned}. \tag{2.21}$$

From (2.21), the SNR is computed to get

$$SNR_{STBC} = \frac{\sum_{i=1}^{2} |h_i|^2}{n^2} \frac{P_t}{2} \tag{2.22}$$

For fair comparison with MRC, the transmit power, P_t, for each antenna is halved. Therefore it can be seen that STBC is similar to MRC however there is a 3 dB penalty from the use of two transmit antennas. Notably, unlike MRC, as the number of antennas increases, transmit diversity causes the received SNR to harden because the combined signal SNR approaches the expectation of the channel gain [38]. It therefore eliminates fading but does not improve SNR [38].

MIMO also enables improved SINR through the use of eigen value decomposition techniques and improved capacity through spatial multiplexing [38].

2.5 Heterogeneous Networks

Heterogeneous networks (HetNets) have grown out of the need to meet the rising demand for high data rates by users. The fact that majority of high data rate users are indoors has raised the research interest in small cells like femtocells and distributed antenna systems which are deployed indoors. Furthermore, the lower costs associated with deploying and operating small cells make them an attractive option for mobile network operators [48].

The macrocell base station (MBS) is a key component of the heterogeneous network. MBSs usually have a coverage anywhere from 500 m up to 10 km with larger coverage areas found mainly in rural areas while smaller ones are in urban areas. MBSs are elevated and are associated with high gains compared to other HetNet components. MBSs are deployed and managed by the operator and are connected to other MBSs via an X2 interface. In HetNets, MBSs are particularly useful for providing coverage to high mobility users that could get dropped calls if they attempt to handover to each SBS they traverse [49, 50]. The five major types of SBSs are microcells, picocells, distributed antenna systems , relays and femtocells [8]. Apart from femtocells, the SBSs are operator deployed and managed low power BSs which are connected to other BSs, including the MBS, via an X2 interface which implies that they are able to benefit from inter-cell interference coordination (ICIC), a technique that mitigates inter-cell interference (ICI). For outdoor deployment, the transmission power for SBSs ranges from 250mW to 2W while for indoor deployments, the maximum power is 100mW [22]. Unlike other SBSs, femtocells can be configured to have restricted access [22,51]. There are two major types of access control for femtocells, open and closed access, and they have an impact on interference management and overall performance of the femtocell network [51]. Open access femtocells allow all UEs in their coverage to connect to them resulting in better network capacity while degrading the QoS of the femtocell owner [51, 52]. Closed access femtocells only allow access to pre-registered UEs resulting in “dead zones” for other UEs [53, 54].

The deployment of SBSs has its own challenges with one of the major ones being interference [48]. Co-channel deployment of SBSs with macrocells results in both co-tier and cross-tier interference [48, 55, 56]. Co-tier interference is the interference between collocated SBSs while cross-tier interference is the interference between SBSs and the MBS [57]. In the downlink, cross-tier interference is between an MBS and an SBS UE

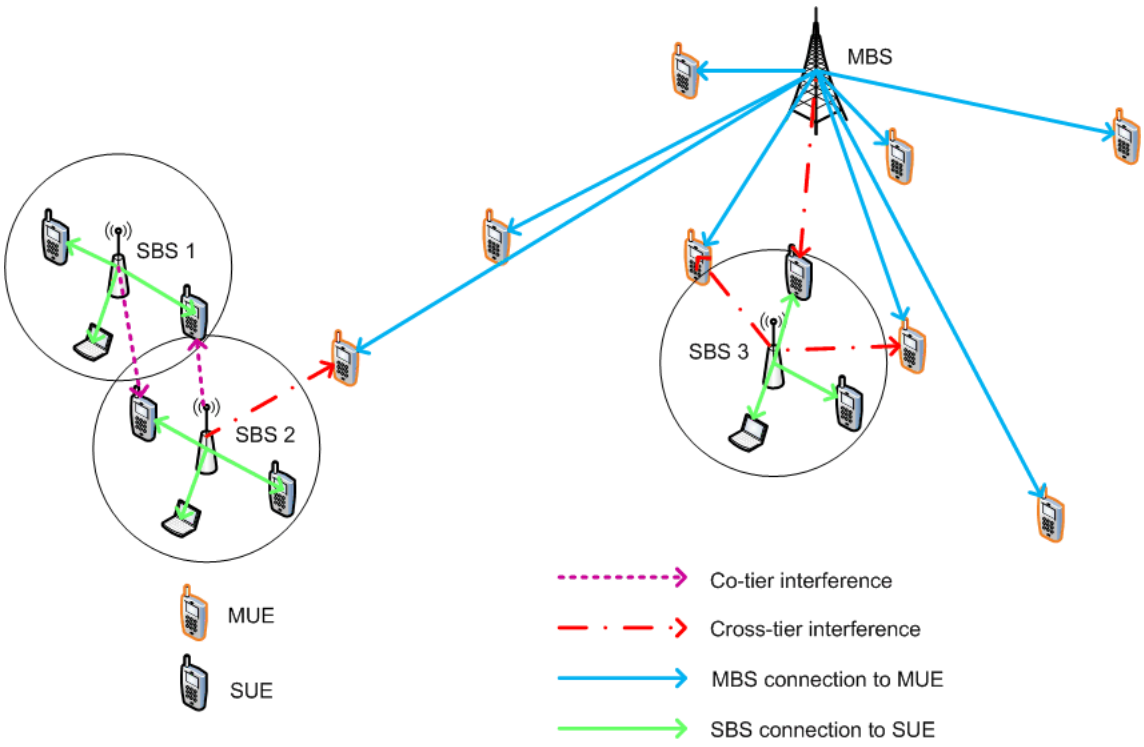


Figure 2.3: Downlink interference in a HetNet

(SUE) and between a macrocell UE (MUE) and an SBS as shown in Figure 2.3 while in the uplink, the reverse is true as shown in Figure 2.4. On the other hand, co-tier interference is between two or more neighbouring SBSs whose coverage areas overlap as shown in Figure 2.4. In the downlink, the SBSs interfere with the SUEs in the overlap area while in the uplink the SUEs interfere with neighbouring SBSs as shown in Figure 2.3 and Figure 2.4 respectively.

Orthogonal deployment of SBSs, on the other hand, only results in co-tier interference because the small cells and the macrocell layer are allocated orthogonal spectrum and thus cannot interfere with each other [57]. However orthogonal deployment of small cells is considered spectrally inefficient because one tier may not need all the spectrum allocated to it but still block the other tier from using the spectrum [57].

2.5.1 Interference Management (IM) in Macrocell-only Networks

In macrocell networks with universal frequency reuse, there is a huge disparity between cell-edge and cell-center user signal to interference plus noise ratios (SINRs).

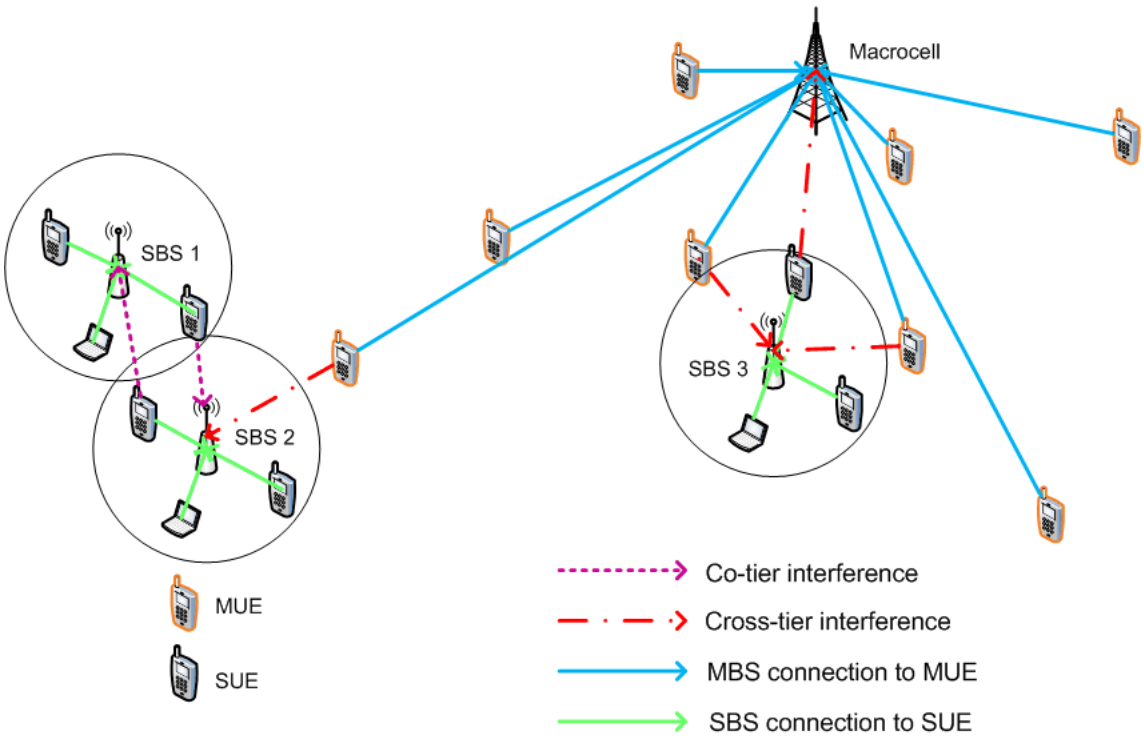


Figure 2.4: Uplink interference in a HetNet

Cell-edge SINRs are low because of either interference, in smaller cells in urban areas, or noise, in case of larger cells in rural areas [28]. Increasing transmitted power can improve the SINRs in the noise-limited scenario. However, in interference-limited networks, increase in power exacerbates the interference problem [28]. The degradation in SINR at the cell-edge is attributed to reducing power from the serving base station due to pathloss and increasing inter-cell interference (ICI) because of the shorter distance to interfering macro base stations (MBSs) compared to a cell-centre user. There are three major categories of ICI management techniques: interference cancellation, interference randomisation and interference avoidance [28]. Interference cancellation requires advanced receivers, with antenna arrays and interference cancellation or suppression or both in order to separate the noise from the desired signal. Additionally suppression errors significantly degrade the performance of this technique [58] . Interference randomisation relies on use of scrambling codes to suppress interference using processing gain at the receiver [28]. Finally interference avoidance techniques rely on algorithms for resource allocation that minimise interference. Therefore interference avoidance techniques mainly affect the scheduler and are easier to integrate into existing systems [58]. Two popular ICI avoidance techniques have emerged i.e. fractional

frequency reuse (FFR) and soft frequency reuse (SFR). They are also referred to as intercell interference coordination (ICIC) techniques because neighbouring macrocells coordinate on resource allocation for the cell edge over the X2 interface [28]. These two techniques rely on partitioning of resources at the cell edge to enhance cell edge performance. The results of numerical analysis presented in [28] make the case for SFR and FFR because they show that a frequency reuse ratio greater than one provides better SINR and spectrum efficiency for cell edge users but offers degradation for cell center users in terms of spectrum efficiency. This is because the cell-centre user's SINR is not dominated by ICI and therefore its elimination will only slightly improve the SINR however the reduction of resources due to partitioning reduces its data rate more significantly [59].

2.5.1.1 Fractional Frequency Reuse

Depending on the reuse factor, FFR will partition resources in such a way that the center users of the cell (or sector), for all cells, use the same frequency band f_1 while the remaining spectrum is divided up to be used exclusively by the cell-edge users of each cell. Therefore, each cell uses a different frequency band in its cell edge in order to minimise ICI. Figure 2.5 shows an example of FFR with reuse factor three. FFR is inefficient in its spectrum usage because some frequency bands are unused in each cell [59].

2.5.1.2 Soft Frequency Reuse

Unlike FFR, in SFR, all frequency bands are used by all cells. Figure 2.6 shows SFR with reuse factor of three. With SFR, the whole bandwidth is divided into three orthogonal bands. The cell-centre users use two of the bands while the cell edge users use one of the bands. To mitigate ICI, the band used by cell edge users in one cell is used as a center user band for the neighbouring cells as shown in Figure 2.6. ICI is further mitigated by transmitting a lower power on the cell-centre frequencies while higher power is transmitted on cell-edge frequency bands [59].

Critiques of the SFR and FFR note that they are built on the assumption that the traffic and its distribution in the cell is constant which is not true. To address this problem,

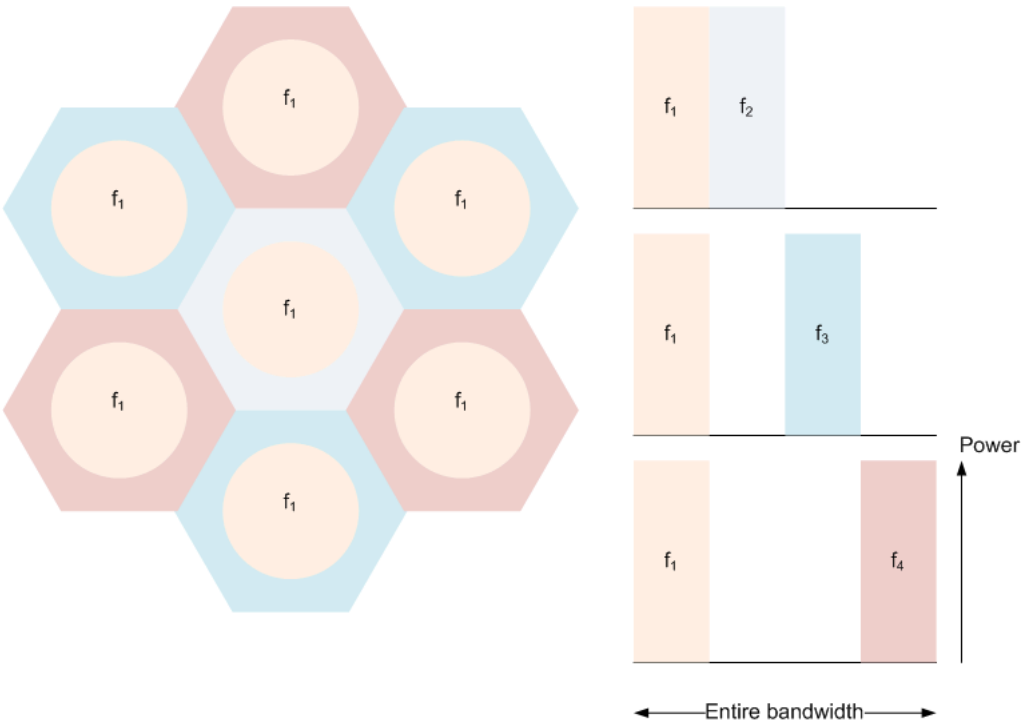


Figure 2.5: Illustration of fractional frequency reuse (FFR) scheme [2]

adaptive frequency reuse schemes have been suggested that partition the cell based on the traffic experienced in the network [59].

2.6 Clustering Techniques

Several definitions of a cluster are defined in the literature, however the most relevant one for this thesis is: a cluster is a set or group of points such that the distance between any two points within the cluster is less than the distance between any point in the cluster and any point outside the cluster [60]. Therefore clustering techniques are used to group data or objects based on a similarity measure e.g. distance [60]. Clustering methods are normally classified as either partitional or hierarchical methods. Hierarchical clustering transforms a proximity matrix (matrix with similarity measure) into a sequence of nested partitions where each partition is nested into the next partition in the sequence [60]. Hierarchical clustering is either agglomerative of divisive where agglomerative clustering with every point being an individual cluster [61,62]. Depending on the choice of clustering algorithm, the proximity matrix is interpreted to determine

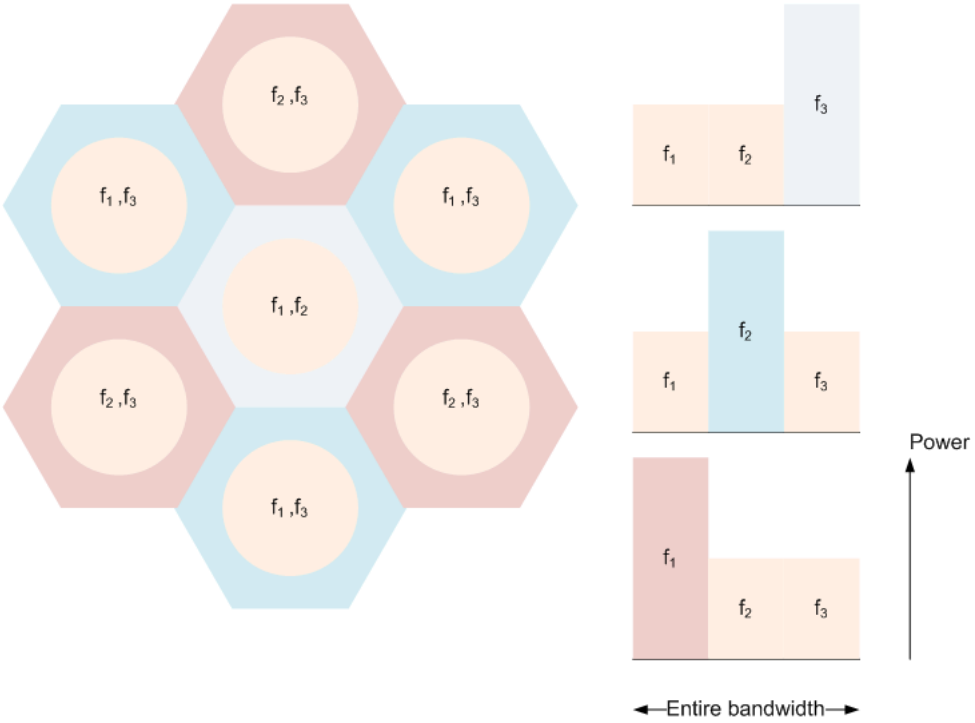


Figure 2.6: Illustration of soft frequency reuse (SFR) scheme [2]

which two clusters to merge first. This process is repeated until there is only one cluster [60, 61]. A divisive clustering starts with one cluster and keeps splitting it until every point is an individual cluster. Hierarchical clusterings are usually represented by a dendogram shown in Figure 2.7. The dendogram consist of layers of nodes, with each layer representing a cluster. Cutting the dendogram horizontally produces a clustering [60]. Where exactly to cut the dendogram, depends on how many clusters are required. For instance in Figure 2.7, if the dendogram is cut at level 1, there will be three clusters, while at level 2, there will be two clusters.

2.6.1 Agglomerative Clustering algorithms

A number of algorithms exist in literature for hierarchical clustering including, but not limited to, single-link, complete-link and Ward's method. Given proximity matrix $\mathcal{D} = [d(i, j)]$ where $d(i, j)$ is the distance between points i and j, it should be clear that the $\mathcal{D}$ is a symmetrical matrix. If it is a dissimilarity matrix then $d(1, 2) < d(1, 3)$ means that points 1 and 2 are more alike than points 1 and 3. Therefore points 1 and 2 will be merged before 1 and 3. For the single link algorithm, inter-cluster distance

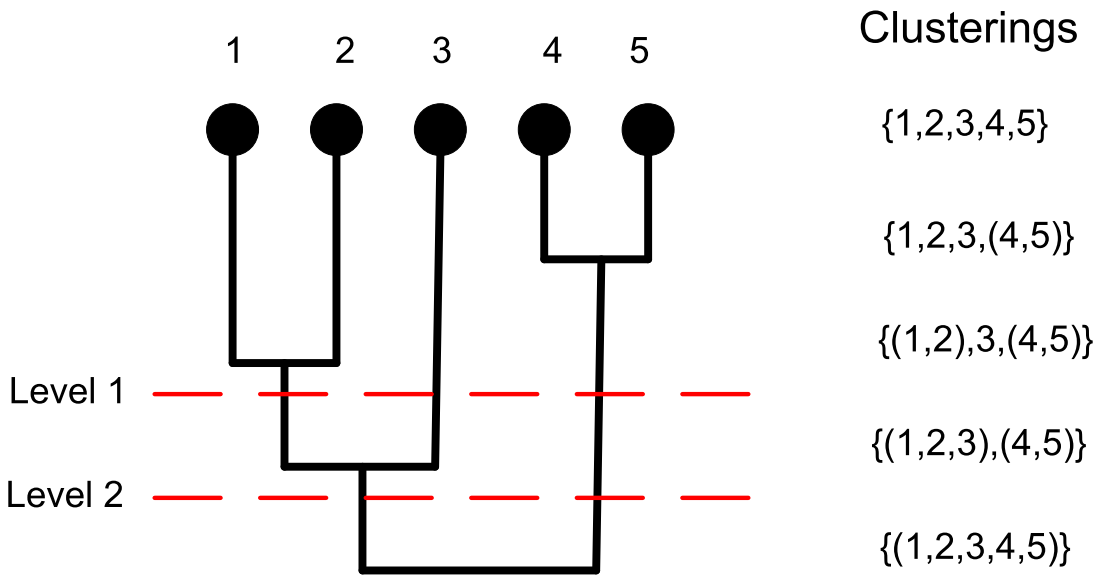


Figure 2.7: Dendogram Example

is the distance between the two objects closest to each other in the different clusters while for the complete link algorithm it is the largest distance between points in the different clusters [61]. Unlike single linkage algorithm and complete link algorithm, Ward's method uses geometric methods to calculate distances because it uses geometric centers to represent clusters. For Ward's method the inter-cluster distance is given by

$$\mathcal{D}_w(C_i, C_j) = \sum_{x \in C_i} (x - c_i)^2 + \sum_{x \in C_j} (x - c_j)^2 - \sum_{x \in C_{ij}} (x - c_{ij})^2, \tag{2.23}$$

where C_i denotes cluster i, x denotes a point in the cluster, c_i denotes the centroid of cluster i, c_{ij} denotes the centroid of merged clusters i and j, and the cluster centroid is the average of points in the cluster. With Ward's distance as the linkage metric, clusters that minimise $\mathcal{D}_w(C_i, C_j)$ will be merged first [61, 62]. Of the agglomerative clustering methods, Ward's method is the only one whose inter-cluster distance metric is based on the sum-of-squares criterion and this causes it to produce groups that minimize within-group dispersion at each merging stage [63].

2.7 WLAN Basics

The basic building block of an IEEE 802.11 LAN is called the basic service set (BSS). It contains at least one access point (AP) and at least one station (STA) in case of infrastructure mode while for ad-hoc mode, it is made up of at least two STAs (referred to as independent BSS (IBSS)) [64]. All STAs in an infrastructure mode BSS communicate through the AP [64]. A distribution system (DS) is used to link multiple BSSs in order to increase the WLAN coverage and provide mobility support. STAs connect to the AP via the wireless medium while the AP has a wired connection to other APs that are connected to the DS [64]. The BSSs connected to the same DS are referred to as an extended service set (ESS).

Information is transmitted in the form of frames in 802.11 networks. The formats of the media access control (MAC) frames or MAC protocol data unit (MPDU) and physical layer protocol data unit (PPDU) are described in following subsections.

2.7.1 MAC Frame Formats

Each MAC frame has three distinct components i.e. MAC header, frame body and frame check sequence (FCS) [64]. The MAC header contains the following information: frame control, duration, addresses, sequence control information [64]. The frame control field contains general control information including the protocol version, what type of frame it is (data, management or control frame), whether the frame is to or from the DS, and other control information [14, 64]. The duration field specifies the duration of the frame in microseconds so that the other STAs can determine how long they have to wait before trying to contend for the wireless medium. There are four possible address fields although only two are mandatory i.e. the source and destination addresses [64]. The other two addresses are the transmitting STA and receiving STA addresses which are not used in infrastructure mode BSSs [14,64]. Sequence control is used in cases where frame fragmentation is implemented to ensure that the proper sequence of the fragmented frame is restored [64]. Regarding the type of frames, control frames include frames like the ACK frame or RTS frames which are used for handshaking while management frames include frames like the beacon frame, association frames, authentication frames, timing frames etc. Finally, data frames are used for transmission of data [14].

The frame body is a variable length field that depends on the type of frame sent. The minimum frame body length is 0 bytes, usually associated with control frames, while the maximum length varies depending on the standard [64].

The FCS is an error detecting code i.e. it is a cyclic redundancy check (CRC) that is calculated over the fields of the MAC header and the frame body [64]. The receiving STA generates its own CRC based on the received MAC header and frame body and compares it to the FCS to determine if the received frame contains errors [64].

2.7.2 PHY Frame Format

The Physical layer frame is referred to as a physical layer converegence procedure (PLCP) protocol data unit (PPDU) format is made of three components i.e. preamble, header and PLCP service data unit (PSDU) [64]. The preamble contains the synchronisation information. The header contains the modulation and coding information for the rest of the PPDU, the length which indicates the size in bytes of the PSDU, tail bits, service bits, parity bits and CRC bits [64]. Finally, the PSDU contains the contents of the PPDU i.e. the frame being transmitted to or from the MAC layer [64].

2.8 Summary

Chapter 2 has presented the relevant background theory required for this thesis. The different types of fading that occur in wireless networks and their effects were presented along with the relevant models for them. Next, various multiple access techniques used in wireless networks were presented and the details of OFDMA, and CSMA were briefly discussed. A brief discussion on multiple antenna techniques was presented followed by an introduction and brief discussion on the concept of heterogeneous networks. This was followed by a brief discussion on clustering techniques which are used for grouping users in later chapters. Finally, some WLAN basics were introduced including frame formats and the components of a WLAN.

Chapter 3

Power Allocation Technique for SON RRM

IN this chapter a power allocation technique for enhancing the throughput of two self organising radio resource management (RRM) techniques is proposed. The work presented in this chapter is based on the work in [65]. The initial part of the chapter presents a review of the literature on interference avoidance techniques which zeros in on self organising network (SON) RRM techniques. This is followed by a review of pilot power allocation techniques. Next, two algorithms that aim to improve the throughput of SON RRM techniques by modifying the interference map that is used to guide resource allocation are presented. The latter part of this chapter presents an evaluation of the performance of the power allocation technique.

3.1 Literature Review

Interference avoidance is an approach to IM which involves elimination or minimisation of the interference experience by users in the network. This can be achieved via hardware based techniques or resource allocation. Hardware based techniques use advanced antennas like antenna arrays or sectored antennas and antenna techniques e.g. beamforming to minimise the interference experienced by users [66, 67]. Hardware based techniques are expensive and would significantly increase the cost of the SBS. Interference avoidance based on resource allocation i.e. power allocation and RRM

does not require any changes to hardware because it involves algorithms that are applied centrally by a controller (e.g. femtocell management system (FMS) or Operator Management System (OMS)) or in a distributed manner by the BSs to minimise the interference received by UEs in the network.

Resource allocation is usually either centralised or distributed. For centralised resource allocation, all the BSs gather information on interferers from their UEs and send it to a central controller which then allocates the resources to ensure that interferers are orthogonal to each other. The central controller then sends the allocation decision back to the BSs for example in the form of forbidden RBs or specification of maximum power transmission or both [67]. With distributed resource allocation, on the other hand, each BS independently allocates its resources. The BS may base its allocation decisions on its perception of the interference environment [67–69] while in other cases, there may be some negotiation or coordination between the BSs in form of utility functions and resource pricing. These utility functions and pricing mechanisms guide resource allocation decisions for BSs e.g. an RB with high interference may have lower utility than one with low interference therefore, the BS will select the one with lower interference. These distributed allocation mechanisms usually need a number of iterations to converge to the optimal solution [70–72]. Distributed allocation offers simplicity and ease of implementation however the centralized allocation offers better spectral efficiency [67, 73] albeit with signalling overhead. Different techniques are used to adapt the control overhead in centralised allocation especially when femtocells are involved. These techniques include clustering of femtocells, event based allocation updates only e.g. when the interference environment changes, and allocation updates to only affected BSs [67].

3.1.1 RRM in HetNets

Allocation of orthogonal resources to the different tiers is a simple approach to eliminating cross-tier interference [67, 68, 74–76]. This is broadly referred to as resource partitioning (RP) [57]. Although it has proved to perform well in macrocells with high SBS density [77], it is spectrally inefficient when the traffic in the tiers is not balanced i.e. when one tier has high traffic while the other has low traffic [67, 75, 76, 78]. This led to a variation of it referred to as partial co-channel deployment where only

MUEs interfered by SBSs or causing interference to SBSs are allocated to orthogonal resources [78–80]. Enhanced ICIC (e-ICIC) is another form of RP where the MBS vacates some subframes, referred to as almost blank subframes (ABS) for use by SBSs [8, 81]. e-ICIC is usually coupled with offloading of UEs from the MBS through the implementation of cell range expansion in order to reap the benefits of installing the SBSs [82, 83]. e-ICIC requires strict synchronisation between the SBS and the macrocell and exchange of information over the X2 interface which is why it doesn't apply to femtocells [84]. In macrocells with FFR (also a form of RP) implemented, cross-tier interference may be mitigated by the femtocells determining the part of the macrocell where they are located, centre or outer, and then using the partitions that are not utilised by MUEs in that region [85, 86]. A more efficient approach is to use all partitions irrespective of location and only avoid partitions when an MUE is detected in the SBS's coverage area [87]. In [88] cognitive femtocells are offered incentives of extra RBs (with no macrocell interference) in order to serve close-by MUEs. The femtocells accept the offer only if they can maximise their utility while providing the meeting the QoS requirements of the MUE.

RP has also been used for co-tier interference mitigation. Of particular interest in this thesis are RP techniques that require the SBSs to have self organising (SON) capabilities. Several definitions for self organising network (SON) exist in literature [89] however, in this thesis it is defined as follows: self organisation is an adaptive functionality that enables the network to make intelligent changes to the network that maximise some network performance criteria in response to detected changes in the network [89]. Self organisation comprises of self configuration, self optimisation and self healing [89, 90]. In order to self optimise, the SBSs should be aware of the interference environment. This may be done by determining neighbouring BSs. For some techniques, the BS itself determines the neighbouring BSs [91,92] while for others, the BSs use measurement reports (MRs) from their UEs. Whereas fourth generation (4G) systems support both periodic and event triggered MRs, the latter are assumed in this thesis where a BS sends a Radio Resource Control (RRC) connection reconfiguration message to a UE which the UE responds to by searching for the neighbouring BSs, identifying the Physical Layer Cell Identities (PCI) and measuring the Reference Signal Received Power (RSRP) and/or Reference Signal Received Quality (RSRQ) [93].

Several techniques have been proposed in literature that utilise SON capabilities to influence RRM decisions. In [91] periodic measurements are used by each BS to create a

neighbouring BSs list, where neighbouring BSs are those that create interference to the BS that is above a specified threshold. The neighbour list along with RB requirements by the UEs are sent to the central controller which models the resource allocation problem as a max-coloring problem in order to maximise RB reuse. The technique proposed in [92] uses the dsatur graph colouring algorithm to determine BS allocation to the different partitions of the resource. The partitions are fixed by the network and do not vary with interference conditions. Unlike [91], there is a strict allocation order which is the descending order of the number of BSs that are interfered by a BS. Zheng et. al [94] partitions the femtocells using a greedy graph colouring algorithm and then allocates orthogonal resources to the each group. Because the group sizes are not uniform, convex optimisation in used to determine how to partition resources in order to boost throughput [94]. The problem with these techniques is that they determine allocation for SUEs based on interference experienced by the SBS. A simple dynamic resource partitioning technique is presented in [95] where each femtocell is restricted to use of only half the available resources. Using either MRs or femtocell sniffing, the femtocell determines the interference received on each subchannel (the measurement takes fading into account), sorts them in ascending order of received interference and then chooses to transmit on half the subchannels with the least interference. This technique can effectively mitigate co-tier interference in low traffic scenarios [95]. In [96] each BS generates an interference map, referred to as the matrix of conflict (MoC), which is based on MR reports from the SUEs. Unlike algorithms in [92, 94], the partitioning is determined based on the SUEs interference environment which means that two SUEs served by the same SBS may have access to different number of resources because of their different interference environments. Using the aggregated interference map, the central controller orders the SBSs in descending order of the number of interfered UEs plus served UEs (detected UEs) by the SBS. One key difference between this technique and the ones previously presented is that the partitions change with traffic conditions. The partition for each UE is determined by dividing the number of available resources by the maximum number of detected users by the BSs that detect the SUE. Next, the allocation is done to ensure maximum resource reuse by giving allocation priority to resources allocated to non-interfering and non-interfered SUEs. After the first round of allocation, SUEs which are not interfered by any neighbouring BS are allocated to all resources to which they're neither forbidden nor allocated. If any other UEs have resources to which they are neither forbidden nor allocated, the allocation process in the first round will be done again until no allocation opportunities

are left. A hybrid technique that is a mix of the algorithm in [31] and [95] is presented in [32]. The interference map generation and determination of the partition for all UEs is determined in the same manner as [31]. However, instead of determining the allocation for the BSs, the central controller sends the number of RBs accessible by each SUE to its serving SBS. Each SBS will then allocate each SUE to the most appropriate RBs based on the channel quality indicator (CQI) reports. The concept is similar to [95] except that the partitioning is not fixed and varies for each SUE depending on the detected UEs by its SBS and other SBSs that cause it interference. In order for the technique in [32] and [95] to work, the SBSs should carry out resource allocation at random times (allocation should not be synchronised) [32]. In [69], BSs sense the uplink to determine if there are any UEs served by neighbouring BSs in their coverage and by determining how many UEs are in its coverage area, it determines how many RBs to allocate to its UEs. All the RP techniques above, for co-tier interference, assume equal power allocation on all radio resources.

Finally, interference randomisation involves distributed radio resource allocation techniques where each femtocell randomly allocates resources to its UEs over the whole spectrum or part of the spectrum [67, 97]. The allocation is done independent of the impact on neighbouring UEs. This technique is simple and fast but it does not protect the users from co-tier or cross-tier interference as shown in [67]. Random allocation may result in acceptable performance in low traffic scenarios where a UE avoids high interference on an RB by randomly being allocated to another RB. If the randomly allocated RB has less interference, the UE will have a higher data rate. However, in dense networks, this technique suffers from poor performance [73]. Chandrasekhar et al. propose that as the femtocell density increases, the fraction of resources that an FBS allocates randomly should be decreased to minimise interference [74].

3.1.1.1 SON RRM Techniques

In this section, the SON RRM techniques used to evaluate the performance of the algorithms presented in Chapters 3 and 4 are presented. An evaluation of the performance of a selection of SON RRM techniques is presented in [31] and it shows that the MoC-based RRM has the best performance in terms of throughput, minimum data rate and energy efficiency. Of the evaluated techniques, the ones used in this thesis are MoC-based RRM [31] and MoC-based self organising RRM [32]. The techniques are

elaborated in this section.

MoC-based RRM

1. UEs send their MR reports containing the interference measurement to their serving BS. The average received power by the UEs from the neighbouring BSs is determined by the serving BS. The MoC for BS i, $\zeta^i \in \mathbb{R}^{X_i \times F}$, where F is the number of SBSs, is generated by the serving BS as follows

$$\zeta^i_{k,j} = \begin{cases} 1 & \frac{p_{i,k}}{p_{j,k}} < \gamma_{th} \\ 0 & \frac{p_{i,k}}{p_{j,k}} \geq \gamma_{th} \end{cases}. \tag{3.1}$$

2. Therefore, if $\frac{p_{i,k}}{p_{j,k}} < \gamma_{th}$, femtocell j is classified as an interferer to UE k and k cannot reuse a RB allocated to a UE served by BS j. Each SBS sends its local MoC to the central controller which aggregates the local MoCs to form an aggregated MoC $\zeta \in \mathbb{R}^{N_u \times F}$, where N_u are the number of UEs in the network and F are the number of SBSs in the network. It is given by [98]

$$\zeta_{k,j} = \begin{cases} 2, & \text{if } j \text{ serves } k, \\ 1, & \text{if } j \text{ interferes } k, \\ 0, & \text{otherwise.} \end{cases} \tag{3.2}$$

3. The detected UEs of each BS j, the number of UEs that the BS interferes plus the UEs that it serves, are determined from ζ, using

$$u_j = \sum_{k=1}^{N_u} I_{\{\zeta_{k,j}>0\}}, \tag{3.3}$$

where $I_{\{q\}}$ is an indicator function such that $I_{\{q\}} = 1$ when the condition q is true; otherwise, $I_{\{q\}} = 0$. To establish the minimum number of RBs to be allocated to each UE, the minimum RB vector, $\mathcal{S} \in \mathbb{R}^{1 \times F}$, is calculated using

$$\mathcal{S}_j = \left\lfloor N_{rb}/u_j \right\rfloor, \tag{3.4}$$

where $\lfloor \cdot \rfloor$ denotes the floor function. Let $\mathcal{S}^{min} \in \mathbb{R}^{1\times K}$ denote a vector with the minimum number of RBs allocated to all UEs, where K is the number of UEs in the network. For each UE k, the minimum number of RBs is given by

$$\mathcal{S}_{k,min} = \min\left(I_{\{\zeta_{k*}>0\}} \odot \mathcal{S}\right), \tag{3.5}$$

where $\odot$ is the componentwise multiplication of vectors and $I_{\{q\}}$ is an indicator function as previously defined.

4. First the allocation $\overline{A}$, forbidden $\overline{F}$, and priority matrices are initialised as $\mathbf{0}^{N_u\times N_{rb}}$ where $\mathbf{0}$ indicates a zero matrix.

5. For the initial allocation, the allocation order is based on descending order of u_j hence the UEs served by the BS with the highest number of detected UEs are served first. The UEs served by the same BS are randomly ordered for allocation. The first UE k is allocated $\mathcal{S}_k^{min}$ randomly selected RBs and $\overline{A}$ is updated accordingly. Once it is allocated, the $\overline{F}$ is updated to indicate which UEs cannot share RBs with the UE. For the next allocations, the procedure is repeated however allocation priority is given to RBs that have been allocated to non-interfering and non-interfered UEs. The priority matrix helps to identify such RBs.

6. The second round of allocation identifies UEs that are not interfered by any femtocell, referred to as inner UEs, and allocates to them any resources that are forbidden to them.

7. The central controller makes the following check to determine if there are any remaining allocation opportunities

$$\sum_{r=1}^{N_{rb}} \sum_{k=1}^{N_u} \overline{A}_{k,r} + \overline{F}_{k,r} < N_{rb} \times N_u. \tag{3.6}$$

8. If the check in (3.6) is true, then procedure for initial allocation will be repeated until the check in (3.6) is false.

9. The central controller sends $\overline{A}$ to all the BSs which then allocate the RBs to their UEs accordingly.

MoC-based self organising RRM

1. Users send their MR reports containing the interference measurement to their serving BS. The average received power by the UEs from the neighbouring BSs is determined by the serving BS. The MoC for BS i, $\zeta_i \in \mathbb{R}^{\chi_i \times F}$ given by (3.1).

2. Therefore, if $\frac{p_{i,k}}{p_{j,k}} < \gamma_{th}$, femtocell j is classified as an interferer to UE k and k cannot reuse a RB allocated to a UE served by BS j. Each SBS sends its local MoC to the central controller which aggregates the local MoCs to form an aggregated MoC $\zeta \in \mathbb{R}^{N_u \times F}$ given by (3.2).

3. The detected UEs of each BS j, the number of UEs that the BS interferes plus the UEs that it serves, are determined from ζ, using (3.3).

4. The number of RBs to be allocated to each UE are given by (3.5).

5. The central controller then sends, to each BS, the number of RBs to be allocated to each of its UEs.

6. UEs send CQI reports containing the SINR measurement at all RBs to their serving BS [99].

7. The serving BS uses the CQI reports to determine RB allocation. The allocation update process begins after a random waiting period in order to ensure that the BSs do not update allocation at the same time. Based on the CQI, the BS updates the interference matrix, $\overline{I}^{i} \in \mathbb{R}^{\chi_i \times N_{rb}}$ where χ_i is the number of UEs served by the BS i and N_{rb} is the number of RBs for allocation in the network. The new RB allocation for BS i is computed based on the following criteria:

$$\min \sum_{k=1}^{\chi_i} \sum_{r=1}^{N_{rb}} \overline{I}^{i}_{k,r} \overline{A}^{i}_{k,r} \tag{3.7}$$

subject to

$$\sum_{k=1}^{\chi_i} \overline{A}^{i}_{k,r} \leq 1, \qquad \forall r$$

$$\overline{A}^{i}_{k,r} \in \{0, 1\} \qquad \forall k, r$$

$$\sum_{k=1}^{\chi_i} \sum_{r=1}^{N_{rb}} \overline{A}^{i}_{k,r} = \mathcal{S}_{k,min}$$

where $\mathcal{S}_{k,min}$ is the number of RBs determined by (3.5), $\overline{A}^{i}_{k,r}$ is the allocation of BS i for UE k at RB r and it is equal to 1 if UE k is allocated to RB r and equal to 0 otherwise.

3.1.2 Power Allocation in HetNets

Power control applied in a centralised or distributed manner [67] is a key interference avoidance approach [100]. Centralised power control requires a central controller to determine transmission power based on the interference environment while distributed power control involves each femtocell determining the transmission power based on measurement results and possibly preset system parameters [57]. In [100], the authors advocate for femtocells to determine their distance from the nearest macrocell and use it to adjust their power to avoid interference to macrocell users and to mitigate co-tier interference. However, if the MBS is far and the MUEs are close to the femtocell, they will not be protected from interference.

One popular technique used for distributed power control is game theory [57, 101]. In [102], a distributed power control non-cooperative game is developed to control interference between femtocells. A utility function of the game is derived using fairness and interference to other femtocells which results in less power for low load femtocells compared to high load femtocells. Whereas the technique ensures fairness in power allocation between femtocells, it does not necessarily ensure that the femtocell users get the required power to transmit which may leave the UE served by the low load femtocell in outage. In [103], a distributed non-cooperative game is developed with a utility function that minimises interference to the macrocell users. Whereas the cross-tier interference is managed, the co-tier interference is ignored and the femtocells are not guaranteed any spectrum in case of a high user density macrocell. To mitigate cross-tier interference and give priority to MBS users using game theory, the use of a Stackleberg game is perhaps the most appropriate and popular game used because it models the MBS as a leader, with priority over the femtocell, modelled as the follower [77, 104–106]. The MBS therefore sets the prices for femtocells based on the interference it receives from them, high price for high interference and low for low interference. If the price is high, the femtocell lowers its power in order to maximise its utility. The power control game is solved when a Nash equilibrium is reached i.e. for all involved players, a change in their power would result in a reduction in their utility.

Different variations of the Stackleberg game are used in [105–107] to solve the power allocation problem however it is clear that finding the solution to the game is complex and some algorithms take long to converge. In [108], the uncertainty in channel information is added to the Stackleberg game formulation for a more realistic model. A suboptimal approach suggested in [107] involves uniform pricing for all femtocells which results in a simpler and faster solution to the game. In [106] a distributed approach is taken to determining the power transmitted whereby the femtocells compete in an uncooperative game, each with the goal of maximising their capacity however the computational burden is huge and a suboptimal approach which uses Lagrangian techniques is used to solve the game. Notably [77] recommends that in high interference scenarios, dedicated spectrum should be used by the MBS to avoid interference because the Stackleberg game solution may not provide adequate protection to MBS users. Most game theoretic algorithms require full channel information to determine the optimal power allocation. Further, they require a number of iterations to converge to the optimal solution which may not be ideal for real time applications.

3.1.2.1 Pilot Power Allocation in HetNets

In the literature, research related to pilot signals in heterogeneous networks is focused on load balancing [33, 34], cross tier interference mitigation [35, 36] and cell coverage [109, 110]. The techniques are implemented in third generation (3G) [34, 36, 109, 110] and 4G networks [33, 35]. A distributed algorithm that optimises femtocell coverage based on the deployment of femtocells is presented in [109]. By adjusting pilot power, it ensures that the femtocell maximises its coverage when there are no neighbouring BSs and minimises the coverage area when there is overlap between two or more femtocells. Cluster-heads use a genetic algorithm to determine the optimal pilot power for the BSs in their clusters in order to balance the load and reduce both coverage gaps and overlap in [110]. The results show significant improvement in the SINR distribution of UEs compared to the case where all femtocells transmit the same pilot power. However, the SINR performance degrades significantly with increase in the number of femtocells in the network. In [33] a distributed algorithm to adapt the pilot power of femtocells so as to offload UEs from the MBS subject to ability to meet the UEs' QoS requirements is presented. The algorithm improves energy efficiency by switching the femtocells without associated UEs to Listen mode until they detect an MUE in their vicinity. The femtocell then sets an initial pilot power to ensure that a UE at a

specific reference point is able to receive the same power from the macrocell as from the femtocell. If it is handed over to the femtocell, its SINR is determined and if it is less than the required threshold, the femtocell reduces power until the UE is handed back to the MBS. Both co-tier and cross-tier interference are reduced by femtocells reducing power. Results show that in high load scenarios, the femtocells offload a significant number of UEs from the MBS [33]. In [36], downlink cross tier interference is minimised by setting the pilot power to ensure that the femtocell UEs receive the same power from the macrocell as from the femtocell at 10 m away from the femtocell. This power setting fixes the coverage at 10 m. Co-tier interference is not addressed in this work because the femtocells are placed far apart from each other. In [35], the authors advocate for adaptive pilot power setting based on the location of the furthest UE. Therefore, each femtocell's coverage radius depends on how far it is from its furthest UE. Their results show that smaller coverage radius results in higher MUE SINRs because of reduced downlink interference while resulting in lower femtocell UE (FUE) SINRs due to reduced transmit power. The reverse is true as well. Therefore, when the coverage radius is low, the cross-tier performance of the algorithm in [35] is better than that of the approach in [36]. The authors in [35] conclude that gains in co-tier SINR from pilot power allocation are diminished when FUEs are located close to interfering femtocells. In [34] a distributed algorithm to adjust the pilot power of femtocells in order to balance the load of the femtocells while minimising coverage holes is presented. Overloaded femtocells reduce their pilot power in order to handover some UEs while underloaded femtocells in the vicinity of the overloaded femtocells increase power in order to provide coverage for the dropped UEs. In order to avoid coverage holes, a femtocell only reduces its pilot power when it is certain there is a neighbouring femtocell that will associate with the UE it drops. The focus of the paper is load balancing through modification of the pilot channel power and it does not address the impact on co-tier interference.

Techniques [33, 35, 36] are ineffective in scenarios where co-tier interference is the dominant form of interference in a given scenario i.e. when the macrocell is far away or when it is using orthogonal resources. The techniques that focus on load balancing provide limited gains when the network is underloaded [34, 110, 111].

This chapter presents an algorithm for reducing the reference signal power and ultimately, the transmit power, of some or all SBSs in the network in order to modify the interference map and ultimately improve the throughput of established RP SON RRM

techniques. The MRs used by SON RRM techniques contain the following information which is used for interference map generation: PCI, RSRP and RSRQ. RSRP is defined as " the linear average over the power contributions (in [W]) of the resource elements that carry cell-specific reference signals (RSs) within the considered measurement frequency bandwidth" in [112] while RSRQ is given by:

$$(N_{rb} \times \text{RSRP}/\text{RSSI}), \tag{3.8}$$

where N_{rb} is the number of RBs of the carrier Received Signal Strength Indicator (RSSI) measurement bandwidth [112]. Carrier RSSI is the average of the total received power (in [W]) in OFDM symbols containing RS for antenna port 0 over N_{rb} RBs by the UE from all sources including co-channel serving and non-serving cells, adjacent channel interference and thermal noise [112]. RSs, also known as pilot signals, are inserted in the RBs to enable channel estimation [28, 43]. Therefore, by adjusting the RS power, the RSRP and RSRQ values change which will then cause the generated interference map to change. This is the basis of the power allocation technique presented in this chapter. Unlike conventional power allocation techniques [113, 114] where the power is varied on each RB depending on the channel gain, our technique allocates the same power to all RBs and it is equal to the power allocated to the reference signals.

3.2 System Model

The downlink of a perfectly synchronous LTE/TDD SBS network is considered. We assume that there are N_{rb} RBs for data transmission in each time slot. Further, a UE is allocated an RB for the duration of one transmission time interval (TTI). The allocation is done on a frame by frame basis and because there are 10 TTIs in a frame, $N_T = N_{rb} \times 10$ RBs are allocated in each allocation cycle [115]. This implies that there's scheduling of resources in each TTI. Therefore, if $N_{rb} = 50$, a UE that is allocated at RB 55 will transmit at RB 5 in the second TTI. Let $\mathcal{F}$ and $\mathcal{K}$ denote the deployed SBSs and UEs in the network such that $\mathcal{F} = \{1, 2, \ldots, i, \ldots j, \ldots, \text{F}\}$ and $\mathcal{K} = \{1, 2, \ldots, k, \ldots, N_u\}$. Let $i \in \mathcal{F}$ denote a random SBS in the set, which serves a UE $k \in \mathcal{K}$ and $j \in \mathcal{F}$ denote a neighbouring SBS to i. Let the number of UEs served by BS i be denoted by $\mathcal{X}_i$. The SBSs are connected to a central controller, e.g. if the SBSs are femtocells, they connect to the FMS with operation and management (OAM)

functionality, in order to manage resource allocation. It is assumed that there is no cross-tier interference. The SBSs are assumed to have SON capabilities and as such are able to find neighbouring SBSs. Therefore, all SBSs and UEs have a neighbouring SBSs list in their memory. Finally the SBSs are assumed to operate in open access mode. The SINR of user k at RB r is given by

$$\Upsilon_{k,r} = \frac{p_{i,k,r} h_{i,k,r}}{\sum_{j=1, j\neq i}^{F} p_{j,k,r} h_{j,k,r} \widetilde{A}_{j,r} + N_0 \mathcal{B}_{rb}} \tag{3.9}$$

where $p_{i,k,r}$ and $h_{i,k,r}$ are the received power and channel gain from BS i to its UE k on RB r, respectively, $p_{j,k,r}$ and $h_{j,k,r}$ are the interference power and channel gain from BS j to UE k on RB r, respectively, $\mathcal{B}_{rb}$ is the bandwidth of an RB and $\widetilde{A}_{j,r} \in \{0, 1\}$ is an indicator which is equal to 1 if a UE served by BS j is allocated to RB r and equal to 0 otherwise. The data rate of a UE k at RB r is given by

$$\mathcal{R}_{k,r} = \frac{\log_2 (1 + \Upsilon_{k,r})}{T_0}, \tag{3.10}$$

where T_0 is the OFDM symbol duration. The throughput, $\mathcal{R}_{tot}$, is given by

$$\mathcal{R}_{tot} = \frac{\sum_{k=1}^{N_u} \sum_{r=1}^{N_T} \mathcal{R}_{k,r}}{N_{sf}} \tag{3.11}$$

where N_{sf} is the number of sub frames in a frame.

3.3 RB Allocation

RB allocation is done using two techniques: MoC based RRM (M-RRM) and MoC-based self organising RRM (MS-RRM), both of which are described in 3.1.1.1.

3.4 Power Allocation Algorithm

If we denote the minimum data rate attained by the RRM technique before RS power adaptation by $\mathcal{R}_{min}$, then the sum rate maximisation problem we intend to solve is

$$\text{maximise}\left(\sum_{k=1}^{N_u}\sum_{r=1}^{N_T}\overline{A}_{k,r}\mathcal{R}_{k,r}\right) \tag{3.12}$$

subject to

$$\begin{aligned}\textstyle\sum_{k=1}^{\chi_i}\overline{A}_{k,r} \leq 1, &\qquad \forall r, k \in \chi_i\\ \overline{A}_{k,r} \in \{0,1\}, &\qquad \forall k, n\\ \textstyle\sum_{r=1}^{N_T}\mathcal{R}_{k,r}\overline{A}_{k,r} \geq \mathcal{R}_{min}, &\qquad \forall k.\end{aligned}$$

The first constraint means that each BS i will allocate at most one UE to an RB. The second constraint means that the elements, $\overline{A}_{k,r}$, of $\overline{A}$, the allocation matrix, can only take on two values: either 1 to indicate that UE k is allocated to RB r or 0 to indicate that it is not. The final constraint means that the minimum data rate obtained after RS power adaptation should be greater than or equal to $\mathcal{R}_{min}$. The third constraint ensures that the data rate of the most interfered UEs is not compromised in order to boost the data rate of other UEs. In order to solve the sum rate maximisation problem for MoC based RRM, we propose a two stage algorithm. Let $\mathcal{U} = \left\{u_1, u_2 \ldots, u_j, \ldots, u_F\right\}$ denote the number of detected UEs from the MoC. It was shown in [98] that the minimum data rate of the network is dependent on $u_{max} = \max\{\mathcal{U}\}$. This is because u_{max} determines the least RBs allocated to a UE as shown by (3.4) and (3.5). The first stage of the algorithm is concerned with reducing u_{max} in order to increase the minimum subcarriers allocated to the UEs. The second stage of algorithm reduces the RS power of SBSs with inner UEs (usually have high SINR) to enable them to reuse RBs that they were initially forbidden to reuse. The first stage is referred to as detected UE minimisation while the second stage is referred to as inner UE throughput maximisation.

3.4.1 Detected UE Minimisation

Let u_{init} denote the value of u_{max} when all BSs transmit the same power and $\mathbb{P} \in \mathbb{R}^{1\times F}$ denote the RS power for each of the SBSs. It is the same for all SBSs at the start

of the algorithm. Moreover, $\overline{\mathcal{U}}$ denotes the number of detected UEs in the updated MoC after each iteration of RS power modification, $\mathcal{F}_{max}$ denotes the SBSs with u_{max} detected UEs. For each SBS j, $u_j = \chi_j + u_{cj}$, where u_{cj} denotes the UEs interfered by SBS j and it is given by

$$u_{cj} = \sum_{k=1}^{N_u} I_{\{\zeta_{k,j}=1\}}, \tag{3.13}$$

where $I_{\{q\}}$ is an indicator function such that $I_{\{q\}} = 1$ when the condition q is true; otherwise, $I_{\{q\}} = 0$. Detected UE minimisation is performed using Algorithm 1. SBSs that reduce their RS power are denoted by $\mathcal{F}_{pc}$ and in the first iteration, $\mathcal{F}_{pc} = \mathcal{F}_{max}$. All SBSs in $\mathcal{F}_{pc}$ reduce their pilot power in steps of $\triangle p$ dB. Each time u_{max} reduces, the transmit power of the SBSs is updated in $\mathbb{P}$. After a power reduction iteration, any changes in the MoCs of the BSs will be sent to the central controller so that ζ and $\overline{\mathcal{U}}$ are updated. UEs served by SBSs in set $\mathcal{F}_{pc}$ have reduced SIRs with respect to other SBSs. When the SIR, $\gamma_{k,j}$, of UE k with respect to SBS j drops below γ_{th} then the UE becomes a detected UE of SBS j. If a UE served by SBS $i \in \mathcal{F}_{max}$ becomes a detected UE of SBS $j \notin \mathcal{F}_{max}$ making $u_j \geq u_{max}$ and $u_i < u_{max}$ then SBS $j \in \mathcal{F}_{max}$ and $i \notin \mathcal{F}_{max}$. If in the next power reduction iteration, $u_j < u_{max}$ and $u_i \geq u_{max}$ then, SBS j is considered conditional member of $\mathcal{F}_{max}$ for as long as SBS $i \in \mathcal{F}_{max}$. Therefore, if SBS i ceases to be a member of $\mathcal{F}_{max}$ because power reduction makes $u_i < u_{max}$, then SBS j will cease to be a conditional member of $\mathcal{F}_{max}$. We denote SBSs such as SBS j as $\mathcal{F}_{dep}$. After each power reduction iteration, changes in ζ and $\overline{\mathcal{U}}$ are used to update $\mathcal{F}_{dep}$. If $\mathcal{F}_{dep} \neq \emptyset$, then in the next iteration $\mathcal{F}_{pc}$ is given by

$$\mathcal{F}_{pc} = \mathcal{F}_{max} \cup \mathcal{F}_{dep}. \tag{3.14}$$

After each power reduction iteration, the decision to stop or continue is based on the following three exit criteria:

$$\begin{aligned} \mathcal{F}_{pc} &= \mathcal{F}, \\ \mathcal{K}_c &\subseteq \mathcal{K}_s, \\ \left\{ j \mid u_j = \chi_j \wedge j \in \mathcal{F}_{max} \right\} &= \emptyset, \end{aligned} \tag{3.15}$$

where $\mathcal{K}_s$ denotes all UEs served by SBSs in $\mathcal{F}_{pc}$ and $\mathcal{K}_c$ denotes all UEs interfered by SBSs in $\mathcal{F}_{pc}$. If any of the exit criteria is true, then the power reduction should stop. If the first criterion is true, it means that all SBSs in the network are simultaneously reducing power by $\triangle p$ which means that u_{max} cannot be reduced any further. If the second criterion is true, it means that all the UEs that are detected by the SBSs in $\mathcal{F}_{pc}$ are served by the SBSs in $\mathcal{F}_{pc}$ which means u_{max} cannot be reduced. Finally, if the third criterion is true, it means that for one or more SBSs in $\mathcal{F}_{max}$, the only detected UEs are those that it serves. Such a SBS will require to reduce its power by more than γ_{th} in order to reduce u_{max}. Once at least one of the exit criteria is met, then $P_{t,i} = \mathbb{P}_i$ for $i = \{1, 2, \ldots, F\}$.

3.4.2 Inner UE Throughput Maximisation

Inner UE throughput maximisation boosts system throughput by letting inner UEs access more RBs and is performed using Algorithm 2. Given $\overline{A}$, the total RB utilisation $S_t = \sum_{k=1}^{N_u} \sum_{r=1}^{N_T} \overline{A}_{k,r}$. The number of inner UEs served by SBS j is denoted by $u_{j,}^{in}$. Let $\mathcal{F}_{in}$ denote SBSs with inner UEs. Each SBS j, such that $j \in \mathcal{F}_{in}$, will take a turn to reduce pilot power in steps of $\triangle p$ in order to reduce u_j. Each time u_j reduces RB allocation is done and S_t and the total power reduction are saved in $\mathcal{T}$ and $\mathcal{W}$ respectively, where $\mathcal{T}$ and $\mathcal{W}$ are vectors that track RB utilisation and power reduction. The pilot power is reduced until at least one of the following exit criteria is met

$$\begin{aligned} u_{max} &= u_{init}, \\ u_j &= \chi_j, \\ u_j^{in} &= \emptyset, \end{aligned} \tag{3.16}$$

where j denotes the SBS that is currently reducing power. The first exit criterion means that the u_{max} is equal to the maximum number of detected UEs before the detected UE minimisation algorithm started. When this happens the worst off UEs will get the same number of RBs as before but with higher interference from inner UEs that have gained access to their RBs. This is why the power reduction is stopped once criterion 1 is met.

Algorithm 1 Detected UE Minimisation

if none of the exit criteria are met **then**
    flag = 1
else
    flag = 0
end if
$\mathcal{F}_{dep} = \emptyset$
while flag **do**
    $\mathcal{F}_{pc} = \mathcal{F}_{max} \cup \mathcal{F}_{dep}$
    for $j = 1 : F$ **do**
        if $j \in \mathcal{F}_{pc}$ **then**
            $P_{t,j} = P_{t,j} - \triangle p$
        end if
    end for
    update ζ using new measurement report
    generate $\overline{\mathcal{U}}$ using (3.3)
    if $\max(\overline{\mathcal{U}}) \geq u_{max}$ **then**
        update $\mathcal{F}_{max}$ by $\{j \in \mathcal{F}_{max} | \overline{\mathcal{U}}_j \geq u_{max}\}$
        update $\mathcal{K}_s$ and $\mathcal{K}_c$ based on the new ζ
        update $\mathcal{F}_{dep}$
        $\mathcal{F}_{pc} = \mathcal{F}_{max} \cup \mathcal{F}_{dep}$
        if none of the exit criteria are met **then**
            $\mathcal{U} = \overline{\mathcal{U}}$
        else
            flag = 0
        end if
    else
        $u_{max} = \max(\overline{\mathcal{U}})$ i.e. u_{max} is reduced
        $\mathcal{U} = \overline{\mathcal{U}}$
    end if
    if any of the exit criteria are met **then**
        flag = 0
    end if
end while

Algorithm 2 Inner UE throughput maximisation

for $j = 1 : F$ **do**
  $\mathcal{T} = \mathbf{0}^{1\times u_m}$ // initialise RB counter
  $\mathcal{W} = \mathbf{0}^{1\times u_m}$ // initialise power tracker
  $\vartheta = 1; \varrho = 1$ // initialise counters
  $\mathcal{T}_{(\vartheta)} = S_t$ // initialise total RB utilisation
  if $j \in \mathcal{F}_{in}$ **then**
    while none of the exit criteria are met **do**
      $\varrho = \varrho + 1$
      reduce pilot power of j by $\triangle p$
      update ζ using new measurement report
      update $\overline{\mathcal{U}}$
      if $\max(\overline{\mathcal{U}}) < u_{init}$ **then**
        if $\overline{\mathcal{U}}_j < \mathcal{U}_j$ **then**
          $\vartheta = \vartheta + 1$
          allocate RBs and calculate S_t
          $\mathcal{T}(\vartheta) = S_t$ // update total RB utilisation
          $\mathcal{W}(\vartheta) = \varrho \times \triangle p$ // update power reduction
          $u_{max} = \max(\overline{\mathcal{U}})$
          $\mathcal{U} = \overline{\mathcal{U}}$
        else
          reduce pilot power of j by $\triangle p$
        end if
      else
        find ϑ' such that $\mathcal{T}_{\vartheta'} = \max(\mathcal{T})$
        $\mathbb{P}_j = \mathbb{P}_j - \mathcal{W}_{\vartheta'}$
        exit while loop
      end if
    end while
  end if
end for

Criterion 2 is the same as criterion 3 in (3.15) while criterion 3 means that because of power reduction the BS has no inner UEs hence any further power reduction will not increase RB utilisation. Once any of the exit criterion are met, the central controller

will find the position $\vartheta^{'}$ in $\mathcal{T}$ where $\mathcal{T} = \max(\mathcal{T})$ i.e. where maximum RB utilisation is achieved. $\mathcal{W}_{\vartheta'}$ is therefore the amount by which the transmit power should be reduced to maximise RB utilisation. The final transmit power of SBS $j \in \mathcal{F}_{in}$ is calculated by $\mathbb{P}_j - \mathcal{W}_{\vartheta'}$.

3.4.3 SIR Difference Matrix

The number of iterations required by the proposed power allocation technique is high especially for small $\triangle p$. One work around that avoids sending the MoC updates is for the SBSs to send a matrix with the SIR difference, $\gamma_{k,j} - \gamma_{th}$, instead, which is denoted by $\mathfrak{I}^i \in \mathbb{R}^{\chi_i \times F}$ for BS i. The aggregated MoC is generated as follows,

$$\zeta_{k,j} = \begin{cases} 2, & \mathfrak{I}_{k,j} = -\gamma_{th}, \\ 1, & \mathfrak{I}_{k,j} < 0 \wedge \mathfrak{I}_{k,j} > -\gamma_{th}, \\ 0, & \mathfrak{I}_{k,j} > 0. \end{cases} \tag{3.17}$$

Algorithms 1 and 2 are then implemented by the central controller and with each power reduction step, $\triangle p$, $\mathfrak{I}$ is updated by

$$\mathfrak{I}_{k,j} = \mathfrak{I}_{k,j} - \triangle p \quad k \in \mathcal{X}_i, j \neq i \tag{3.18}$$

for UEs served by power reducing SBS i while for other UEs it is update by

$$\mathfrak{I}_{k,i} = \mathfrak{I}_{k,i} + \triangle p \quad k \notin \mathcal{X}_i \ . \tag{3.19}$$

Therefore, power reduction iterations will be done by the central controller without the need for back and forth MoC updates.

3.4.3.1 Signalling Overhead - SIR Matrix and MoC Comparison

The total amount of information exchange between the central controller and BS i for DMIM using the MoC is given by

$$b_{in,\zeta} = \left(\chi_i \cdot b_\zeta \cdot F + b_{\triangle p}\right) \cdot N_{it} \tag{3.20}$$

while the information exchange between the central controller and the BSs MMIM using the SIR difference matrix is given by

$$b_{in,\mathfrak{I}} = \chi_i \cdot b_{\mathfrak{I}} \cdot F + b_{\triangle p} \tag{3.21}$$

where b_ζ is the number of bits required to transmit an element in the MoC, $b_{\triangle p}$ is the number of bits for the power reduction instruction, $b_{\mathfrak{I}}$ is the number of bits required to transmit an element in the SIR difference matrix and N_{it} is the number of iterations. There are two bits for each element in the MoC because it has only 3 possible values hence $b_\zeta = 2$. Let us assume that the SIR difference values can be as large 128 dB and therefore require up to 7 bits meaning that $b_{\mathfrak{I}} = 7$. Therefore, if N_{it} is large, $b_{in,\zeta}$ may be significantly larger than $b_{in,\mathfrak{I}}$.

3.4.4 Convergence Analysis

In this section, $\wedge$ denotes 'AND' while $\vee$ denotes 'OR'. To prove convergence, the approach used in [116] is adopted.

3.4.4.1 Detected UE Minimisation

The goal of Algorithm 1 is to reduce the maximum number of detected UEs by the SBSs in the network. Convergence is reached when further reduction in power by the $\mathcal{F}_{pc}$ does not result in reduction of detected UEs. The exit criteria in (3.15) are indications of convergence and are explained here. The first condition states that convergence is reached when all SBSs in the network are elements of $\mathcal{F}_{pc}$. In order to reduce the number of detected UEs by SBSs in $\mathcal{F}_{pc}$, the transmit power of the SBSs is reduced by

$\triangle p$. For simplicity, suppose that the network has two SBSs serving one UE each and that SBS 1 is an element of $\mathcal{F}_{pc}$ while SBS 2 is not. This implies that $\gamma_1 \geq \gamma_{th} \wedge \gamma_2 < \gamma_{th}$ i.e. BS 1 has 2 detected UEs while BS 2 has 1 detected UE. From (3.9), it can be seen that by SBS 1 reducing power by $\triangle p$, the SIR of UE 1 reduces while the SIR of UE 2 increases. There are four possible outcomes from the power reduction. Outcome 1: $\gamma_1 \geq \gamma_{th} \wedge \gamma_2 < \gamma_{th}$ i.e. no change in detected UEs; Outcome 2: $\gamma_1 < \gamma_{th} \wedge \gamma_2 < \gamma_{th}$ i.e. increasing the number of detected UEs by SBS 2 to 2; Outcome 3: $\gamma_1 \geq \gamma_{th} \wedge \gamma_2 \geq \gamma_{th}$ i.e. reducing the number of detected UEs by BS 1 to 1; Outcome 4: $\gamma_1 < \gamma_{th} \wedge \gamma_2 \geq \gamma_{th}$ i.e. increasing the number of detected UEs by BS 1 to 1 while increasing the number of detected UEs by BS 2 to 2. Outcome 1 implies that further power reduction is required by BS 1 until one of the other outcomes is reached. Outcome 4 means that SBS 1 is an element of $\mathcal{F}_{dep}$ which implies both SBSs are elements $\mathcal{F}_{pc}$. Although Outcomes 2 and 3 also mean that both SBSs are elements of $\mathcal{F}_{pc}$, Outcome 2 is more desirable because it reduces the number of detected UEs. When both SBSs are elements of $\mathcal{F}_{pc}$ as stipulated in the first exit criteria in (3.15), further reduction in transmit power will result in no change in detected UEs. This is because neither γ_1 nor γ_2 changes because the change in interference and transmit power are the same for both UEs in (3.9). The noise power is ignored because it is much less than both the desired and interference power which is usually the case in small cell networks because of proximity to both the serving and interfering SBS. In case of Outcomes 3 and 4, it is best to increase the power of BS 1 to restore the SIRs that were obtained before power reduction started.

The second exit criteria in (3.15) states that if all the detected UEs by the SBSs in $\mathcal{F}_{pc}$ are served by SBSs in $\mathcal{F}_{pc}$ then the maximum number of detected UEs cannot be reduced. This is similar to the previous criteria because it means that for all detected UEs in this case, the power reduction by SBSs in $\mathcal{F}_{pc}$ does not alter their SIRs because the reduction in interference and serving SBS power are the same and they cancel out. Therefore, in such a scenario it is useless to increase power reduction as it will only reduce the SIRs of the UEs with respect to other SBSs that are not elements of $\mathcal{F}_{pc}$. This may result in increased detected UEs by these SBSs which would result in reduced reuse without a reduction in u_{max}.

The third exit criteria in (3.15) states that if the detected UEs by one or more of the SBSs in $\mathcal{F}_{pc}$ are only UEs served by the SBS then further reduction in power will not result in a reduction of u_{max}. This is because for such an SBS to reduce u_{max}, it has to reduce its transmit power by at least $2 \times \gamma_{th}$ in order to handover the UE and reduce

interference to it by γ_{th}. This is a huge reduction in power and is usually unattainable when more than two SBSs are involved. In the event that it is possible, it results in a large number of iterations to achieve the desired reduction in u_{max}. This is why it is included as the third exit criteria.

The above analysis shows that when Algorithm 1 is implemented, it always converges because one or more of the exit criteria will be met. The above analysis can easily be extended to include more SBSs and more UEs per SBS and similar outcomes would be achieved and the algorithm would still converge if the exit criteria are applied.

3.4.4.2 Inner UE Throughput Maximisation

The goal of Algorithm 2 is to reduce the number of detected UEs by SBSs with inner UEs. This effectively increases the number of RBs the inner UEs can access while it increases interference for the UEs already allocated to these UEs. In order to avoid the violation of the minimum data rate performance, it is necessary to ensure that the minimum number of RBs allocated to UEs after both Algorithms 1 and 2 is higher than RBs allocated after both algorithms are implemented. According to (3.4), the minimum number of RBs before the algorithms are implemented is given by

$$N_{rb,min} = \lfloor N_{rb}/u_{init} \rfloor, \tag{3.22}$$

Therefore if $u_{max} \geq u_{init}$ after Algorithm 2, it means that the minimum data rate will be less because of higher interference from the inner UEs coupled with either the same or reduced RBs. For simplicity, consider a scenario with four SBSs each serving one UE. Assume that the current MoC is given by

$$\zeta = \begin{matrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 1 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \end{matrix} \tag{3.23}$$

From (3.23) it can be seen that $u_{max} = 3$ and SBS 1 has an inner UE. Assume that

$u_{init} = 4$. Let us denote the SIR of the UEs by $\gamma_1, \gamma_2, \gamma_3, \gamma_4$. In order to increase the reuse of UE 1, SBS 1 needs to reduce transmit power in order to increase the either $\gamma_{3,1}$ or $\gamma_{4,1}$ or both so that they are greater or equal to γ_{th}. There are three possible outcomes. Outcome 1 is that $\gamma_{1,2} \vee \gamma_{1,3} \vee \gamma_{1,4} < \gamma_{th}$ i.e. meaning that UE 1 is no longer an inner UE. This result reduces the number of UEs that UE 1 can share RBs with and not favourable for reuse increase. This is exit criteria 3. In particular if $\gamma_{1,3} < \gamma_{th}$ then $u_{max} = 4$. This result is not desirable because it reduces $N_{rb,min}$. This is captured by exit criteria 1. Also if $\gamma_{1,2} < \gamma_{th}$ it means that UE 1 can no longer share RBs with UE 2 which it could share resources with before the inner UE maximisation and this is not desirable. Outcome 2 is that $\gamma_{3,1} \geq \gamma_{th} \vee \gamma_{4,1} \geq \gamma_{th}$. This outcome is desirable because it increases the number of RBs that UE 1 can access. However, once this result is attained SBS 1 still reduces power further in attempt to reduce the detected UEs further and in the process it ends up with either Outcome 1 or Outcome 3. Outcome 3 is that $\gamma_{3,1} \geq \gamma_{th} \wedge \gamma_{4,1} \geq \gamma_{th}$. This outcome is desirable because it means that UE 1 can share RBs with all UEs served by neighbouring SBSs. Further power reduction at this stage is futile because the only detected UE is the one served by SBS 1. This is exit criteria 2.

Therefore when the SBS with an inner UE iteratively reduces power, at least one of the exit criteria will be met and this shows that the algorithm always converges. The above analysis can easily be extended to scenarios with more SBSs and more UEs per SBS and the same exit criteria would ensure that ther algorithm converges.

3.5 Performance Analysis

The probability density function (PDF) of the downlink SIR conditioned on the UE location in an $\ell \times \ell$ building was derived in [31]. This PDF was used to derive relationship between the minimum data rate attained and SIR threshold when MoC-based RRM is used.

3.5.1 SIR Model

The distance between UE k and its serving SBS i is denoted by $d_{k,i}$, the distance between the UE and a neighbouring SBS j is denoted by $d_{k,j}$, and the indoor pathloss exponent is denoted by α. Assuming free space path loss (refer to (2.2)), ignoring shadowing and fading, and equal transmit power by all SBSs, the SIR of UE k with respect to SBS j is given by

$$\gamma_{k,j} = \frac{p_{i,k}}{p_{j,k}} = \left(\frac{d_{k,j}}{d_{k,i}}\right)^{\alpha}. \tag{3.24}$$

From (3.24), for UE k to be classified as undetected by SBS j the following condition should be met

$$\gamma_{k,j} \geq \gamma_{th} \tag{3.25}$$

Combining (3.24) and (3.25) gives the following condition for UE k not to be a detected UE of BS j

$$\frac{d_{k,i}}{d_{k,j}} \leq \left(\frac{1}{\gamma_{th}}\right)^{\frac{1}{\alpha}}. \tag{3.26}$$

The ratio $\frac{d_{k,i}}{d_{k,j}}$ is referred to as the distance ratio between k, i and j and is denoted by D_r. Therefore UE k is a detected UE of BS j if $\left(\frac{1}{\gamma_{th}}\right)^{\frac{1}{\alpha}} \leq D_r \leq 1$ and it is not detected when $0 \leq D_r \leq \left(\frac{1}{\gamma_{th}}\right)^{\frac{1}{\alpha}}$. However, if $P_{t,i}$ and $P_{t,j}$ are different, the condition (3.26) becomes

$$\frac{d_{k,i}}{d_{k,j}} \leq \left(\frac{P_{t,i}}{P_{t,j} \cdot \gamma_{th}}\right)^{\frac{1}{\alpha}}. \tag{3.27}$$

From (3.27), it is clear that decreasing the transmit power of BS j will increase the distance ratio at which UE k will be detected by BS j. On the other hand, decreasing the transmit power of BS i will decrease the distance ratio at which UE k is detected by BS i. Therefore UE k is a detected UE of BS j if $\left(\frac{P_{t,i}}{P_{t,j} \cdot \gamma_{th}}\right)^{\frac{1}{\alpha}} \leq D_r \leq 1$ and it is not detected when $0 \leq D_r \leq \left(\frac{P_{t,i}}{P_{t,j} \cdot \gamma_{th}}\right)^{\frac{1}{\alpha}}$.

3.5.2 Distance Ratio Analysis

Consider an enterprise deployment model with an $\ell \times \ell$ building and two BSs randomly and uniformly placed in the building. Assuming that the BSs transmit equal power and that the BSs are positioned on the x,y-plane with BS1 located at $(0, 0)$, and BS2 located at (x, y) and a UE k is randomly and uniformly distributed around BS1 within a distance d_{max}, then the PDF of D_r is given by [31]

$$f_{D_r}(d_r) = \begin{cases} \frac{4\ell}{3d_{max}} \arcsin(1) + \frac{4\ell}{10d_{max}} - \frac{2\ell}{d_{max}}, & 0 < d_r < \frac{d_{max}}{\ell}, \\ \frac{4d_{max}^2}{3\ell^2 d_r^3} \arcsin(1) + \frac{4d_{max}^4}{10\ell^4 d_r^5} - \frac{2d_{max}^3}{\ell^3 d_r^4}, & d_r \geq \frac{d_{max}}{\ell}, \\ 0 & \text{elsewhere.} \end{cases} \tag{3.28}$$

3.5.3 Minimum Data Rate Analysis

In the enterprise scenario with F SBSs and χ_j UEs served per BS, the number of detected UEs by a given BS is given by [31]

$$u_j = \chi_j \left[1 + (F-1) \left\{ 1 - F_{D_r} \left(\left(\frac{1}{\gamma_{th}} \right)^{\frac{1}{\alpha}} \right) \right\} \right] \tag{3.29}$$

where $F_{D_r}(d_r)$ is the cumulative distribution function (CDF) of $f_{D_r}(d_r)$. The minimum number of RBs, $N_{rb,min}$, that can be allocated to a UE using MoC based RRM is lower bounded by

$$N_{rb,min} \geq \left\{ \frac{N_{rb}}{u_j} = \frac{N_{rb}}{\chi_j \left[1 + (F-1) \left\{ 1 - F_{D_r} \left(\left(\frac{1}{\gamma_{th}} \right)^{\frac{1}{\alpha}} \right) \right\} \right]} \right\} \tag{3.30}$$

Assuming an AWGN channel, the data rate of UE k is lower bounded by

$$\mathcal{R}_k \geq N_{rb,min} \mathcal{B}_{rb} \log_2 (1 + \Upsilon_{min}) \tag{3.31}$$

where Υ_{min} is the minimum SINR [31]. Assuming that $p_{i,k} \gg N_0\mathcal{B}_{rb}$,

$$\Upsilon_k = \frac{p_{i,k}}{p_{int,k}}, \tag{3.32}$$

where $p_{int,k}$ is the interference power from co-channel SBSs. Υ_{min} will result from the highest interference caused to UE k which occurs when it receives maximum interference power, $p_{j,k}^{max}$ from all the SBSs sharing the RBs with it, where $p_{j,k}^{max}$ is given by [31]

$$p_{j,k}^{max} = \frac{p_{i,k}}{\gamma_{th}} \tag{3.33}$$

Therefore, the minimum data rate $\mathcal{R}_{min}$

$$\mathcal{R}_{min} = \frac{N_{rb,min}\mathcal{B}_{rb}\log_2\left(1+\frac{p_{i,k}}{\hat{p}_{j,k}b_{int}}\right)}{\chi_j\left[1+(F-1)\left\{1-F_{D_r}\left(\left(\frac{1}{\gamma_{th}}\right)^{\frac{1}{\alpha}}\right)\right\}\right]} \tag{3.34}$$

where F is the number of SBSs, $\hat{p}_{j,k}$ is the maximum power that is received from an interferer for a given γ_{th} and b_{int} is the number of BSs that share the RB with UE k and it is derived in [31]. Maintaining $\mathcal{R}_{min}$, while increasing RB utilisation (i.e. increasing $p_{int,k}$) necessitates an increase in $N_{rb,min}$. If the transmit power, $P_{t,j}$, of BS j with the highest number of detected UEs is reduced so that $\bar{u}_j < u_j$, where $\bar{u}_j$ is given by

$$\bar{u}_j = \chi_j\left[1+(F-1)\left\{1-F_{D_r}\left(\left(\frac{P_{t,i}}{P_{t,j}\gamma_{th}}\right)^{\frac{1}{\alpha}}\right)\right\}\right], \quad P_{t,j} < P_{t,i}\ , \tag{3.35}$$

then the new $\mathcal{R}_{min}$ is given by

$$\mathcal{R}_{min} = \frac{N_{rb,min}\mathcal{B}_{rb}\log_2\left(1+\Upsilon_{min}\right)}{\bar{u}_j}. \tag{3.36}$$

According to (3.36), a reduction in $\bar{u}_j$ increases $\mathcal{R}_{min}$ linearly while a reduction in the Υ_{min} due to increased interference reduces $\mathcal{R}_{min}$ logarithmically. This presents the basis for algorithms 1 and 2.

3.6 Results and Performance Evaluation

The 3GPP suburban 5×5 grid model is used to represent an SBS deployment in an apartment building where each apartment is 10 m $\times$ 10 m [117]. One SBS and one user are randomly and uniformly dropped in each apartment. Moreover the choice of apartments to activate in each simulation run is also uniformly distributed. The data rate is calculated using (3.10). The pathloss between SBSs and UEs is given by (2.2). The remaining simulation parameters are given in 3.1. The initial transmit power for all SBSs is given by P_{max}/N_{sub}, where N_{sub} is the number of subcarriers in N_{rb} and P_{max} is the maximum transmit power of each SBS. Our technique is denoted by DMIM and its performance is shown for M-RRM [31] and MS-RRM [32]. With DMIM, P_t is different for different SBSs however it is equally allocated to all RBs. The performance of DMIM is compared to SM where all the SBSs transmit the same power, DM where only algorithm 1 is implemented and MP where all RBs are allocated maximum power that can be allocated to pilot symbols i.e. 10% of P_{max}. The ECR of BS j, is given by [98]

$$ECR_j = \frac{P_{DLj}}{\mathcal{R}_j}, \tag{3.37}$$

where $\mathcal{R}_j$ is the SBS's total data rate and P_{DLj} is the power consumed during downlink transmission given by

$$P_{DLj} = \frac{P_{RF} + P_{SP}}{\mu_{ps}}, \tag{3.38}$$

where μ_{ps} is the power supply efficiency, P_{SP} is the power used by the femtocell for signal processing and P_{RF} is the RF component of the power consumed and it is given by

$$P_{RF} = \frac{P_{t,j}}{\mu_{pa}}, \tag{3.39}$$

where μ_{pa} is the power amplifier efficiency.

Figure 3.1 shows the minimum data rate performance of the proposed technique as the number of SBSs increases. DMIM improves the minimum data rate attained by SP while MP reduces it. By ensuring that the detected users are less than those attained

Table 3.1: System Level Simulation Parameters

Parameter	**Value**	**Parameter**	**Value**
Resource blocks, N_{rb}	25	Number of subframes, N_{sf}	10
UEs per SBS, $\mathcal{X}$	1	Bandwidth	5/10 MHz
Carrier frequency, f	2.3 GHz	Amplifier efficiency, μ_{pa}	20%
Subcarrier Bandwidth	15 kHz	Power supply efficiency, μ_{ps}	85%
Frame period	10×10^{-3}s	Signal processing power, P_{SP}	3.35 W
Shadowing, φ	8 dB	SBS transmit power, P_{max}	20 dBm
Delay spread	10% of symbols	MBS transmit power	46 dBm
Fading	Rayleigh	OFDM symbol period, T_0	1.43×10^{-4}s
Thermal Noise Density, N_0	-174 dBm/Hz	MBS radius	1000 km
MBS-grid separation	35 m	MBS-UE separation	35 m
Wall Loss	20 dB	Pathloss exponent,α	3

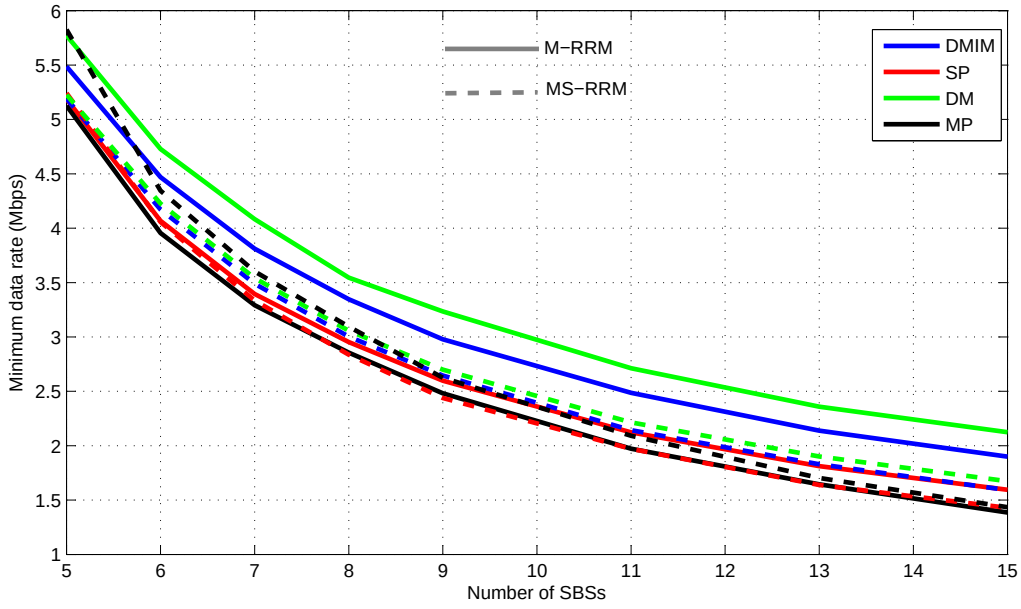


Figure 3.1: Minimum data rate performance with varying number of SBSs

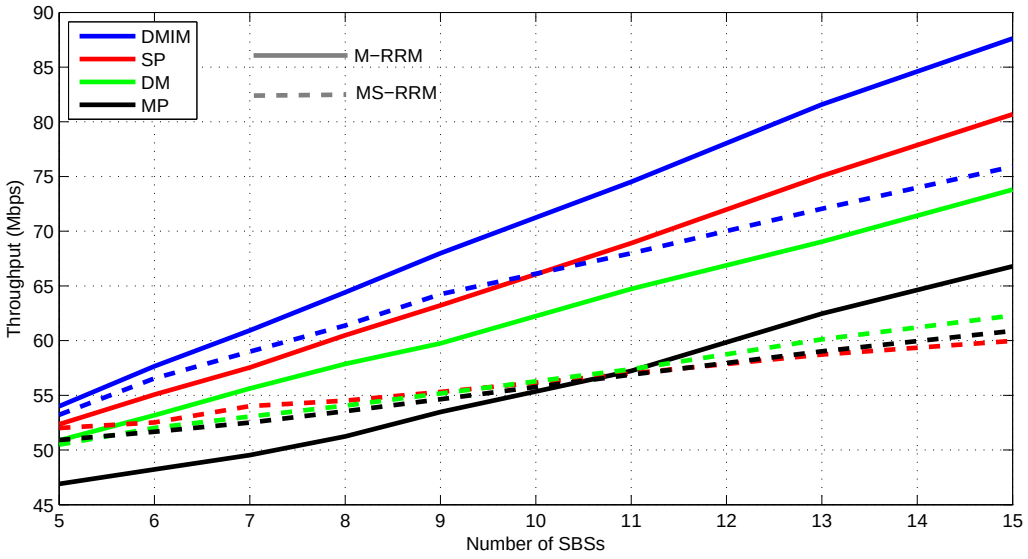


Figure 3.2: Throughput performance with varying number of SBSs

by SP, both DMIM and DM increase the minimum RBs allocated to the UEs which increases the minimum data rate. Unlike DM, DMIM also implements algorithm 2 which increases the level of interference and reduces the minimum data rate compared to DM. The minimum data rate for MP is less than that for SP when M-RRM is used and this is because more power is allocated to each RB which limits the number of RBs that can be allocated by each BS as a result of the power constraint. This reduces RB utilisation which reduces the minimum data rate. However, for lower BS density, MP has a high minimum data rate compared to other techniques. This is because each SBS uses less RBs and hence there is less reuse which results in higher data rates attained by the UEs. However as the BS density increases, the minimum data rate reduces because of increased interference from sharing RBs and reduced RBs accessed by each SBS.

Figure 3.2 shows that DMIM results in increased throughput compared to SP. This is because of the increased reuse of RBs by inner UEs which tend to have high SINRs. However, unlike M-RRM which has modest gains of up to 8.5% in throughput, MS-RRM has up to 25% improvement in throughput. This is because MS-RRM unlike M-RRM is distributed and does not maximise reuse opportunities for all UEs. However with DMIM decreasing detected UEs for inner UEs, the number of RBs they have access to increases significantly hence leading to higher gain in throughput. However, because it is not centralised, not all reuse opportunities are utilised resulting in

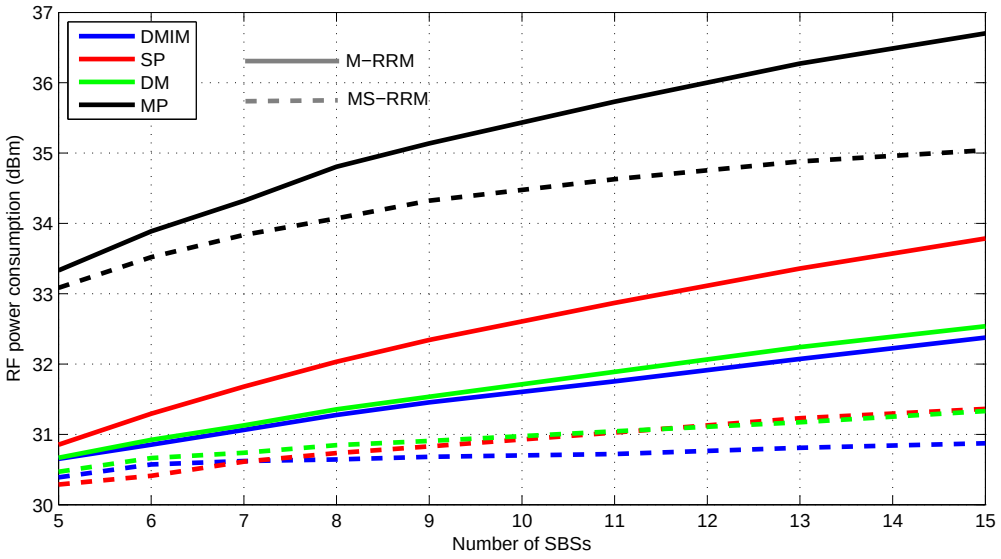


Figure 3.3: RF power consumption performance with varying number of SBSs

less throughput than M-RRM. Another reason for lower throughput is that with MS-RRM, if a UE k is interfered by a BS but that BS serves a UE that is not interfered by the serving BS of UE k, then, if UE k is allocated before that UE, it has no protection in a scenario where the interfering BS allocates its UE on the same RBs. DM has the lower throughput than SP because by reducing detected users, it increases the minimum RBs allocated to all UEs including those that are highly interfered and this effectively reduces RB reuse. For M-RRM, the gain in minimum data rate is traded off for a reduction in RB reuse hence resulting in lower throughput. However for MS-RRM, for higher density of SBSs, it has higher throughput than SP because it lowers the level of interference and boosts the data rate of some highly interfered UEs in the process. MP has the lowest throughput as a result of the limitation on the number of RBs allocated by each BS as a result of the power constraint.

The RF power consumption of DMIM is lower than that of the other techniques in Figure 3.3 for M-RRM while for MS-RRM, it is less only for 8 or more SBSs. For MS-RRM this is due to the initial increment in RB utilisation coupled with a lower reduction in power. DMIM has lower RF power consumption than DM because of the added power reduction during algorithm 2. MP consumes the most RF power because unlike the other techniques, the BSs allocate fewer RBs but utilise all their power or a higher percentage of their power in the process.

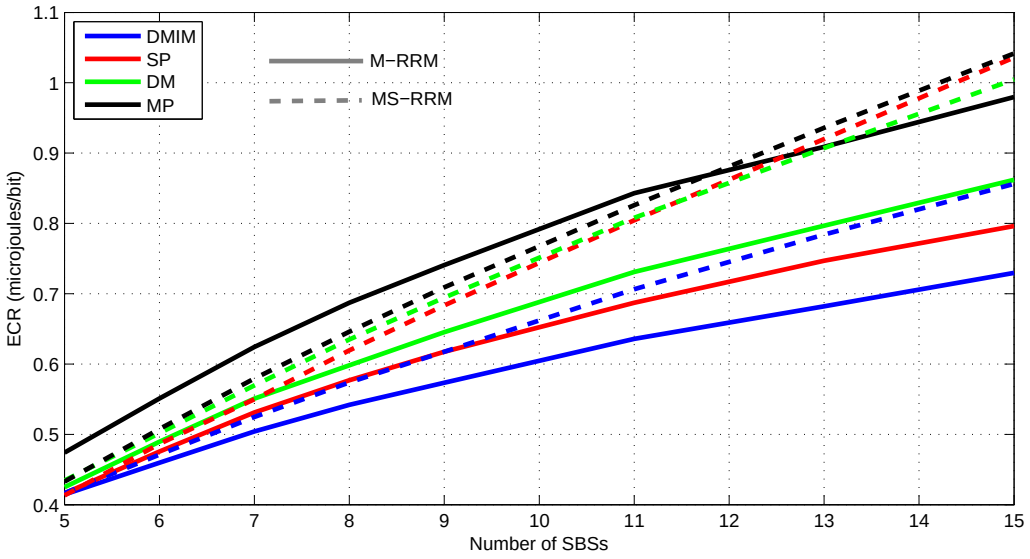


Figure 3.4: ECR performance with varying number of SBSs

The ECR performance of the technique is shown in Figure 3.4. The DMIM technique has the lowest ECR of all the techniques and the ECR improvement is higher as the number of SBSs increases. This is mainly attributed to both its lower RF power consumption and high throughput. For both MS-RRM and M-RRM, the performance of DMIM is better than that attained by the other techniques. DM and MP have higher ECR for M-RRM than DMIM and SP because they have lower throughput. On the other hand, for MS-RRM, DM has higher throughput than SP for higher BS densities and thus has better ECR for higher BS densities.

The CDF of the average UE data rate for the techniques is shown in Figure 3.5 for 15 deployed SBSs. For M-RRM, it can be seen that DMIM increases the data rate for all UEs and has a small percentage of UEs with very high data rates (higher than 30 Mbps). DM improves the data rate for low data rate UEs while MP reduces the data rates for low rate UEs because they have access to fewer RBs compared to SP. For MS-RRM, DMIM only slightly improves the data rates for low data rate UEs and then significantly improves the data rates of about 10% of the other UEs (the inner UEs). DM and MP have lower percentage of high data rate UEs compared to DMIM for both M-RRM and MS-RRM. Therefore the high data rates are mainly due to the few inner UEs that are allocated to a larger portion of the RBs.

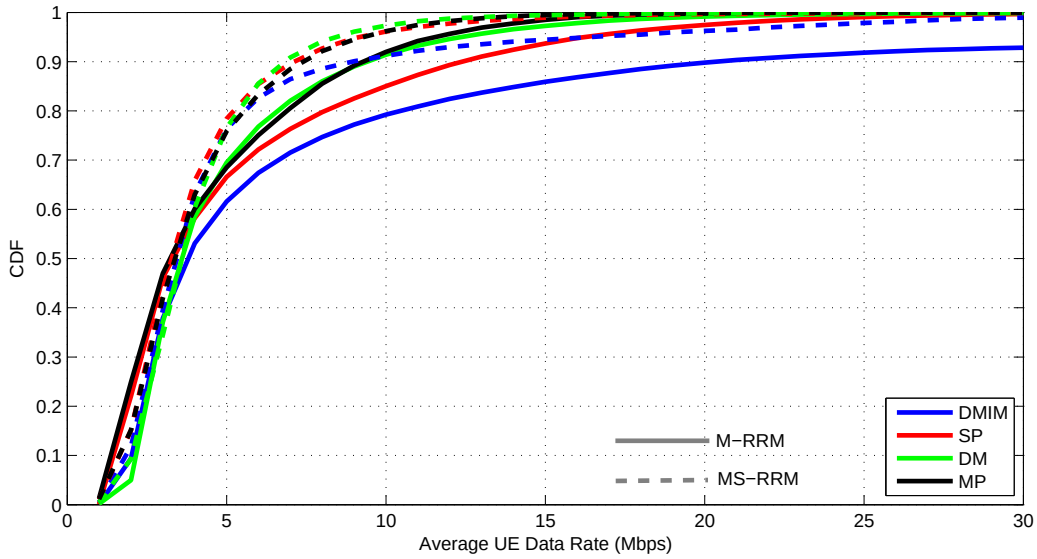


Figure 3.5: Average UE data rate CDF for 15 deployed SBSs

3.6.1 Impact of $\triangle p$ size

Figure 3.6 shows the impact of the size of $\triangle p$ on the sum rate of DMIM. It shows that the sum rate is higher for smaller $\triangle p$ and this is because the smaller the power step size allows more precision to be attained. Therefore, more reuse opportunities can be identified.

Figure 3.6 shows that smaller $\triangle p$ requires more iterations to reach convergence than the larger one. With each iteration, all the BSs need to send their updated MoC to the central controller which then instructs all BSs $j \in \mathcal{F}_{pc}$ to reduce power further or to stop. More iterations imply more delay associated with power allocation. The delay associated even with the larger values of $\triangle p$ is still too high. That is why for practical implementation of the technique, we recommend that the SBSs send the SIR difference matrix so that the central controller can carry out the algorithm without the need for back and forth updates. Once it is done, it sends each SBS instructions on how much power it should transmit. Considering the required iterations in Figure 3.6 and equations (3.20) and (3.21), it can be seen that the signalling overhead is lower when the SIR matrix is used because it is sent only once and the cental controller can then work out the power reduction needed by the SBSs.

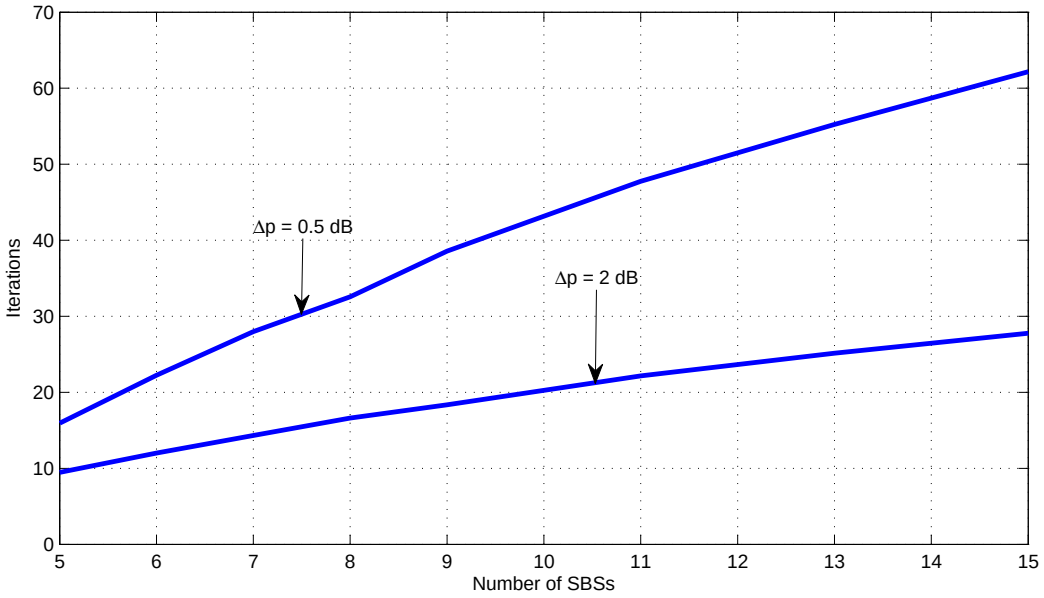


Figure 3.6: Required iterations for convergence

3.6.2 Heterogeneous Network Simulation

For the heterogeneous network simulation, we consider a macrocell overlaid with indoor SBSs in the 5×5 grid model. One grid is deployed in the macrocell coverage area at a distance of at least 35m from the macrocell. UE and SBS deployment is the same as in the homogeneous scenario. The activation parameter for each SBS is uniformly distributed in the range [0,1] and only SBSs with an activation parameter greater than 0.5 are considered in each simulation run. A scenario with 7 macrocells is considered with the grid deployed in the centre cell and a total of 30 UEs are deployed in the outdoor area of each macrocell in a random and uniform manner. A total bandwidth of 10 MHz is assumed for this simulation with 50 RBs available for allocation in each time slot. Both the macrocell and SBSs are equipped with two antennas space-time-block-code (STBC). The Stanford university interim (SUI) path loss model for Terrain type C is used to model the pathloss between the macrocell and all UEs and is given by [118]

$$PL_{j,k} = PL(d_0) + 10\Psi \log_{10}\left(\frac{d_{j,k}}{100}\right) + Xf + Xh + s + WL, \tag{3.40}$$

where $PL_{j,k}$ denotes the pathloss between macrocell j and UE k, $PL(d_0)$ is given by (2.2) in which $d = d_0 = 100$ m, $Xf = 6\log_{10}\left(\frac{f_c}{2000}\right)$, $Xh = 20\log_{10}\left(\frac{2}{\hbar_k}\right)$, $\Psi = 3.6 - 0.005\hbar_j + \frac{20}{\hbar_j} + 0.50\mathcal{N}_a$, $s = \mathcal{N}_b\,(8.2 + 1.6\mathcal{N}_c)$, where f_c is the carrier frequency in Hz, $\mathcal{N}_a$, $\mathcal{N}_b$ and $\mathcal{N}_c$ represent zero mean Gaussian random variables with unit variance, $\hbar_k$ and $\hbar_j$ are the UE and macrocell heights given by 2 m and 30 m respectively. The macrocell employs FFR where the coverage of the MBS is divided into inner and outer regions based on the FFR threshold which is presumed to be 20dB above the noise level. The resources allocated to the inner and outer regions are $N_{rb}/2$ and $N_{rb}/6$ respectively because a reuse factor of 3 is used.. The SBSs use the $N_{rb}/3$ resources allocated to the cell-edge UEs of neighbouring cells. Therefore the interference experienced by SBS UEs (SUEs) is from the neighboring MBSs. There is no coordination between the SBSs and the MBSs. To assess the performance of the technique, the maximum distance between the grid and the MBS is varied from 250 m to 1000m, the MBS coverage radius. The grid is randomly placed within the selected radius from the MBS. Only M-RRM is considered for the HetNet scenario.

The performance of our technique is compared to SP as described previously. Figure 3.7 shows the throughput results for the techniques and the results show that DMIM still has the higher throughput than SP. For both SP and DMIM, the throughput gradually reduces as the radius increases and this is attributed to the increased interference experienced by the SUEs when they are closer to the cell-edge. In all cases, the increase in throughput is approximately 8%.

Figure 3.8 shows the CDF of the UE data rates in the HetNet. The MUEs are not affected by the SBSs in this case and the throughput is the same for both DMIM and SP. This is because these are mainly other cell UEs and their most significant interferers are the neighbouring MBSs. The SUEs are seen to have slightly higher data rate than when SP is used and this explains the increase in throughput shown in Figure 3.7.

3.7 Summary

In this chapter, a power allocation technique to increase the throughput of a network without reducing the minimum data rate for two SON RRM techniques is presented. The technique improves throughput by increasing RB utilisation of UEs. This is achieved in two stages. The first stage increases the minimum RBs allocated to UEs

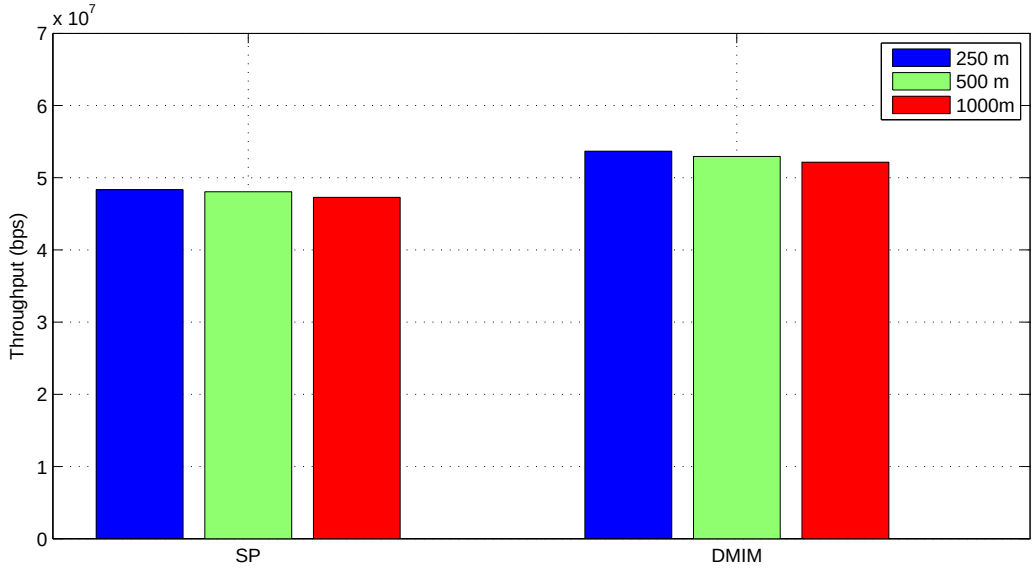


Figure 3.7: Throughput for SUEs

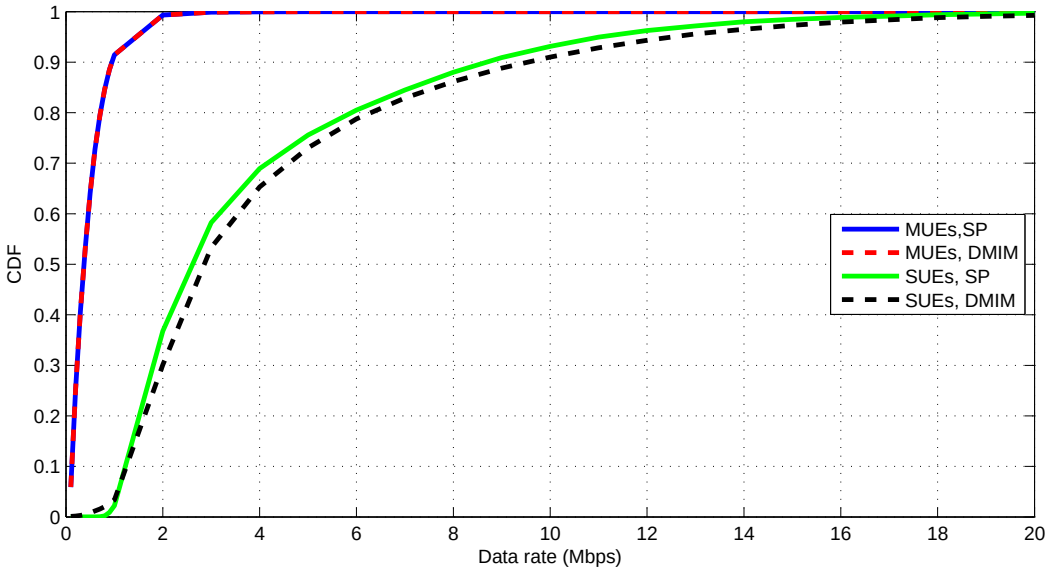


Figure 3.8: CDF of throughput in HetNet

while the second stage increases RB reuse by inner UEs. The proposed technique is compared to two other techniques i.e. one where the maximum pilot power is allocated to the RBs and another where the same power to transmitted by all SBSs. For both the decentralised and centralised SON RRM techniques, the results show that the proposed technique improves the throughput while simultaneously reducing RF power consumption and improving the minimum data rate and ECR of the network. Similar results in terms of throughput are obtained when the technique is applied in the HetNet. The technique is shown not to affect the MUEs while improving the SUE data rates and network throughput.

Chapter 4

Throughput Maximisation in Small Cell Networks using Power Control

IN this chapter, two power control algorithms with the goal of maximising throughput in a small cell network are presented. The first algorithm uses power control to improve RB reuse and ultimately increase throughput while the second technique uses power control to improve the throughput of lower data rate UEs by reducing interference from high data rate UEs. The work presented in this chapter is based on the work in [119]. The technique uses a distributed version of the MoC-based RRM technique to reduce overhead.

4.1 Literature Review

The use of power control to increase reuse and throughput or to minimise power consumption is the subject of a number of research studies [120–125]. In [120] a self organising rule to minimise downlink transmit power is presented where lower power is allocated to UEs close to the BS or those with low throughput demand. Following this rule, femtocells are able to allocate their UEs RBs in a distributed manner after a number of iterations. The paper [120] also introduces a coordinated RB allocation algorithm where femtocells coordinate through exchange of the High Interference Indicator which informs the neighbouring femtocells the RBs where they are creating too much interference for a neighbouring UE. The results show improvement in terms

of outage and throughput compared to other techniques like random allocation. A joint RB and power allocation algorithm whose aim is to minimise power consumption while meeting QoS requirements of UEs is presented in [121]. Clustering is used to manage co-tier interference with the cluster-head allocating orthogonal RBs to highly interfering femtocells. High priority UEs are allocated RBS first and are allocated only as much power as needed to meet the target SINR on a given RB. If they cannot meet the target SINR, they will be rescheduled for the next scheduling cycle. The best effort UEs are fairly allocated the remaining RBs and their power allocation is determined using linear programming techniques to minimise allocated power. In [122], a joint resource allocation and power control algorithm which mitigates the minimum sum interference level of the network is proposed. This is achieved by ensuring that the UE selected for allocation at each RB is the one that minimises the increase in interference in the network. The algorithm requires information exchange between the femtocells so as to generate a power transfer matrix which is used to determine the transmit power and consequently, which UE to allocate at a given RB. Similarly, power transfer matrices are used to solve the power allocation problem in [124]. If the solution is not feasible, the femtocell progressively removes UEs with the minimum SINR. To avoid dealing with power transfer matrix inversions, scheduling and femtocell grouping are introduced in [123]. The femtocells are grouped using a distributed graph colouring algorithm in order to protect cell-edge UEs by allocating them orthogonal resources [123]. A weighted proportional fair algorithm is used to schedule UEs while iterative power control is used to minimise transmitted power to the power needed to meet the target SINR. If the necessary power to meet the target SINR is greater than the maximum transmit power per RB, the maximum transmit power is allocated to the RB. The algorithm is shown to have lower outage than [122] in high interference scenarios with high data rate requirements. The general concept in all these papers is to minimise interference and power consumption by allocating only the necessary power to get the desired SINR so that UEs can meet their desired data rates. Further, some papers include scheduling when available RBs and power cannot meet all the UEs demands during the current RB allocation cycle [121–124]. Another approach to power and interference minimisation is presented in [116] where a power adaptation game based on game theory is used to minimise the transmit power based on the traffic load of the network. In the distributed power adaptation game, each femtocell reduces its power by a given factor and afterwards determines if any of the UEs served before the power reduction is now in outage. If any of the UEs are in outage,

it will increase its power by the factor it scaled it down with otherwise it will continue to reduce power until one or more of its UEs are put in outage. This power adaptation game is done in a distributed manner by only the femtocells that currently meet their blocking probabilities. Therefore, power adaptation game is activated in response to a change in traffic load in the network where the traffic is modelled using the Poisson distribution. Unlike the previously discussed techniques, reuse maximisation and ultimately throughput maximisation is the motivation for power control in [125]. The initial RB allocation is done using the MoC based RRM (M-RRM) technique. Power control is used to increase reuse by allocating RBs to UEs that were formerly forbidden from using them if their serving femtocells can reduce the transmit power on those RBs in order to create negligible interference to already allocated UEs.

In this chapter, an RB maximisation technique that is similar to [125] for M-RRM is proposed. Inspired by the work in [116], a distributed power adaptation algorithm is proposed to reduce the interference in the network by high data rate UEs which is similar to approach taken in [121–124].

4.2 System Model

The downlink of a perfectly synchronous LTE/TDD SBS network is considered. There are N_{rb} RBs for data transmission in each time slot. Further, a UE is allocated an RB for the duration of one transmission time interval (TTI). Let $\mathcal{F}$ and $\mathcal{K}$ denote the deployed SBSs and UEs in the network such that $\mathcal{F} = \{1, 2, \ldots, i, \ldots j, \ldots, \mathrm{F}\}$ and $\mathcal{K} = \{1, 2, \ldots, k, \ldots, N_u\}$. Let $i \in \mathcal{F}$ denote a random SBS in the set, which serves a UE $k \in \mathcal{K}$ and $j \in \mathcal{F}$ denote a neighbouring SBS to i. Let the number of UEs served by BS i be denoted by $\mathcal{X}_i$. The SBSs are connected to a central controller in order to manage resource allocation. It is assumed that there is no cross-tier interference. The SBSs are assumed to have SON capabilities and as such are able to find neighbouring SBSs. Therefore, all SBSs and UEs have a neighbouring SBSs list in their memory. Finally the SBSs are assumed to operate in open access mode. The signal-to-interference-plus-noise-ratio (SINR) of user k at RB r is given by

$$\Upsilon_{k,r} = \frac{p_{i,k,r} h_{i,k,r}}{\sum_{j=1, j\neq i}^{F} p_{j,k,r} h_{j,k,r} \widetilde{A}_{j,r} + N_0 \mathcal{B}_{rb}} \tag{4.1}$$

where $p_{i,k,r}$ and $h_{i,k,r}$ are the received power and channel gain from BS i to its UE k on RB r, respectively, $p_{j,k,r}$ and $h_{j,k,r}$ are the interference power and channel gain from BS j to UE k on RB r, respectively, $\mathcal{B}_{rb}$ is the bandwidth of an RB and $\widetilde{A}_{j,r} \in \{0, 1\}$ is an indicator which is equal to 1 if a UE served by BS j is allocated to RB r and equal to 0 otherwise. To calculate the achievable data rate at the RB r, adaptive M-ary quadrature amplitude modulation (MQAM) is used. The MQAM level that ensures the required BER, ρb, is selected. The achievable data rate per subcarrier is given by:

$$\mathcal{R}_{k,r} = \frac{\overline{\mathcal{R}}_{k,r}}{T_0}, \tag{4.2}$$

where T_0 is the OFDM symbol duration and $\overline{\mathcal{R}}_{k,r}$ is is the discrete capacity of UE k on RB r and it is given by [40]

$$\overline{\mathcal{R}}_{k,r} = \mathfrak{b}_{k,r} - E_{k,r}\rho s_{,k,r} \tag{4.3}$$

where $\mathfrak{b}_{k,r}$, $E_{k,r}$ and $\rho s_{k,r}$ are the number of bits derived from the SINR, the equivocation of the symbol and the symbol error rate at RB r respectively. $\rho s_{k,r}$ is given by [40]:

$$\rho s_{k,r} = \mathfrak{b}_{k,r} \cdot \rho b_{k,r}. \tag{4.4}$$

$E_{k,r}$ is given by [31, 40]:

$$E_{k,r} = \; -\rho s_{k,r} \cdot \log_2\left(\frac{\rho s_{k,r}}{2^{\mathfrak{b}_{k,r}} - 1}\right) - (1 - \rho s_{k,r}) \cdot \log_2\left(1 - \rho s_{k,r}\right) \tag{4.5}$$

The throughput, $\mathcal{R}_{tot}$, is given by 3.11.

4.3 RB Allocation

To reduce the overhead associated with the completely centralised M-RRM (CM-RRM) technique, the central controller sends ζ to the BSs. It is assumed that all SBSs have M-RRM allocation algorithm installed. Therefore, given ζ each SBS can determine where to allocate its UEs. To ensure that each SBS ends up with the same result, allocation of RBs is done consecutively instead of randomly as is done in the

original centralised MoC. This ensures that all SBSs generate the same $\overline{A}$ and are able to determine exactly which RBs to allocate their UEs without breaking the threshold constraint.

4.4 Power Control and Reuse Maximisation

The reuse maximisation algorithm is presented in subsection 4.4.1 while the power adaptation algorithm is presented in 4.4.2.

4.4.1 Reuse Maximisation Algorithm

Each BS determines the unique columns in $\overline{A}$ which are denoted by $\breve{A} \in \mathbb{R}^{K \times \overline{N}}$ where $\overline{N}$ denotes the number of unique columns in $\overline{A}$ which represent the unique UE allocations on the RBs. The SBSs with allocated UEs at a column n in $\breve{A}$ are denoted by $\mathcal{F}_n$. Using the MoC and $\breve{A}$, each SBS generates a potential allocations matrix, $\Lambda \in \mathbb{R}^{K \times \overline{N}}$, as follows

$$\Lambda_{k,n} = \begin{cases} 1, & \breve{A}_{k,n} = 0 \wedge \sum_{i=1, i \in \mathcal{F}_n}^{F} \zeta_{k,i} = 0, \\ 0, & \text{otherwise}, \end{cases} \tag{4.6}$$

where $\wedge$ denotes the 'AND' operator. Each UE with a 1 in $\Lambda_{k,n}$ has the potential to share the RBs with the combination of UEs in $\breve{A}_{k,n}$ allocated to it. Let the femtocells that serve such UEs be denoted by $\mathcal{F}_{pot}$. From the MRs received from their UEs, each SBS $i \in \mathcal{F}_{pot}$ determines the power reduction step, $\triangle p_i$, that would make the SIR of its UE equal to the γ_{th} with regard to its highest interferer in $\mathcal{F}_n$. If a femtocell serves more than one UE with $\Lambda_{k,n} = 1$, it will choose the one with the higher power reduction step. The SBSs send $\triangle p_i$ to the central controller in order for it to determine which BS will be allocated to the RBs with $\breve{A}_{k,n}$ first. At the same time, SBSs in $\mathcal{F}_n$ determine the power reduction that each of the SBSs in $\mathcal{F}_{pot}$ needs to make in order to maintain its current average MQAM level based on the MR. Each SBS will therefore create a vector, $\triangle \overline{p}^i \in \mathbb{R}^{1 \times F}$ with the necessary power reduction step, $\triangle \overline{p}^i_j$, by each BS so that its allocated UE can maintain its current MQAM level according to its MR. Any SBS $j \notin \mathcal{F}_{pot}$ will have $\triangle \overline{p}^i_j = 0$. The vectors are aggregated by the central controller. The

aggregated power reduction matrix is denoted by $\Omega \in \mathbb{R}^{F \times F}$ and it is given by

$$\Omega_{i,j} = \triangle \overline{p}^{i}_{j}. \tag{4.7}$$

The central controller deletes any SBS from $\mathcal{F}_{pot}$ where

$$\triangle p_i < \max(\Omega_{*i}) \quad i \in \mathcal{F}_{pot} \tag{4.8}$$

The condition in (4.8) means that the SBS will achieve an SIR less than γ_{th} if it reduces its power tomax (Ω_{*i}). If $\left|\mathcal{F}_{pot}\right| > 1$, the central controller will select $i \in \mathcal{F}_{pot}$with $\max\{\triangle p_i - \max(\Omega_{*i})\}$ to allocate its UE to the RBs where $\overline{A}_{k,r} = \breve{A}_{k,n}$ with the power reduced by $(\max(\Omega_{*i}) + p_c)$ in dB, where p_c denotes a constant that caters for increase in interference. After the BS allocates its UE, it updates the central controller with $\triangle \overline{p}_{i,}$ to ensure that any subsequent allocations do not reduce the data rate of the newly allocated UE. Based on the new allocation, $\mathcal{F}_{pot}$ is updated and if it is not empty, the SBSs in $\mathcal{F}_{pot}$ will update $\triangle p_i$ to cater for the new UE allocation. The process will be repeated until $\mathcal{F}_{pot} = \emptyset$. The process is then repeated for the remaining $\overline{N} - 1$ columns in $\breve{A}$. The procedure is detailed in Algorithm 3.

Algorithm 3 Reuse Maximisation

for r=1 until $\overline{N}$ **do**
  for i=1 until F **do**
    if $\sum_{k=1,k\in\chi_i}^{|\chi_i|} \Lambda_{k,n} > 0$ **then**
      $i \in \mathcal{F}_{pot}$
      if $\sum_{k=1,k\in\chi_i}^{|\chi_i|} \Lambda_{k,n} = 1$ **then**
        $\triangle p_i$ is the power reduction for UE $k \in \chi_i$ to attain γ_{th}
      else if $\sum_{k=1,k\in\chi_i}^{|\chi_i|} \Lambda_{k,n} > 1$ **then**
        $\triangle p_i$ is the maximum power reduction value for $k \in \chi_i$ to attain γ_{th}
      end if
    else if $\breve{A}_{k,n} = 1$ **then**
      Determine $\triangle \overline{p}^i$
    end if
  end for
  for i=1 until F **do**
    if $\triangle p_i < \max(\Omega_{*i})$ **then**
      Delete i from $\mathcal{F}_{pot}$
    end if
  end for
  while $\mathcal{F}_{pot} \neq \emptyset$ **do**
    Find $j \in \mathcal{F}_{pot}$ with $\max\left(\triangle p_j - \Omega_{*j}\right)$
    Allocate UE and reduce power of j by $\max\left(\Omega_{*j} + p_c\right)$ on RBs where $\overline{A} = \breve{A}_{k,n}$
    Delete j from $\mathcal{F}_{pot}$
    Update Ω for newly allocated UE
    Update $\triangle p_j$ for all $j \in \mathcal{F}_{pot}$ to include newly allocated UE
    repeat 14 to 18
  end while
end for

4.4.2 Power Adaptation

Let Υ_1^L denote the minimum SINR for the highest MQAM level. Some UEs have SINRs that are greater than the SINR required for the maximum MQAM level. Despite

these high SINRs, the maximum number of bits they can transmit is limited by the maximum MQAM level. This means that the power transmitted to such UEs can be reduced without reducing their data rates in order to reduce interference to other UEs. In order to achieve this goal, the power transmitted to the UEs is reduced until their SINR is equal to Υ_1^L. There is no need for coordination between the SBSs to implement this power reduction and one iteration is enough. In the first iteration, all SBSs with UEs with $\Upsilon_{k,r} > \Upsilon_1^L$ at RB r reduce the power on the RB so that its SINR is equal to Υ_1^L. The power reduction is determined using $\triangle p_{k,r} = \Upsilon_{k,r} - \Upsilon_1^L$. If any of the UEs allocated at RB r has increased SINR such that it is greater than Υ_1^L then it too reduces power. To avoid an endless cycle of power reduction, each BS reduces power only once. The procedure is detailed in Algorithm 4. The difference between this algorithm and the ones presented in [121–124] is that there is no target data rate here so the target SIR (because the noise is very low) is focused on interference reduction for lower rate UEs while minimising reduction in data rate for the high data rate UE. However, the loss in data rate cannot be avoided because power reduction reduces the SINR which increases $\rho b_{k,r}$ which ultimately reduces the data rate as seen in (4.3). Therefore, fractions of bits per RB are sacrificed for the higher data rate UEs to improve the data rate of some lower data rate UEs.

Algorithm 4 Power Adaptation Algorithm

```
for i=1 until F do
    for k=1 until χ_i do
        for r=1 until N_rb do
            if Υ_{k,r} > Υ_1^L then
                Reduce power until Υ_{k,r} = Υ_1^L
            end if
        end for
    end for
end for
```

4.4.3 Convergence Analysis

Algorithm 3 requires no convergence analysis because the power reduction is done in once for each affected SBS and the amount of power reduction is determined by the central controller. On the other hand , Algorithm 4 some convergence analysis is

required. Suppose an RB is shared by three UEs, a number of scenarios are possible. Scenario 1 is that all UEs have SINRs that are greater than Υ_1^L. In this case, all SBSs will reduce their transmit power in order to reach Υ_1^L however because their interferers are also reducing their power, Υ_1^L is unattainable. To avoid the unnecessary further reduction in power, Algorithm 4 imposes the rule that power reduction is done only once. This is necessary because the SBSs do not share information at this stage and implement the power reduction in a distributed manner. Scenario 2 is one where some of the UEs have SINRs that are less than Υ_1^L. In this case the UEs with SINRs greater than Υ_1^L will reduce power in the first round. Depending on the interference reduction, some of the UEs that initially had SINRs less than Υ_1^L may now have SINRs that are greater than Υ_1^L. Therefore in the second iteration, these UEs would also reduce power. Because each SBS only reduces power once, the next iteration will be for only those UEs whose serving SBSs have not yet reduced power. If any of such UEs have the SINRs greater than Υ_1^L then their SBS reduces power otherwise no SBS will reduce power and the algorithm will have converged for a given RB.

4.4.4 Signalling Overhead Comparison

The RB allocation used is denoted by DM-RRM, the semi-distributed version of CM-RRM which aims to minimise the signalling overhead associated with CM-RRM. However it only works if the allocation algorithm is known to the BSs. The total amount of information exchange between the central controller and BS i for DM-RRM is given by

$$b_{in,dm} = \chi_i \cdot b_\zeta \cdot (F-1) + N_u \cdot b_\zeta \cdot F \tag{4.9}$$

while the information exchange between the central controller and the BSs for CM-RRM is given by

$$b_{in,cm} = \chi_i \cdot b_\zeta \cdot (F-1) + N_u \cdot b_{\overline{A}} \cdot N_{rb} \tag{4.10}$$

where b_ζ is the number of bits required for the MoC, $b_{\overline{A}}$ is the number of bits required to encode the allocation matrix. Unlike the $\overline{A}$ which is made up of only ones and zeros, ζ requires two bits because it has three unique values. Therefore, as long $2F < N_{rb}$, then there is a reduction in the control overhead when DM-RRM is used. The information required for the power allocation algorithm in [125] and the one proposed in this

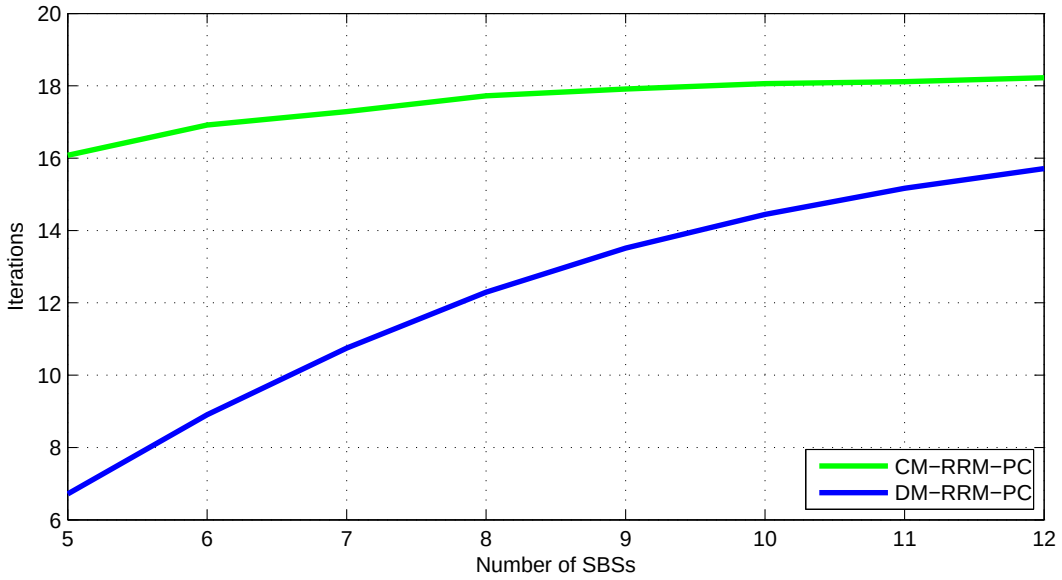


Figure 4.1: Required iterations for reuse maximisation

paper is nearly the same. However, the algorithm in [125] requires that information for each RB where a potential UE may potentially be allocated while the proposed one only needs to identify the RBs with unique allocations. The power allocation for the new UE is the same for all the RBs where $\overline{A} = \breve{A}_{*n}$. This reduces the number of iterations needed. Figure 4.1 shows the difference in iterations required for the reuse maximisation by centralised M-RRM with power control (CM-RRM-PC) [125] and distributed M-RRM with power control (DM-RRM-PC) from which it can be seen there is a reduction in the required information by DM-RRM-PC.

4.4.4.1 Central Controller Implementation

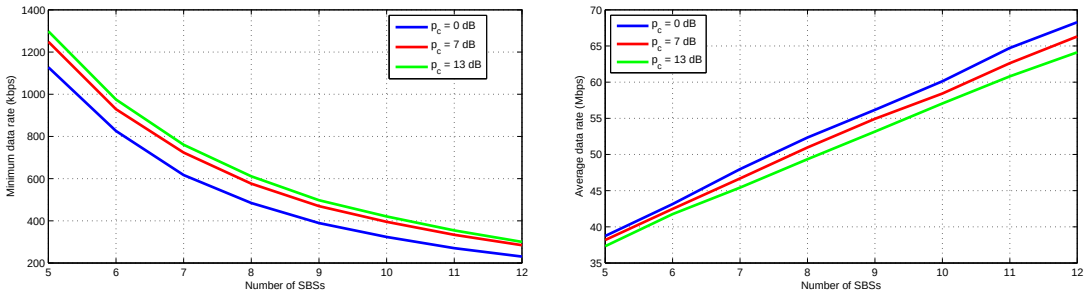
To further speed up this the reuse maximisation algorithm, the SBSs may share with the central controller the SIR difference matrix, $\mathfrak{I}$, instead of the ζ. That way, the central controller can determine which UEs to allocate to RBs and the required power reduction without the need for repeated information requests to the SBSs. Once it determines the power reduction step, the central controller will then inform the selected BSs how much power they need to allocate on the RBs where they're allocated.

Table 4.1: System Level Simulation Parameters

Parameter	**Value**	**Parameter**	**Value**
Resource blocks, N_{rb}	25	Number of subframes, N_{sf}	10
UEs per SBS, $\mathcal{X}$	2 per BS	Bandwidth	5 MHz
Carrier frequency, f	2.3 GHz	Amplifier efficiency, μ_{pa}	20%
Subcarrier Bandwidth	15 kHz	Power supply efficiency, μ_{ps}	85%
Frame period	10×10^{-3}s	Signal processing power, P_{SP}	3.35 W
Shadowing, φ	8 dB	SBS transmit power, P_{max}	20 dBm
Delay spread	10% of symbols	OFDM symbol period, T_0	1.43×10^{-4}s
Fading	Rayleigh	Pathloss exponent, α	3
Thermal Noise Density, N_0	-174 dBm/Hz	SIR threshold, γ_{th}	20 dB

4.5 Results and Performance Evaluation

The scenario used for this evaluation is an enterprise of 60 m $\times$ 60 m with a varying number of SBSs. It is assumed that there is no cross-tier interference which may be due to orthogonal allocation of RBs or the absence of a macrocell. The SBSs are uniformly distributed throughout the enterprise. The proposed algorithm is evaluated for open access SBSs with two UEs per SBS. Two UEs are uniformly dropped in a circle of radius 10 m around each SBS. The pathloss between SBSs and UEs is modelled by (2.2). All SBSs are open access. The remaining simulation parameters are given in 4.1. The performance of the technique DM-RRM with power control (DM-RRM-PC) is compared to centralised M-RRM with no power control (CM-RRM-NP) [31] and centralised M-RRM with power control (CM-RRM-PC) [125]. The initial transmit power for all SBSs is given by P_{max}/N_{sub}, where N_{sub} is the number of subcarriers in N_{rb} and P_{max} is the maximum transmit power of each SBS. The ECR is calculated using (3.37).



(a) Minimum data rate (b) Throughput

Figure 4.2: Impact of p_c

4.5.1 Impact of p_c

The impact of the choice of p_c on the minimum data rate and throughput of the technique is shown in Figures 4.2a and 4.2b respectively. A lower value of p_c results in a higher throughput but a lower minimum data rate. This is because the lower the value of p_c, the higher the level of interference to the already allocated UEs which reduces their SINRs and consequently their data rates. However for the newly allocated UEs, the lower p_c means that their SINRs are higher and this leads to an increase in the throughput. Therefore the choice of p_c depends on what the aim of the power control is. If it is to maximise throughput irrespective of the impact on allocated UEs then p_c should be set to zero. For the simulations in the next section, p_c = 10dB for 8 or less BSs and p_c = 13dB for 9 or more BSs.

4.5.2 Minimum Data Rate

Figure 4.3 shows the variation of the minimum data rate performance with increasing BS density. DM-RRM-PC has the highest minimum data rate for low SBS density with a 3% increment in minimum data rate for 5 SBSs over the minimum data rate attained by CM-RRM-NP. The increment is as a result of the power adaptation by SBSs that served high data rate UEs which reduces the interference for other UEs. The increment gradually reduces and beyond 11 SBSs, the minimum data rate attained by DM-RRM-PC is about 1.5% less than that obtained by CM-RRM-NP. On the other hand, CM-RRM-PC results in approximated 1.5% reduction in the minimum data rate of CM-RRM-NP. This is because in determining minimum interference, the average

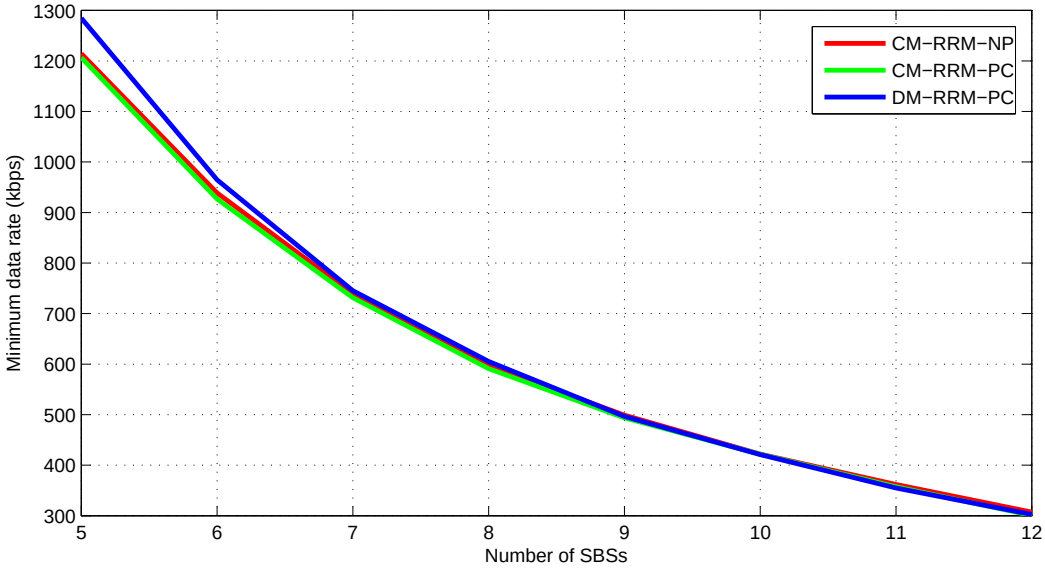


Figure 4.3: Minimum data rate

channel gain is considered yet some subcarriers may have high channel gain from the newly allocated BS causing the SINR to reduce.

4.5.3 Throughput

Figure 4.4 shows the throughput variation with increase in BS density. DM-RRM-PC is shown to have the highest throughput of the techniques that are compared. This is because the reuse maximisation increases the data rate of UEs that are allocated extra RBs after power control and power adaptation increases the data rate for lower rate UEs by reducing interference from higher rate UEs. The combination of the two algorithms results in an increase that ranges from $11 - 13\%$ compared to CM-RRM-NP as the BS density increases. CM-RRM-PC results in nearly the same throughput as DM-RRM-PC for low BS densities but it gradually reduces up to 3% lower throughput than DM-RRM-PC for 12 SBSs. This is because it is more restrictive in its reuse policy than DM-RRM-PC because it takes the channel conditions in to account yet DM-RRM-PC considers the average received power.

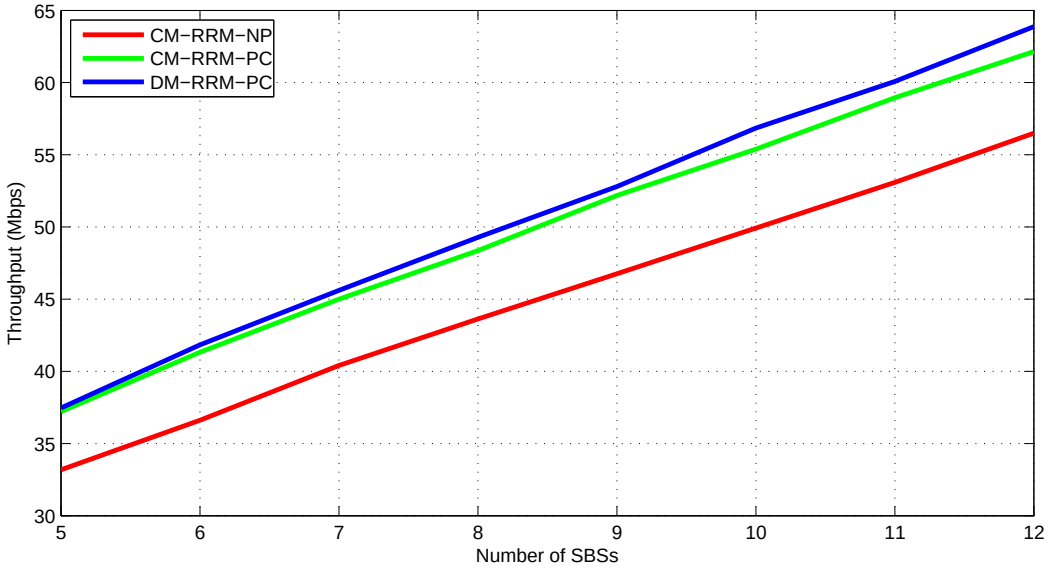


Figure 4.4: Throughput

4.5.4 RF Power Consumption

Figure 4.5 shows the RF power consumption of the techniques and it shows that DM-RRM-PC has the lowest RF power consumption. This is because of the power adaptation process where SBSs with high SINR UEs lower the transmit power in order to attain minimum SINR Υ_1^L. CM-RRM-PC has slightly higher RF power consumption than CM-RRM-NP. This is because during the initial round of allocation, the CM-RRM-PC consumes the same amount of power as CM-RRM-NP but when power control is used to increase resources, more power is allocated to the new allocations and this increases the total RF power consumption.

4.5.5 ECR

Figure 4.6 shows the ECR of the techniques with DM-RRM-PC attaining the lowest ECR which means it has the highest energy efficiency. This is mainly due to a combination of higher throughput and lower RF power consumption than the benchmark techniques. CM-RRM-PC has a lower ECR than CM-RRM-NP despite having a higher RF power consumption than CM-RRM-NP. This is because the ECR performance is dominated by the throughput which is a much bigger value than P_{DLj} in (3.37).

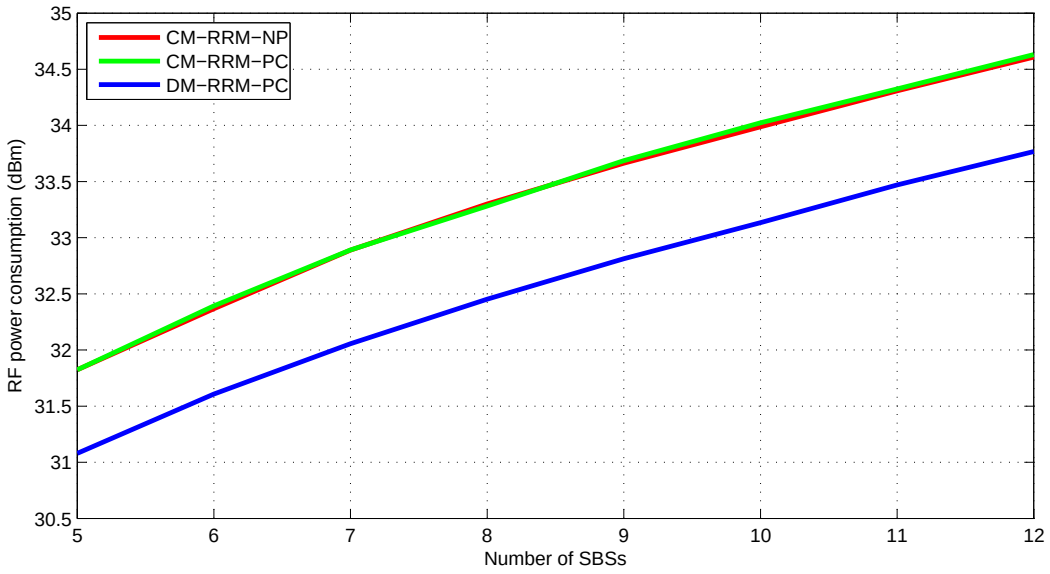


Figure 4.5: RF power consumption

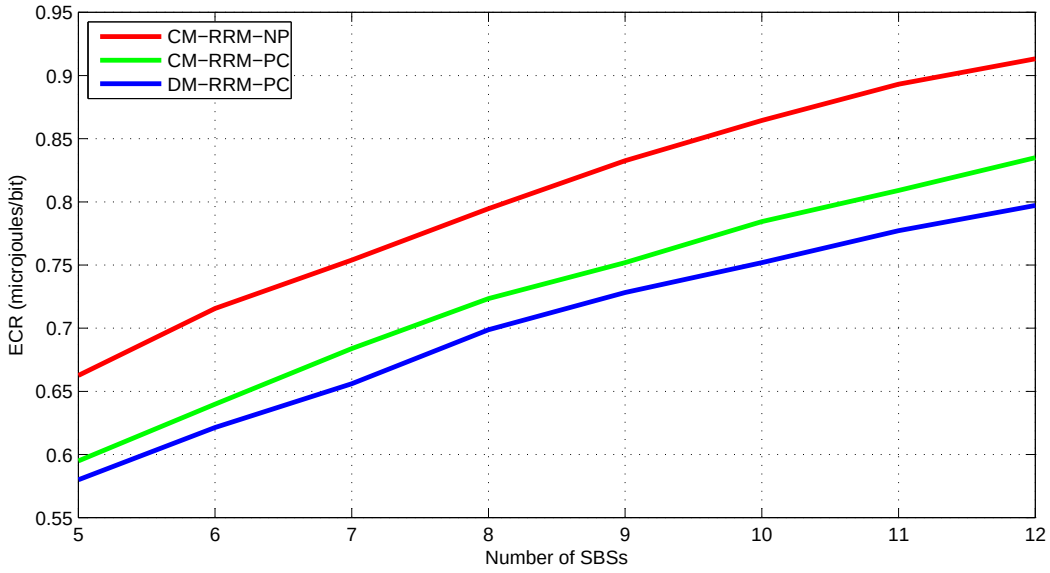


Figure 4.6: ECR

4.6 Summary

In this chapter, two algorithms have been presented i.e. a reuse maximisation algorithm and a power adaptation algorithm. In addition, a semi-distributed form of the MoC algorithm where each SBS only needs the aggregated MoC from the central controller in order to determine which RBs to allocate its UEs to, is presented. The reuse maximisation algorithm increases reuse based on the power received in the MR reports while the power adaptation algorithm relies on the BSs to reduce power in order to improve the data rates of the lower rate UEs. The results show that the proposed technique increases the throughput by $11 - 13\%$ compared to network with no power control applied and it improves the throughput by 3% compared to another bench mark technique that improves throughput by increasing reuse. For lower SBS densities, the increase in throughput is attained along with improvement or maintenance of the minimum data rate when no power control is applied. However for more than 10 SBSs, the minimum data rate reduces by about 1.5% which is similar to the performance of the benchmark technique. Therefore the performance of the proposed technique is shown to improve that of the benchmark technique even though it requires less information exchange to achieve the same result.

Chapter 5

AID-based Backoff and Grouping in 802.11ah Networks

IN this chapter, a new approach to setting backoff counters in 802.11ah networks is presented. The work presented in this chapter is based on the work in [126].

5.1 Introduction

Although the Internet has traditionally been dominated by human-to-human (H2H) communication, the past decade has witnessed a rapid rise in machine type communications (MTC) which include human-to-machine (H2M), machine-to-human (M2H) and machine-to-machine (M2M) communications. This has given rise to the Internet of things (IoT), a network that connects devices which carry out both H2H communications and MTC [4, 12, 24]. For a fully-fledged IoT, communication networks should be able to support heterogeneity in devices ranging from transmit power to required data rates [12, 26]. Whereas several existing wireless local area network (WLAN) technologies can support high transmit power devices over both short ranges and others support low transmit power devices over short ranges, most cannot support low transmit power devices over long ranges [12, 20]. Further, the cost of licensed spectrum is prohibitive for some IoT applications, many of those involving MTC, which makes using license-exempt bands an attractive alternative [4]. However, the 2.4 GHz and 5 GHz bands are congested and owing to their propagation characteristics, are more suited for high

data rate short range communications [127]. On the other hand, the sub-1 GHz frequencies have the required propagation characteristics to cater for the heterogeneity of transmit power and coverage range required by IoT applications [26]. This is why this band was selected by the 802.11ah standard [21, 128]. This standard has made a number of modifications to the existing 802.11 standards in order to support low transmit power, long range MTC. Some of the key modifications made include increasing the maximum number of stations (STAs) or devices supported by an access point (AP) from about 2000 STAs to 8192 STAs by using a new association identifier (AID) format [21]. Other modifications include the introduction of slotted medium access using a restricted access window (RAW) to reduce contention, group sectorisation to mitigate the overlapping basic service set problem (OBSS), short beacons and null data packets (NDP) for some common frames like the ACK frame to reduce the control overhead [26].

Similar to other 802.11 standards, 802.11ah uses the distributed coordination function (DCF), a contention-based medium access control (MAC) protocol, for asynchronous data transmission without central coordination [21, 44]. However, such a contention-based protocol cannot cope with thousands of STAs contending for channel access simultaneously because of the excessive number of collisions. Furthermore, the large separation distance between low power STAs results in the hidden node problem in which collisions occur because STAs cannot detect each other [27]. One of the techniques used to address the above challenges is group sectorisation. Sectorisation involves the partitioning of the coverage of a basic service set (BSS) into non-overlapping sectors. This is achieved by using a set of antennas which focus the transmission and reception of the AP on one sector at a time [26]. This enables stations to transmit in a time division multiplexing (TDM) manner [26, 129]. With group sectorisation, a beacon is transmitted at the start of each sector interval and only STAs in that sector have access during the sector interval [26, 130]. Further, STAs can only gain access to the transmission medium during the time allocated to their group. If the grouping is location-aware, it can solve the hidden node problem [26]. Even without hidden nodes, DCF suffers from collisions that result from contention-based channel access which requires STAs to randomly choose a backoff timer value for the backoff process.

In this chapter, in order to completely eliminate collisions in sectorised 802.11ah networks, a new approach to setting the backoff timers for contending STAs is proposed.

The STAs in the BSS are first grouped using hierarchical clustering and the Welsh-Powell algorithm in order to eliminate hidden nodes. To eliminate collisions, each STA sets its backoff timer based on the position of its AID in the group. The results show that the groups created are not uniform in size. Because the channel access time for all groups is the same [26], non-uniform sized groups violate fairness principles since the STAs in the smaller groups have more access to the channel than those in larger groups. To remedy this, the Welsh-Powell algorithm is modified to ensure that group sizes do not exceed a target group size. A merge-and-split algorithm is then used to create uniform-sized groups for the groups that are smaller than the target group size. An analytical model for the throughput attained by the AID-based backoff counters with groups that are free of hidden nodes is developed and it is verified through simulations. Finally, the performance of DCF with AID backoff counters is compared to its performance with random backoff and PCF. The main contributions in this chapter are:

- Present a grouping technique for uniform sized hidden node-free groups in 802.11ah networks.
- Present a novel approach to setting backoff timers of STAs using their AIDs in sectorised 802.11ah networks to eliminate collisions and reduce idle time.
- Present a mathematical model that accurately captures the normalised throughput for DCF with AID-based backoff counters.

The rest of this chapter is organised as follows. An overview of related work in 802.11ah networks is presented in Section 5.2. The system model is presented in Section 5.3, while Section 5.4 presents the grouping process. The AID-backoff technique is presented in Section 5.5, while Section 5.6 presents the analytical model for the throughput of AID-based backoff technique. Validation of the throughput model through simulations is presented in Section 5.8 which also presents the performance of the grouping technique as well as a comparison of the performance of AID-based backoff to other techniques. Finally, concluding remarks are presented in Section 5.10.

5.2 Literature Review

Grouping STAs in order to improve network performance metrics has been studied for legacy 802.11 standards. For instance, in [131], STAs are grouped based on their data rate with each group having different initial backoff window, frame size and maximum backoff stage in order to improve throughput by fairly allocating the channel occupancy time [131]. In [132], STAs are grouped according to their saturation status, where it has been shown that in order to maximise throughput, the initial backoff window should be increased for saturated groups. This is in agreement with Bianchi's finding that the optimal window size is proportional to the number of contending STAs [37]. However, the optimal window size for thousands of devices would be too large resulting in inefficient resource utilisation due to increased idle time [4].

In [133], a grouping-based DCF (GB-DCF) mechanism is proposed, where each STA randomly selects a group and each group accesses the channel in a TDM manner. Further, individual ACK messages are replaced with block acknowledgements in order to improve throughput. Results show an 81% improvement in throughput over DCF. However, the gain in throughput diminishes when the number of groups is large because empty groups lead to wastage of resources [133]. Similarly, the authors in [4] found that when STAs choose their groups, the throughput reduces because of empty group slots. Their results also show that group slot sizes should be carefully selected because if they are larger than required, they can result in reduced throughput from idle time, especially when a no-crossing slot policy is applied. This implies that the transmissions of one group are not permitted to exceed the group's allocated slot boundary [4].

In [24], two MAC protocols that use grouping to reduce the level of contention in a smart meter network are proposed. One is the same as in [4] while the other involves grouping STAs and selecting a group leader for each STA. The group leaders contend for the channel and the one that successfully transmits a packet wins the channel for its group members. Results in this chapter show that both protocols significantly increase the throughput and reduce the delay when compared DCF without grouping [24]. It is noteworthy that in both [24] and [4], the details of how groups are formed are not given. More importantly, it is assumed that the groups have no hidden nodes. The authors in [134] propose sector-based ready-to-send and clear-to-send frame transmissions to reduce the number of collisions in the network. Their results show that

throughput is maximised for up to one hundred STAs but reduces beyond that due to the increased level of contention. A distributed grouping algorithm based on received power is presented in [135]. The AP determines the number of groups and selects the same number of random STAs to act as group heads. The group heads are then assigned slots in the RAW during which they transmit a pilot signal. The remaining nodes then choose which group to join based on the strongest signal received. Their work shows higher throughput compared to random grouping. However the parameters like slot time and duration of interframe spaces are shorter than those of the standard. Further, the pathloss model is different from that recommended by the standard.

All the techniques discussed in this section aim at improving throughput by reducing contention in the network and therefore reducing collision probability. Furthermore, most of them assume that uniform-sized groups can be created without any hidden nodes. Irrespective of reduced contention and absence of hidden nodes, collisions—which waste device power and reduce the throughput—still occur because of the random backoff timers. The main contribution in this chapter is therefore a simple approach to eliminate collisions by setting the backoff timer using STA AIDs for hidden node-free groups.

5.2.1 Outage in 802.11ah

Yet another key problem faced by 802.11ah is coverage for low power devices that are at the cell edge [25]. Some solutions to the problem suggested by [25] include reducing the packet size and reducing the required reliability. Reducing the packet size reduces the PER (refer to (5.3)) which will increase the reliability and reduce the number of STAs in outage. On the other hand, reduced reliability reduces outage by allowing STAs with higher PER than that set by the standard, to connect to the AP. Both techniques result in reduction of throughput i.e. reducing packet size increases the overhead to payload ratio while reduced reliability increases retransmissions due to packet errors [25]. To improve coverage, the standard introduced support for relays [26]. The relays are installed closer to the cell edges and because they're closer to cell edge STAs, it means that the received signal has a higher SINR and hence lower PER. A number of studies have been carried out to examine the performance of relays in 802.11ah networks. In [136], the authors developed a model for transmit opportunity (TXOP) sharing in 802.11ah networks with relays. Using this model and simulations,

they were able to show the improvement in throughput when relays are used, owing to the higher MCS used because of higher SINR between the relay and the cell-edge STAs than between the AP and STAs. In [137] a hierarchical MAC protocol with multi-channel assignment for relays in order to improve throughput in 802.11ah networks is proposed. The protocol first determines the threshold distance from the AP beyond which a UE is not permitted to transmit directly to the AP. The relays are allocated slots in the RAW during which they forward packets from the STAs to the AP [137]. Outside their allocated RAW slots, the relays use orthogonal channels to receive data from the STAs. Although the technique attains enhanced throughput, it assumes all the required orthogonal channels are available yet this may not always be the case. Similarly, the authors in [138] also present an algorithm to determine multi-channel assignment to relays in 802.11ah. The relays are also allocated slots in the RAW for forwarding data from the STAs. Further, the authors present a mathematical model to determine the optimal RAW and slot size to cater for heterogeneity in traffic from different groups. Their results show enhanced throughput for 802.11ah networks.

Network virtualisation is a technique used in Wi-Fi networks to get around the connectivity model that restricts a STA to connect to one AP by letting a network card connect to multiple networks simultaneously and in a manner that is transparent to the operating system [139] . Further, it can also be used to define different roles for the same STA. In [140], network virtualisation is used in the creation of infrastructure mode ad-hoc networks. It involves the definition of virtual network interfaces to enable the STA to operate as both an AP and STA. The down side is that the STA can only be in one mode at a time implying that it needs to switch modes in order to fulfil both duties [140].

Although relay use has been shown to improve throughput and reduce outage in the network, the work in this chapter considers the problem of reducing outage in a network without relays. It is assumed that the STAs have network virtualisation functionality and therefore can self configure to act as relays of STAs. This functionality enables a relay STA to associate with the outage UE in its neighbourhood and thereafter relay its packets to the AP.

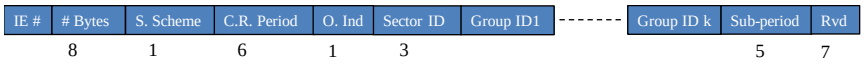


Figure 5.1: Sectorisation Information Element [3]

5.3 System Model

A macro deployment of an IEEE 802.11ah network with N_u STAs accessing the wireless channel is considered. Only single hop communication between the AP and an STA is considered. The channel is non-ideal resulting in packet errors. All packets have the same length and all STAs use the same modulation and coding scheme (MCS) class. It is assumed that the network is dominated by uplink traffic which implies that the STA is the transmitter and the AP is the receiver [141, 142]. The network is saturated i.e. each STA has at least one packet to transmit during each sector interval.

5.3.1 Group Sectorisation

The AP implements group sectorisation and divides its coverage into six equal non-overlapping sectors. A beacon frame is transmitted at the start of each sector interval. The beacon frame has a sectorisation information element (IE) that indicates which sector is active and which groups can have access to the channel during the allocated time. Only STAs in that sector have access during the sector interval. Further, the sector interval is divided into group slots where the allocated group of STAs has exclusive access to the channel. It is assumed that the use of sectorisation mitigates the OBSS problem. The format of the sectorisation IE is shown in Fig. 5.1, where the number of bits allocated to each field is indicated. The S. Scheme indicates the sectorisation scheme with a zero indicating group sectorisation, the C.R. period is the complete rotation period (i.e. the number of beacon intervals for all sectors), O. Ind indicates whether or not the AP transmission is omni-directional, Sector ID indicates the ID of the current sector, Group ID indicates the groups that are permitted to transmit in the current sector interval, the sub-period indicates the sub-period for the current sector and finally Rvd refers to the reserved bits [3].

5.3.2 802.11ah Pathloss Model

The pathloss between the AP and an STA is given by [27, 128]

$$PL_{as}(d) = 8 + 37.6 \log_{10}(d), \tag{5.1}$$

where d is the distance between the transmitter and the receiver. The pathloss between two STAs is calculated using [27]

$$PL_{ss}(d) = -6.17 + 58.6 \log_{10}(d). \tag{5.2}$$

Assuming that the fast fading effect is averaged out, the power received at the receiver is given by (2.1).

5.3.3 Medium Access Between Group Slots

In order to avoid collisions between STAs in different groups, STAs are strictly forbidden from using other groups' time slots. STAs will freeze their backoff counters if time left in their group slot is less than the minimum duration for successful transmission and reception of one packet .

5.4 STA Grouping

In a group without hidden nodes, all STAs should be able to decode the preamble of frames sent by other group members. This is possible if the received power is above the minimum receiver sensitivity. For 802.11ah, the minimum receiver sensitivity defined as the input power level with which an STA successfully receives a 256 byte packet with packet error rate (PER) of 10% [143]. This means that 10% of the time, some STAs may be hidden from other STAs in the group. If an STA detects only the preamble but is unable to decode the data frame, it will use the Response Indication Deferral (RID) to determine when to sense the channel again. In case of a data frame, the RID is set to a *Long Response* which means waiting for 27.84 ms [26]. If the STA

receives the ACK frame from the AP thereafter, it can reset the time. If the AP does not send an ACK because of an error, the STA will assume that the transmission is still on-going and will not make any attempt to access the channel even though the channel is idle. Therefore, when it detects the next ACK frame, it will cause a collision. To avoid such a scenario, the members of the groups created must satisfy the requirement of being able to detect the other members data frames for 99.9% of the time i.e. the PER=0.01.

5.4.1 Received Power Threshold

The bit error rate (BER), ρb , of a packet of length l bits is given by

$$\rho b = (1 - \rho p)^{\frac{1}{l}}, \tag{5.3}$$

where ρp is the PER. For uncoded binary phase shift keying (BPSK) modulation in an additive white Gaussian channel, the ρb is given by

$$\rho b = Q(\sqrt{2E_b/N_0}), \tag{5.4}$$

where $Q(\cdot)$ denotes the Q-function, E_b is the bit energy and N_0 is the noise density. Using (5.4), the required E_b/N_0 to meet the target BER is related to the received signal-to-noise ratio (SNR), denoted by Υ, as given by

$$\Upsilon = E_b/N_0 + 10 \cdot \log 10(\mathcal{R}/\mathcal{B}), \tag{5.5}$$

where $\mathcal{R}$ is the data rate and $\mathcal{B}$ is the system bandwidth. Using Υ, the minimum required receiver sensitivity to meet the desired reliability is given by

$$P_{sen}(\mathcal{B}, \mathcal{R}) = \Upsilon + 10 \cdot \log 10(N_0\mathcal{B}) + NF, \tag{5.6}$$

where NF is the noise figure. Equation (5.6) shows that $P_{sen}(\mathcal{B}, \mathcal{R})$ is both bandwidth mode and data rate dependent [143]. To cater for fast fading, a fading margin (FM) is added to $P_{sen}(\mathcal{B}, \mathcal{R})$. Therefore, the received power threshold, P_{th}, is given by

$$P_{th} = P_{sen}(\mathcal{B}, \mathcal{R}) + FM, \tag{5.7}$$

where P_{th} is the minimum power that a STA should receive from a neighbouring STA in order to consider it as a detectable STA.

5.4.2 Detection Graph

To determine whether or not there are hidden nodes in a given group of STAs, the AP needs to know whether the received power by each STA from other STAs is at least P_{th} or less than P_{th}. One of the ways to gather this information is for the AP to poll each STA in the group which will then respond with a null data packet (NDP) at a stipulated time. The STAs can determine the polled STAs from the POLL frame header and they will listen after a SIFS duration for the NDP. Each STA will determine if the received power from the NDP transmission is at least equal to P_{th} or not. Each STA then generates a detection map (similar to an interference map), ζ, which contains only ones and zeros as given by

$$\zeta_{i,j} = \begin{cases} 1 & p_{i,j} < P_{th} \\ 0 & p_{i,j} \geq P_{th} \end{cases}, \tag{5.8}$$

where $p_{i,j}$ denotes the power received by STA i from STA j, 1 indicates that STA i and STA j are hidden from each other while 0 indicates that they detect each other. The AP then polls each STA for its graph and then aggregates them to form one graph which contains all STAs in the group.

For a large number of STAs, to obtain the measurements of received power from all STAs in the sector would imply listening to up to 1000 STAs which are up to 1km away. Yet, for power-constrained STAs, the coverage distance may not exceed 150 m [27]. Therefore, power spent listening to transmissions farther away is wasted. Assuming that the distance between STAs is known, the AP can initially group STAs using the separation distance and then use the detection graph based on received power to regroup in order to eliminate hidden nodes. In this work, clustering is used for distance-based grouping [60, 62].

5.4.2.1 Reducing Group Formation Overhead - Clustering

Assuming that the distances between the STAs in each sector are known, they can be used to generate a proximity matrix which is then used to group the STAs using agglomerative clustering with Ward's method [60, 62]. The resulting dendogram is cut to produce a clustering with N_g clusters, where N_g is the target number of groups per sector. The AP will then poll the STAs in each group in order to retrieve the detection graph.

5.4.3 Grouping

The Welsh-Powell algorithm [144] (see Appendix B) is used to create groups free from hidden nodes. Given ζ, the first step involves determining the number of undetected STAs by each STA by summing up the elements in each row of the detection graph. To determine the allocation order, $\mathcal{A}_s$, the STAs are sorted in descending order of the number of undetected STAs. The first member of $\mathcal{A}_s$ is the first member of the first group, $\mathcal{G}_1$. If the second STA in $\mathcal{A}_s$ detects the first STA, it is added to $\mathcal{G}_1$, otherwise the next STA in $\mathcal{A}_s$ is considered. For an STA to be added to $\mathcal{G}_1$, it must detect all the STAs in $\mathcal{G}_1$. When the last STA in $\mathcal{A}_s$ is reached, the STAs in $\mathcal{G}_1$ form a group without hidden nodes. ζ and $\mathcal{A}_s$ are updated by deleting the rows and columns of the STAs in the group. The process is then repeated for the remaining STAs in ζ to from groups $\mathcal{G}_2$, $\mathcal{G}_3$ and so on, until ζ is empty, implying that all the STAs have been allocated to a group.

The Welsh-Powell algorithm creates non-uniform-sized groups with some groups being bigger than or smaller than a target group size. For a given target group size, g_t, the Welsh-Powell algorithm is modified to ensure that the size of the created groups does not exceed g_t. To achieve this, once a group's size is equal to g_t, no more STAs are added to the group and its members will be deleted from ζ and $\mathcal{A}_s$.

5.4.4 Merge-and-Split Algorithm

The goal of the algorithm is to merge groups when $\left|\mathcal{G}_j\right| < g_t$ where j denotes the group index and $|\cdot|$ denotes the number of elements in the group. However, if the merger of

two or more groups results in a group size greater than g_t, the smallest group will be split into two groups. Let $\mathbb{G}$ denote the set of all groups with $\left|\mathcal{G}_j\right| < g_t$. The groups are sorted in descending order of the number of STAs in the group. Take the first group, denoted by $\mathcal{G}_1$, the group with the highest number of STAs, and find a group in $\mathbb{G}$, $\mathcal{G}_j$, such that $|\mathcal{G}_1| + \left|\mathcal{G}_j\right| = g_t$. If a group $\mathcal{G}_j$ that meets the criteria is found, it is merged with group $\mathcal{G}_1$ and both groups are deleted from $\mathbb{G}$. If none of the group sizes meet the first criteria, then the smallest group in $\mathbb{G}$, denoted by $\mathcal{G}_{end}$, is the next candidate for merging. If $|\mathcal{G}_1| + |\mathcal{G}_{end}| < g_t$, then the two groups are merged and $\mathcal{G}_{end}$ is deleted from $\mathbb{G}$. However, if $|\mathcal{G}_1| + |\mathcal{G}_{end}| > g_t$, then $\mathcal{G}_{end}$ is split into two groups with one group having $g_t - |\mathcal{G}_1|$ STAs and the other having the remaining STAs. The group with $g_t - |\mathcal{G}_1|$ STAs is merged with $\mathcal{G}_1$ and both groups are deleted form $\mathbb{G}$ while the other group becomes the new $\mathcal{G}_{end}$ in $\mathbb{G}$. The process is then repeated until there are no groups in $\mathbb{G}$.

The merging algorithm is suboptimal, however it has a much lower complexity than an exhaustive search for groups or combinations of groups whose elements sum up to g_t. The purpose of the merging procedure is to create uniform sized groups so that the slot allocation is fair. However merging groups introduces hidden nodes in the group. This can be addressed if the STAs in the group are informed that the group is merged and that they can only access specific group slots or parts of the group slot. The STAs are informed of group members by listing their AIDs and then by sending a merged detection map. The detection map only contains STAs in the merged group. Suppose a merged group has two merged groups, $\mathcal{G}_1$ and $\mathcal{G}_2$, then the merged detection graph, ζ^{MG}, will be generated as follows

$$\zeta_{i,j}^{MG} = \begin{cases} 1 & i \in \mathcal{G}_1 \wedge j \in \mathcal{G}_2 \\ 0 & \{i,j\} \in \mathcal{G}_1 \vee \{i,j\} \in \mathcal{G}_2 \end{cases}, \tag{5.9}$$

where the $\wedge$ means 'AND' operator and the $\vee$ means 'OR' operator . $\zeta_{i,j}^{MG}$ takes the value of 1 when STAs i and j are not in the same group and takes a value of 0 when STAs i and j are in the same group (either $\mathcal{G}_1$ or $\mathcal{G}_2$).

5.5 Backoff Timer Assignment - AID-based Approach

It is proposed, that once the groups are formed and the received power graph for the STAs is updated, the AP sends the grouping frame to each group. The frame shall have AIDs of the group members along with the received power graph for the members. If the received power graph is a zero matrix, the STA will determine that there are no hidden nodes in the matrix. However if there are some ones in the matrix, the STA will determine that it is part of a merged group. It will then determine which STAs are detectable in the group based on the zeros in its row in the received power graph.

For the groups without any hidden nodes, the backoff time is set based on the AIDs of other group members [126]. If the AIDs of the STAs in a group are 1538, 1540 and 1541, they will set their backoff timers to values 1, 2 and 3 respectively. In other words, the backoff timer will be set in ascending order of AIDs of group members. This implies that during the group's allocated time slot, after a DIFS interval, STA 1538 will reduce its timer from 1 to 0 while the remaining STAs will also reduce their timers by one, and STA 1538 will then have access to the network. In case STA 1538 did not correctly decode the beacon and therefore does not access the channel, then STAs 1540 and 1541 will reduce their counters twice, and STA 1540 will gain access to the channel. This will eliminate the collisions in the network completely, and will also reduce the time wastage during the backoff countdown process. Once an STA is finished with its transmission, it should set its backoff counter to the number of STAs in the group.

For merged groups, the group slot will be accessed by both groups in proportion. For instance, if group $\mathcal{G}_1$ is a merged group with two subgroups $\mathcal{G}_{11}$ and $\mathcal{G}_{12}$ with 3 STAs and 1 STA respectively, then $\mathcal{G}_{11}$ and $\mathcal{G}_{12}$ will access the group slot with a 3:1 proportion. Therefore group $\mathcal{G}_{11}$ will access the group slot 3 times more than $\mathcal{G}_{12}$. STAs in the subgroup that is not active should freeze their backoff counters.

5.6 Performance Analysis

A saturated scenario dominated by uplink data transmission is assumed. Each STA sets its backoff timer value using the information in the received power graph as explained

in Section 5.5. The PERs for the *beacon*, *data* and *ACK* frames, denoted by ρp_b, ρp_d and ρp_a respectively, are calculated using (5.3). The duration of the *beacon*, *data* and *ACK* frames are denoted by t_b, t_d, and t_a, respectively. The elimination of both hidden nodes and random backoff timers using grouping and AID-based backoff means that there are no collisions in the network. However, not all transmissions will be successful because of packet errors caused by fading. The analysis used is similar to that in [4]. An STA will not be able to transmit during its group slot if it does not decode the beacon correctly. Finally, it is assumed that all groups formed have a uniform size, g.

5.6.1 Distribution of Number of Transmissions in a Slot

Let T_s denote the duration, in time slots, of the group slot during which a group may access the channel. A transmission refers to the transmission of a *data* frame whether successful or not. The number of transmissions within T_s is a random variable and is denoted by M. Let ϕ_s denote the duration of a successful data transmission excluding the backoff slots i.e. $\phi_s = (t_d + t_{sifs} + t_a + t_{difs} + 2 \cdot \wp)$ and ϕ_f denote the duration of a failed data transmission i.e. $\phi_f = (t_d + t_{difs} + \wp)$ where $\wp$ denotes the propagation delay, t_{sifs} is the duration of SIFS and t_{difs} is the duration of DIFS. The maximum number of successful transmissions, $\mathcal{M}_s$, that can take place in a group slot of duration T_s is given by

$$\mathcal{M}_s = \left\lfloor \frac{T_s}{\phi_s + 1} \right\rfloor. \tag{5.10}$$

In (5.10), the backoff time is equal to one time slot and $\phi_s + 1$ is the shortest duration of a successful *data* frame transmission. Assuming that there is at least one successfully received *data* frame in a group slot of duration T_s, the maximum number of transmissions, $\mathcal{M}_1(T_s)$, that can take place is given by

$$\mathcal{M}_1(T_s) = \mathcal{M}_s + I_{\left\{T_s \geq \left(\phi_s + 1 + \mathcal{M}_s(\phi_f + 1)\right)\right\}}, \tag{5.11}$$

where $I_{\{q\}}$ is an indicator function such that $I_{\{q\}} = 1$ when the condition q is true; otherwise, $I_{\{q\}} = 0$. This is derived from the fact that $\phi_s > \phi_f$. Therefore, the shortest

duration for i transmissions is $i(\phi_f + 1)$ which means that there are i failed transmissions and each has a backoff duration of one slot. Depending on T_s, after i failed transmissions, the time left may be long enough for one more successful transmission. Because of the no-crossing rule, if the duration of the remaining time is shorter than the duration of a successful data transmission plus the backoff duration of the STA with the shortest backoff timer value, then there will be no extra transmission and $I_{\{q\}} = 0$.

The longest duration of a *data* frame transmission occurs when there is a successful data transmission ($\phi_s > \phi_f$) and only one STA in the group is transmitting, and it is the last STA in the queue. This implies that the backoff time will be equal to g slots, where g is the group size. Therefore, the minimum number of transmissions that can take place in a group slot of duration T_s is given by

$$\mathcal{M}_2(T_s, g) = \left\lfloor \frac{T_s}{\phi_s + g} \right\rfloor. \tag{5.12}$$

For convenience, $\mathcal{M}_1(T_s)$ is denoted by $\mathcal{M}_1$ and $\mathcal{M}_2(T_s, g)$ by $\mathcal{M}_2$ in the rest of this section. Depending on g and T_s, $\mathcal{M}_1$ and $\mathcal{M}_2$ may or may not be equal. Therefore, M, the number of transmissions in T_s, is given by

$$M = \begin{cases} \{0, \mathcal{M}_1, \mathcal{M}_2\}, & \mathcal{M}_1 > \mathcal{M}_2 \\ \{0, \mathcal{M}_1\}, & \mathcal{M}_1 = \mathcal{M}_2 \end{cases}, \tag{5.13}$$

where the 0 represents the case where all the STAs in the group fail to decode the beacon and therefore cannot transmit in the allocated group slot.

Let $T_{t,m}$ denote the duration, in time slots, of m transmissions and let $T_{b,i}$ denote the number of backoff slots for the ith transmission in a group slot. Then,

$$T_{t,m} = \sum_{i=1}^{m} \Phi_i + T_{b,i} = \sum_{i=1}^{m} \Phi_i + \sum_{i=1}^{m} T_{b,i}, \tag{5.14}$$

where Φ_i is a random variable with two possible values: ϕ_s and ϕ_f and $T_{b,i}$ is a random variable with values $\{1, 2, \cdots g\}$. $\sum_{i=1}^{m} \Phi_i$ depends on the number of successful transmissions and the number of failed transmissions. The distribution of the number of

successful transmissions is binomial and the probability of success is $(1-\rho p_d)$. On the other hand, $T_{b,i}$ is geometrically distributed with success probability $\left(1-\rho_b^p\right)$.

The conditional distribution of the number of successful events, S, in a group slot given g is given by

$$P_{S|G}(S = s|g) = (1 - p_0) \times \sum_{m=\mathcal{M}_1}^{\mathcal{M}_2} P_{S|MG}(S = s|mg), \tag{5.15}$$

where p_0 is the probability that none of the STAs in the group received the beacon frame and is equal to $(\rho p_b)^g$. Moreover, $P_{S|MG}$ is the probability distribution for the number of successful transmissions given the number of transmissions in a group slot and the group size, and it is given by

$$P_{S|MG}(S = s|mg) = \begin{cases} C_s^m (\rho p_d)^{(m-s)} (1-\rho p_d)^s P_{B|MGS}, & m \in M, \\ 0, & \text{otherwise}, \end{cases}, \tag{5.16}$$

where, $P_{B|MGS}$ is the probability distribution of the sum of the backoff slots, B, given m, g and s, and C_s^m is the binomial coefficient given by

$$C_s^m = \frac{m!}{(m-s)!s!}. \tag{5.17}$$

For a given m, g and s, the number of backoff slots, b, that can be accommodated is calculated using

$$b = T_s - s\phi_s - (m-s)\phi_f. \tag{5.18}$$

The probability $P_{B|MGS}$ has the following distribution

$$P_{B|MGS} = \begin{cases} P_{B|MGS}(B \le b \mid mgs), & m \le b \le \phi_s + 1, \\ P_{B|MGS}(B \ge b - \phi_s \mid mgs), & b - \phi_s \le m \cdot g, \\ 0, & \text{otherwise}, \end{cases}. \tag{5.19}$$

The first case of (5.19) states that if b is greater than the minimum number of required slots (m for m transmissions) and it is less than the minimum time for another transmission ($\phi_s + 1$), then it is possible to have m transmissions if the probability that $P_{B|MGS}(B \leq b \mid mgs)$ is greater than zero. The second case of (5.19) states that if $b \geq \phi_s + 1$ then the only way that there will be m transmissions is if $B \geq b - \phi_s$. In other words, B should be large enough to reduce the remaining time after transmission to a value less than $\phi_s + 1$.

The sum of backoffs for a given number of events B, is a random variable because of ρp_b. To determine the distribution of B, the way backoff counters are set is examined. The backoff counters of active STAs are set to values ranging from 1 to g. Once an STA completes its transmission, it sets its backoff value to g. Therefore, at the point where $B \geq g$, all active STAs would have transmitted a data frame. Further, if some or all STAs can transmit more than one data frame, the backoff slot pattern will be the same as in the first cycle. To illustrate this, let $g = 4$, $m = 4$ and let the backoff duration pattern be $[1, 3]$ for the first two transmissions. It should be fairly obvious that the pattern when the next two transmissions take place is $[1, 3, 1, 3]$. This means that for each group size, there is a definite set of backoff slot patterns that are feasible. Let $\mathbf{V}$ denote a matrix with the viable patterns for a given group size. For $g = 2$, $\mathbf{V}_2^2$ is given by

$$\mathbf{V}_2^2 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 2 \end{bmatrix}. \tag{5.20}$$

When $m > g$, the new $\mathbf{V}_m^g$ is generated by repeating the patterns in the previous columns. For example, if $m = 4$ and $g = 2$ then $\mathbf{V}_4^2$ is given by

$$\mathbf{V}_4^2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \end{bmatrix}. \tag{5.21}$$

The second row seems strange though it depicts the scenario where only one of two

STAs decodes the beacon and it is the first one to transmit in the group. After the initial transmission, it has to wait two backoff slots each round in order to transmit because the second STA cannot transmit. If $m < g$ e.g if $m = 1$, then $\mathbf{V}_1^2$ is given by

$$\mathbf{V}_1^2 = \begin{bmatrix} 1 \\ \\ 2 \end{bmatrix}. \tag{5.22}$$

Having generated $\mathbf{V}_m^g$, $P_{B|MGS}$ can easily be determined. For each row in $\mathbf{V}_m^g$, denoted by $\mathbf{v}_i$, sum up the values in the row and only keep the values that meet the required value of B. The new matrix, $\bar{\mathbf{V}}_m^g$ is given by

$$\bar{\mathbf{V}}_m^g = \begin{cases} \mathbf{v}_i \mid \sum \mathbf{v}_i \leq b, & m \leq b \leq \phi_s + 1, \\ \mathbf{v}_i \mid \sum \mathbf{v}_i \geq b - \phi_s, & b - \phi_s \leq m \cdot g, \end{cases}. \tag{5.23}$$

For each row, $\bar{\mathbf{v}}_i$, in $\bar{\mathbf{V}}_m^g$, the number of STAs that received the beacon, $n_{o,i}$, and those that did not receive the beacon, $n_{d,i}$, can be determined. Therefore, $P_{B|MGS}(B \leq b)$ will be given by

$$P_{B|MGS}(B \leq b \mid mgs) = \sum_{i=1}^{R(\bar{\mathbf{V}}_m^g)} (\rho p_b)^{n_{d,i}} (1 - \rho p_b)^{n_{o,i}} \ \forall \bar{\mathbf{v}}_i \ , \tag{5.24}$$

where, $R(\cdot)$ means rows in $(\cdot)$. The probability $P_{B|MGS}(B \geq b \mid mgs)$ will also be given by (5.24) and the only difference is that $\bar{\mathbf{v}}_i$ will contain rows where the summation of values is less than b.

5.6.2 Throughput per Group

Using the equations derived in Section 5.6.1, the expected number of successful transmissions for a group of size $g \geq 2$ STAs within a group slot, $\mathbb{E}_{S|G}$, is given by

$$\mathbb{E}_{S|G}(s \mid g) = \sum_{s=1}^{\mathcal{M}_1} \sum_{m=\mathcal{M}_2}^{\mathcal{M}_1} s \cdot P_{S|MG}(S = s|mg) \tag{5.25}$$

The normalised network throughput is given by

$$Th(g) = \frac{\mathcal{L}}{T_{tot}} \cdot \frac{N_u}{g} \cdot \mathbb{E}_{S|G}(s \mid g) \tag{5.26}$$

where $\mathcal{L}$ is the transmission time for the payload, T_{tot} is the total time for all groups to access the channel, and $\mathcal{L}$ is given by

$$\mathcal{L} = \frac{Q_d \times 8}{\mathcal{R}}, \tag{5.27}$$

where Q_d is the payload in bytes and $\mathcal{R}$ is the data rate. Moreover, T_{tot} is given by

$$T_{tot} = T_s \cdot \frac{N_u}{g} + \kappa \cdot t_b, \tag{5.28}$$

where κ is the number of sectors.

5.7 STA-assisted Packet Transmission

In this section, STA-assisted packet transmission is used to reduce outage in an 802.11ah network which utilises AID based backoff. The MAC header for frames sent from the Relay STA and outage STA will have six extra bytes compared to other frames in the network. As shown in Figure 5.2, the MAC header has an optional third address (A3) that is used to identify the final destination of the packet from the outage STA and to identify the outage STA to the AP [127]. This implies that the frame transmission time is longer than for other packets when the transmission rate is kept constant. To avoid having to increase the group slot sizes in order to cater for the longer transmission times, implicit acknowledgement is used for frames to be forwarded to the AP [26]. This means that when a packet is received by the relay STA, instead of sending an ACK message to the outage STA, it sends the packet to the AP. Once the outage STA detects the packet transmission to the AP, it will implicitly know that the packet was successfully received by the relay STA.

The association procedure is presented in Figure 5.3 and described in Section 5.7.1.

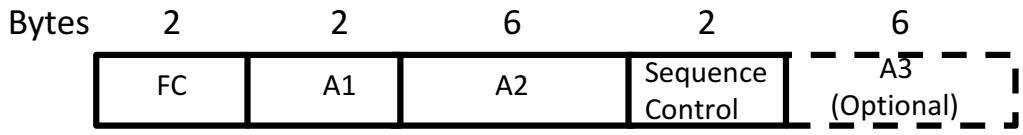


Figure 5.2: 802.11ah MAC header format

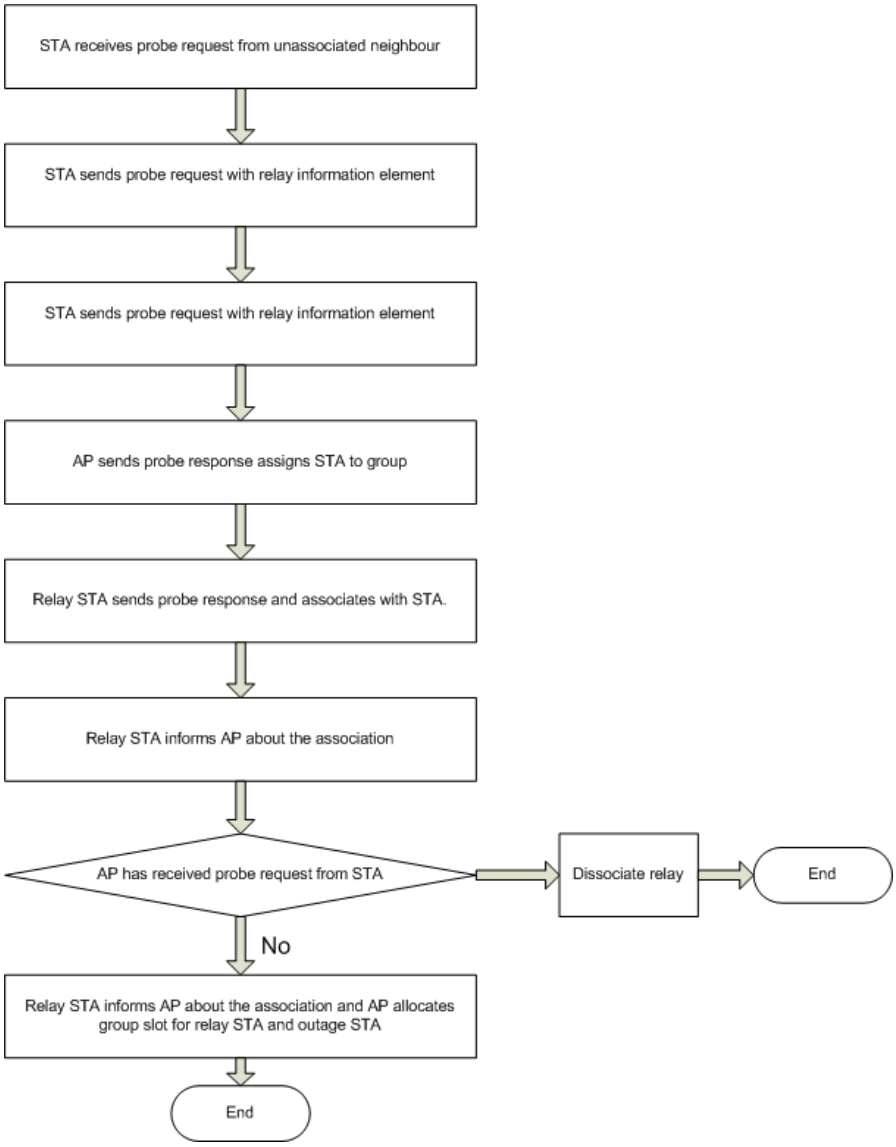


Figure 5.3: Association Process

5.7.1 STA discovery procedure

The STAs are assumed to perform active scanning whereby they send probe requests and then wait for responses from APs that receive the probe responses. After a timeout period, if no probe response is received, the STA sends another probe request. After a number of timeouts due to absence of a probe response from any AP, the STA may then send a probe request with relay discovery element. This is meant to discover relays in the STAs neighbourhood. The neighbouring STAs determine that the STA that is sending probes is in outage based on the fact that the probe requests have a relay discovery element. At this stage, it is proposed that the STAs receiving a good signal from the outage STA request the AP to operate in relay/STA mode. This will enable the AP to allocate a time slot for the data exchange between the outage STA, the relay STA and the AP. This work proposes that at this juncture, the STA uses network virtualisation to become a relay and send a probe request to the AP. The AP will then send a probe response and complete the rest of the association procedure, including giving the relay an AID. The newly associated relay STA will then respond to the probe request of the outage STA. Once the association is completed, the outage STA can then inform the AP that it can be reached through the relay STA using the address resolution protocol (ARP) as explained in [26]. Once this is completed, all packets destined for the STA will be relayed through the relay STA. This means during the beacon interval, the AP will allocate time for the relay to receive data from the outage STA and then relay it to the AP.

To quicken the association process, this work proposes that the relay STA uses the first time slot allocated to it for sending the probe response. Because of AID-based backoff and the grouping approach, there will be no chance of collisions and this will hasten the association process.

5.7.2 Grouping update

Once the association is done, the AP updates the grouping information. In this case, it will allocate the relaying STA and the outage STA to the same group. If the group contains other STAs, the group update should make the relay STA and the outage STA orthogonal to those UEs. This means that the other group STAs will not be able to use the slots allocated to the relay STA and outage STA.

5.7.3 Relaying procedure

The sectorisation IE in the beacon contains the information on when the relay STA and outage STA can access the channel. The relaying procedure will follow normal DCF procedure where by first the outage STA will send a packet to the relay STA, which the relay packet will acknowledge. In the next slot, the relay STA will forward the received packet to the AP. The AP will determine which user the packet is from by examining the MAC header of the received packet i.e. the third address in the MAC header informs the AP that the packet was forwarded.

5.8 Performance Evaluation of AID-based Backoff

5.8.1 Simulation Setup

The parameters used for the simulations are presented in Table 5.1. STAs are uniformly distributed in a 1 km radius circle with the AP located in its centre. Both the control frames (i.e., *beacon*, *ACK*) and *data* frames are OFDM-modulated [47] with OFDM symbol parameters presented in Table 5.1. Although the FFT length is 32 points, only 24 subcarriers are used for data modulation. Both data and contol packets are BPSK modulated with 1/2 rate low density parity bit coding meaning that each OFDM symbol carries 12 bits. The *beacon* and *data* packet duration are calculated using

$$\left(t_b = \left\lceil \frac{(Q_b \times 8 + 22) + \mathcal{S}_{ie}}{12} + \mathcal{H}_p \right\rceil\right) \times T_o, \tag{5.29}$$

$$\left(t_d = \left\lceil \frac{(Q_d \times 8 + 22) + \mathcal{H}_m \times 8}{12} + \mathcal{H}_p \right\rceil\right) \times T_o, \tag{5.30}$$

where the 22 bits are for the tail and service bits, $\mathcal{H}_p$ is the PLCP header size in the OFDM symbols, $\mathcal{H}_m$ is the MAC header in bytes, T_o is the OFDM symbol duration, Q_b and Q_d are the bytes in the beacon and data frames, respectively. Moreover, $\mathcal{S}_{ie}$ are the sectorisation IE bits with 31 bits for information fields and the group bits which depend on the number of groups in a sector. The frame check sequence (FCS) is 4 bytes in length. Only STAs with $\rho p_d \leq 0.1$ transmit data as per the 802.11ah standard.

Table 5.1: Simulation Parameters

Parameter	**Value**	**Parameter**	**Value**
Bandwidth	1 MHz	Noise density	-174 dBm/Hz
Carrier frequency	0.9 GHz	Noise figure	3 dB
AP power	30 dBm	Fade margin	5.6 dB
STA power	10 dBm	Shadowing	8 dB
FFT length	32	CW_{min}	16
Transmitter gain	0 dB	Propagation delay	3 μs
Receiver gain	3 dB	CW_{max}	1024
OFDM symbol duration	40 μs	Beacon size	36 Bytes
SIFS	160 μs	Packet size	256 Bytes
PIFS	212 μs	ACK size	14 symbols
DIFS	264 μs	PHY header size	14 symbols
Slot time	52 μs	MAC header size	12 Bytes
CF-END size	20 Bytes	POLL size	26 Bytes
NDP size	14 symbols	POLL + ACK size	26 Bytes

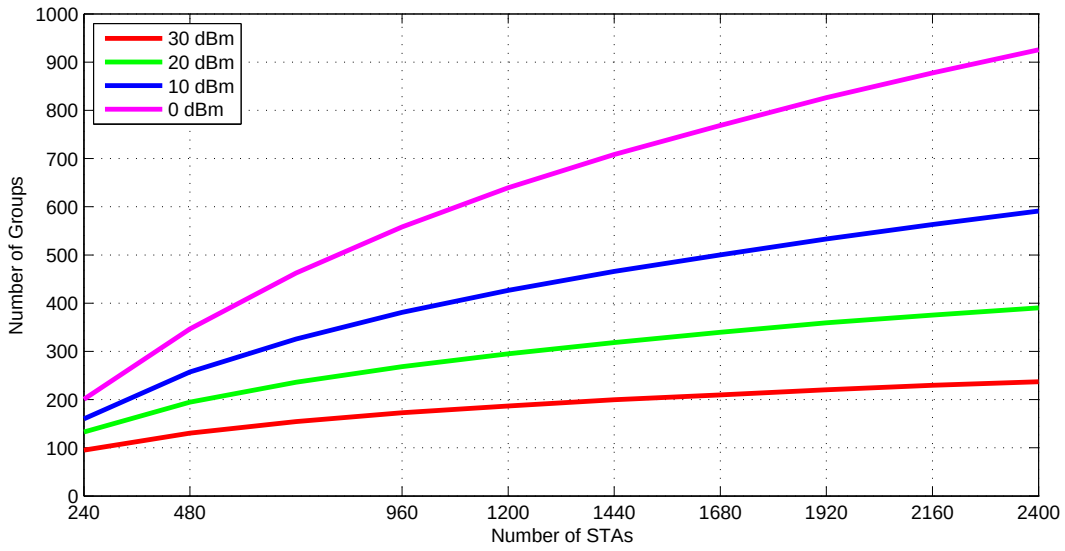


Figure 5.4: Group variation with transmit power

5.8.2 Grouping Results

This section presents the performance of the grouping.

Fig. 5.4 shows the variation of the number of groups with the number of STAs for different STA transmit powers. The lower the transmit power, the higher the number of groups created because the STA coverage is reduced, and thus more STAs become hidden, which leads to formation of smaller groups. As the number of STAs increase, the number of groups increase non-linearly. This is due to the fact that as the number of STAs increases, there will be more STAs in the coverage area of an STA, which results in the formation of bigger groups.

Fig. 5.5 shows the impact of the system bandwidth on the number of groups formed for a fixed transmit power. The higher the bandwidth, the higher the number of groups formed for all transmit powers. This is because the higher bandwidth the higher the required receiver sensitivity, as seen in (5.6). The higher sensitivity reduces the number of STAs that an STA can detect, and thus increases the groups created.

To examine the cost of partitioning users in distance-based groups for received power measurements, the groups created when there is no partitioning (NP) are compared to those created when there is partitioning (P). For partitioning, clustering is used as explained in Section 5.4.2.1 and in each case the STAs are divided into 10 clusters i.e.

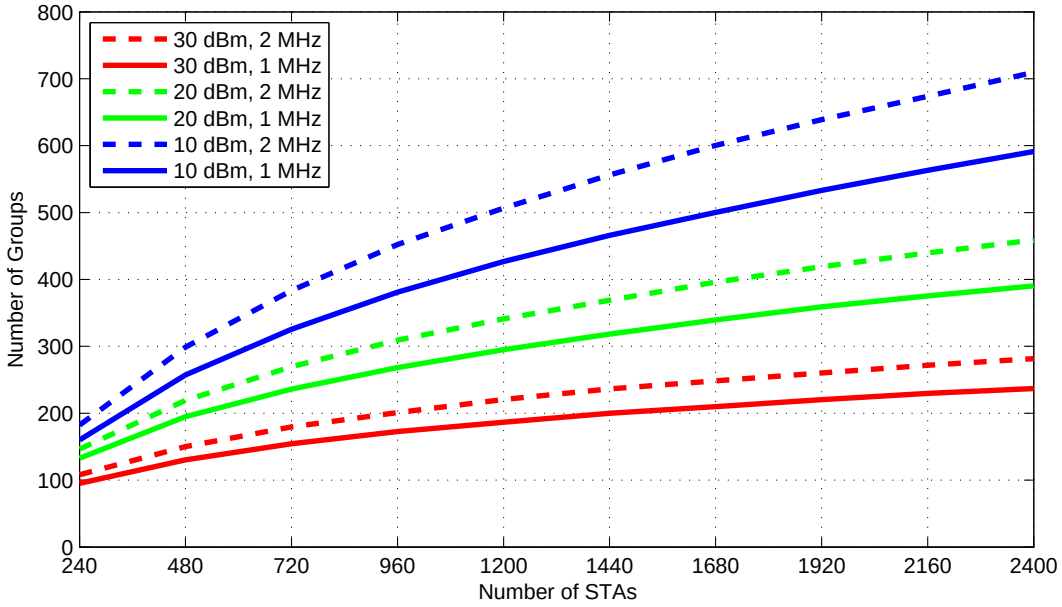


Figure 5.5: Group variation with transmit bandwidth

$N_g = 10$. Fig. 5.6 shows that the number of groups created when users are partitioned is slightly higher than when no partitioning is used. This is because partitioned STAs can only inform the AP about detectable STAs in their partition and therefore, cannot form groups with detectable STAs in another partition. This slight increment in the number of groups can be traded off for the energy cost of the no partition case where an STA would have to develop a received power graph for all STAs in its sector.

Fig. 5.7 demonstrates the cumulative distribution function (CDF) of groups formed for $N_u = \{240, 1200, 2400\}$ STAs and $P_t = \{10 \text{ dBm}, 30 \text{ dBm}\}$. It is shown that groups created are not of uniform group size irrespective of the transmit power or the number of STAs. It is further shown that the group sizes for lower transmit power are smaller which results from reduced STA coverage.

Fig. 5.8 illustrates that smaller group sizes increase the number of groups, as expected. For $N_u < 2160$ STAs, the number of groups created when the maximum group size, $g_t = 16$ is nearly the same as when there is no group restriction. This is because all the created groups have 16 or less STAs and therefore the restriction group size will not increase the number of groups that would have been created if there was no group size restriction.

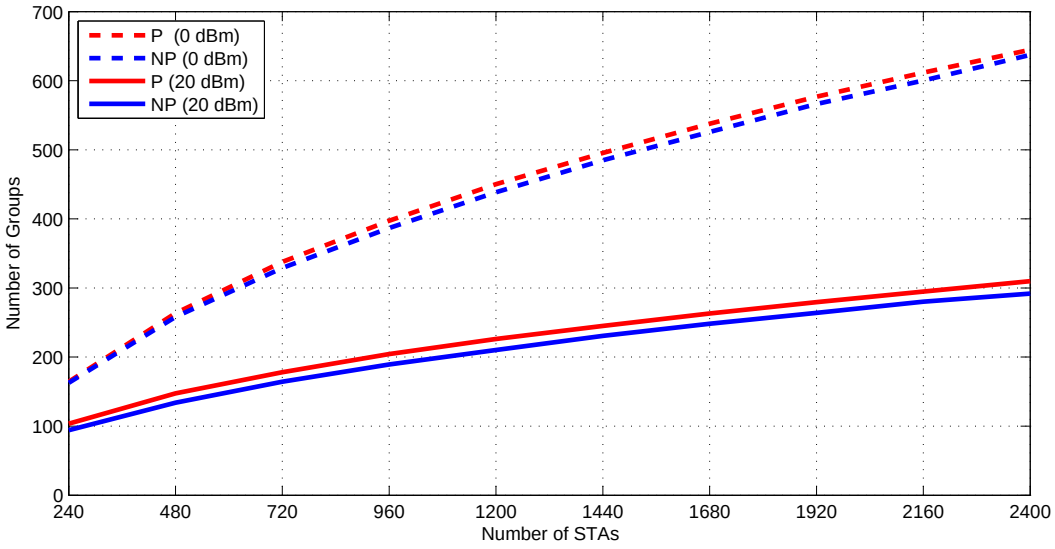


Figure 5.6: Number of groups

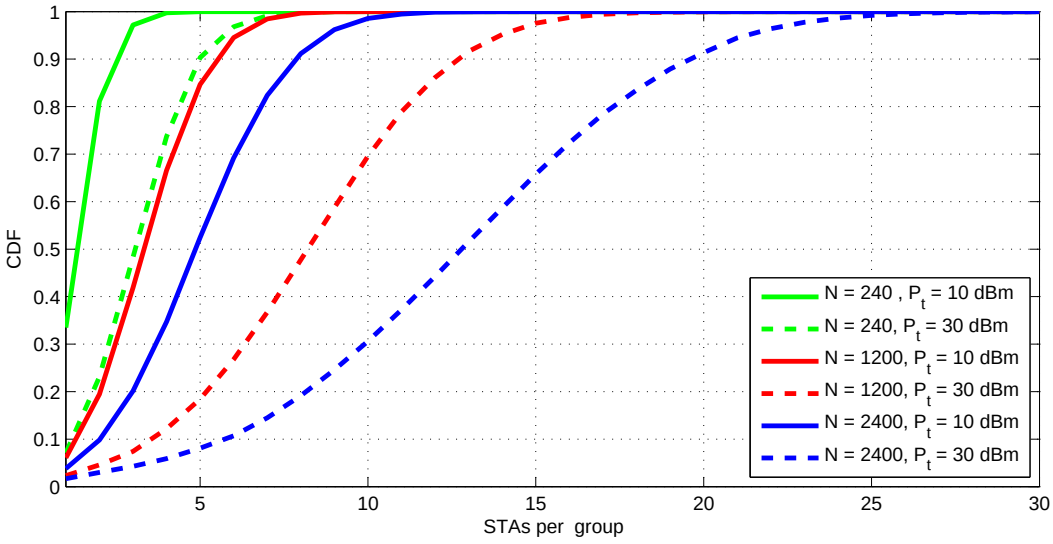


Figure 5.7: CDF of group size

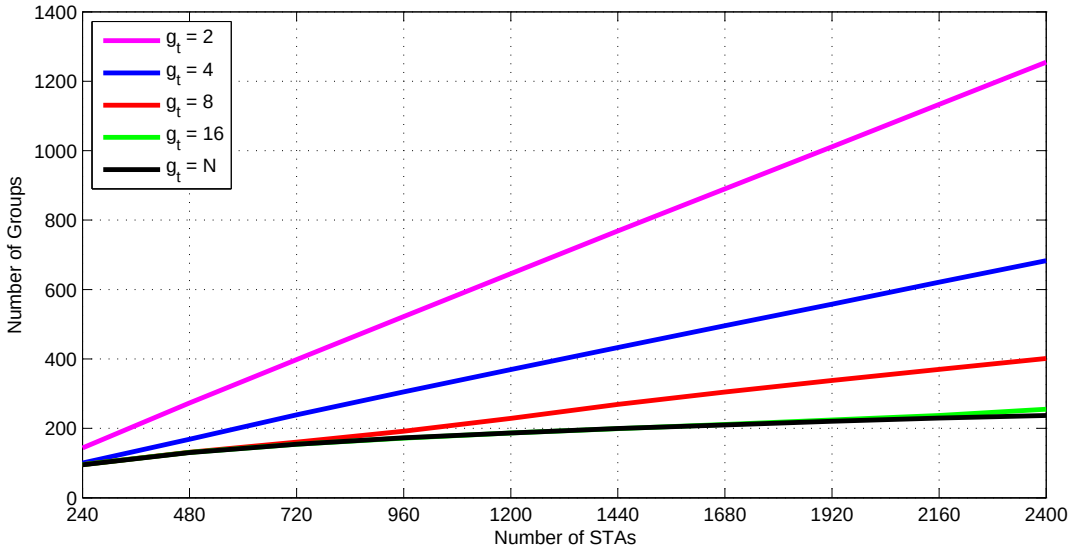


Figure 5.8: Impact of group size restriction ($P_t = 30\,dBm$)

Fig. 5.9 presents the CDF for the resulting groups when the group size restriction is applied, i.e. $g_t = \{2, 4, 8\}$, for 2400 STAs. The results show that for $P_t = 30$ dBm, at least 97% of the groups created are of size g_t for all g_t. However, for $P_t = 10$ dBm, 98% of the groups are of the required size when $g_t = 2$ which drops to 32% when $g_t = 8$. This is because the low transmit power reduces STA coverage making it difficult to form larger groups. The lack of uniformity in group size makes the determination of group slot duration complicated. Fair resource allocation would require variable group slot sizes to match the variable group sizes. However the 802.11ah standard does not support this. This is why the merging of groups to ensure that all groups have the same size was proposed.

5.8.3 Model Validation

To validate the proposed throughput model, the analytical and simulation results are compared for $N_u = \{240, 2400\}$ STAs with $g_t = \{2, 8, 16\}$ and the transmit power is 10dBm. STAs with $\rho p_d > 0.1$ are not included in the simulation because the 802.11ah standard defines the maximum PER as 10% for 256 byte packets. For the analytical evaluation, the expection of E_b/N_0 denoted by $\mathbb{E}(E_b/N_0)$ (see derivation in Appendix A) is used to determine ρp_b and ρp_d. The group size is limited to 16 for two reasons: first, the grouping results show that for $P_t = 10$ dBm, none of the groups created before

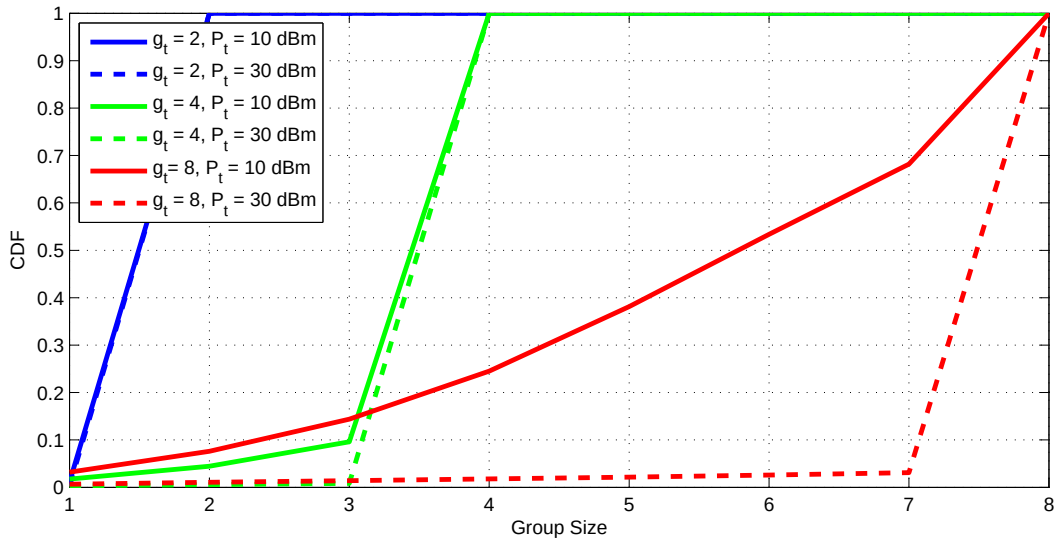
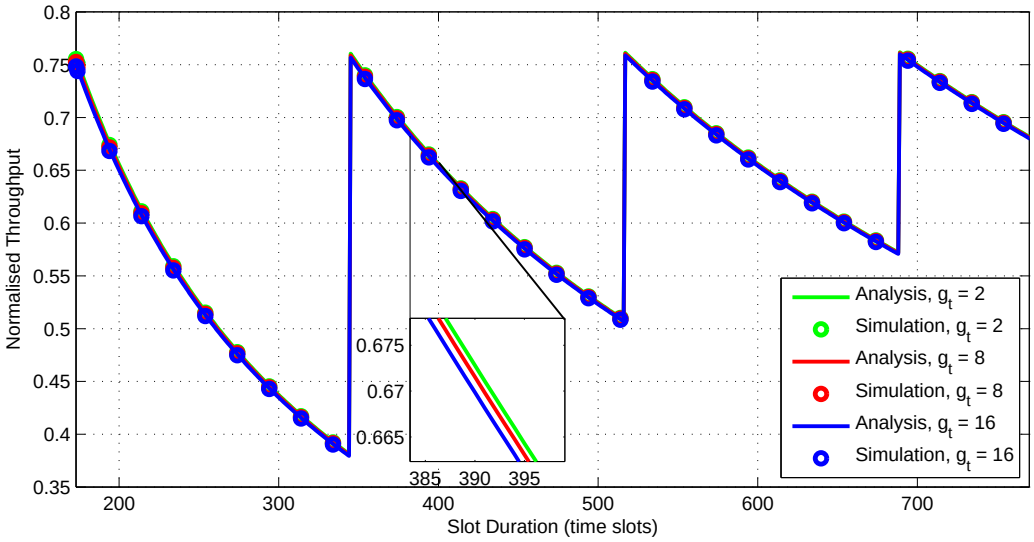


Figure 5.9: CDF of group size

the merging process has more than 16 STAs.

Figs. 5.10 and 5.11 show that the simulation results and the analytical model results match quite well for all group sizes. The normalised throughput fluctuates as the group slot size is increased. This is because the increase in slot duration increases idle time as it is too short for successful transmission of another packet before the group slot boundary. Peak throughput occurs at multiples of $(\phi_s + 1)$ because ρp_b and ρp_d are very small; hence, for most transmissions, successful packet transmission takes place after one backoff slot. For both 240 and 2400 STAs, $g_t = 2$ has the highest normalised throughput. This is because, for any group slot duration, the smaller group size will send more packet transmissions every sector interval before the sectorisation beacon is sent. Its impact is more pronounced when the number of STAs is small as shown in Fig. 5.11 when $N_u = 240$. Assuming that there are 12 groups for $g_t = 2$ and 3 groups for $g_t = 8$, if the slot duration is big enough for only one STA at a time, then to transmit 12 packets when $g_t = 2$, the beacon will be sent once yet when $g_t = 8$, it will be sent 4 times because 3 packets will be sent during each sector interval.



(a)

Figure 5.10: Throughput variation with slot duration for 2400 STAs

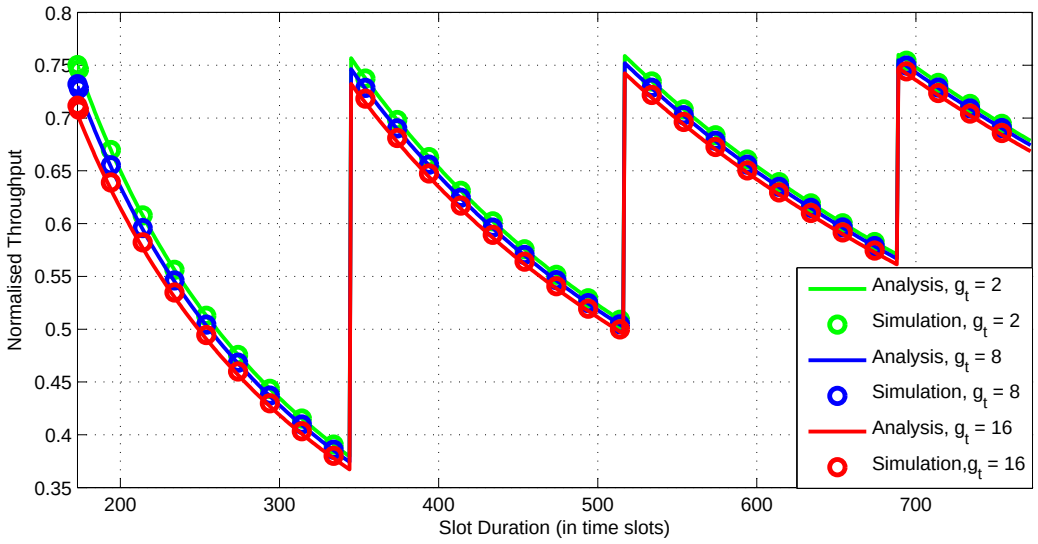


Figure 5.11: Throughput variation with slot duration for 240 STAs

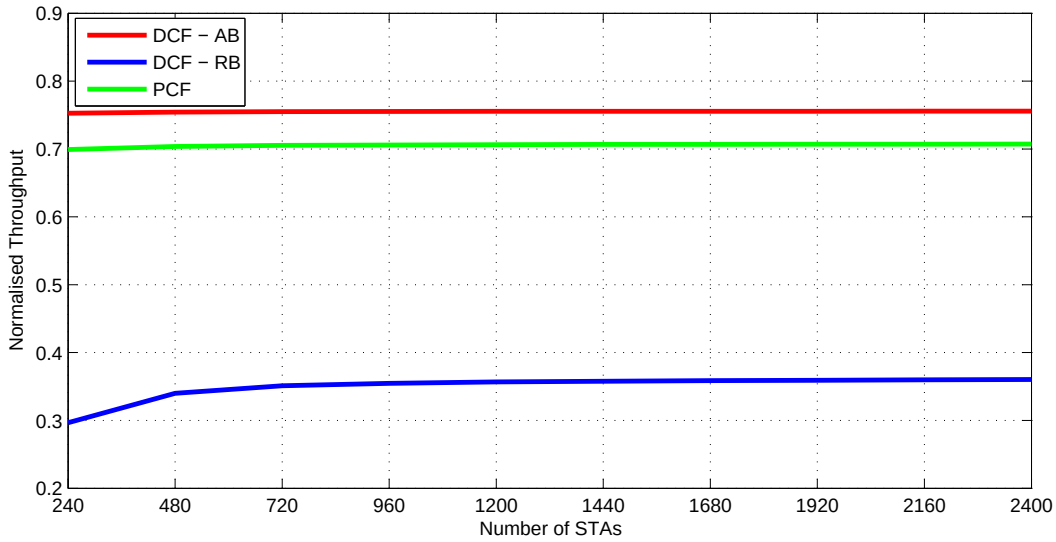


Figure 5.12: Normalised throughput

5.8.4 AID-Based Backoff Performance

In this section, the performance of DCF with AID-based backoff (DCF-AB) is compared to conventional DCF with random backoff (DCF-RB) [4, 24] and the point coordination function (PCF). For DCF-AB and DCF-RB, the grouping technique in Section 5.4 is used to group STAs while no grouping is done for PCF because STAs have to be polled before they can transmit a packet. For all the results, $g_t = 2$ and the time slot duration for DCF-RB and DCF-AB is $T_s = (\phi_s + 1) \times g_t$ slots. This is chosen because during each sector interval, every STA in the group has an opportunity to transmit a packet. Additionally, with each new sector interval, STAs have the opportunity to reset their backoff timers.

Fig. 5.12 shows that DCF-AB has the highest normalised throughput which is 40% more than that attained by DCF-RB. The gain is attributed to elimination of collisions and reduction of idle slots due to backoff and the no-crossing policy in DCF-AB. On the other hand, a large CW_{min} coupled with the no-crossing policy reduces resource efficiency for DCF-RB. PCF throughput is 7% lower than DCF-AB because it has a slightly higher overhead that results from the fact that the control frames used like the POLL+ACK, POLL and CF-END frames cannot be replaced by NDP frames.

Fig. 5.13 shows that the head of line (HOL) delay increases with the number of STAs

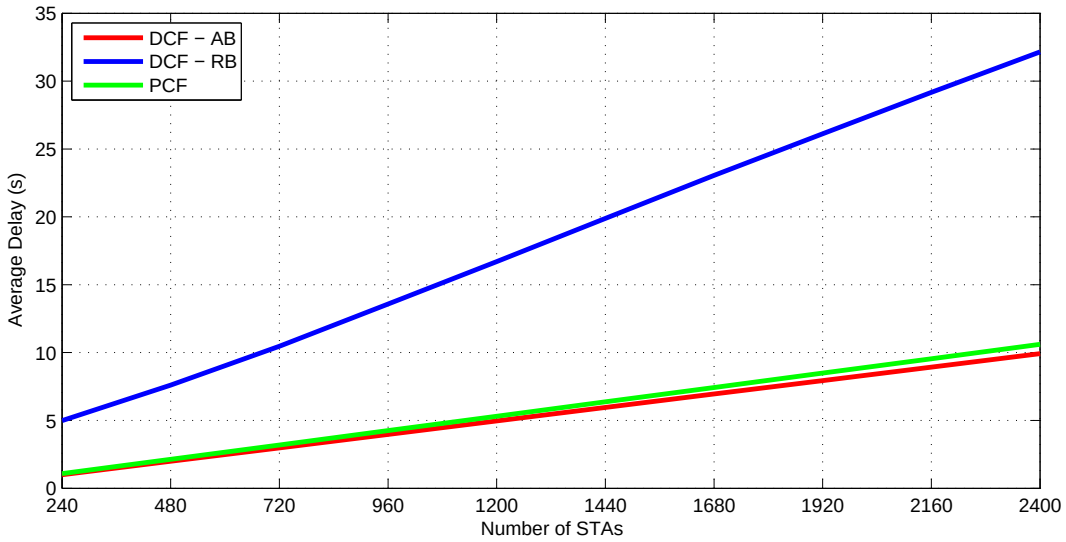


Figure 5.13: HOL delay

for all techniques. This is because there are more STAs to be served in each sector interval, which increases the amount of time it takes for the AP to make a complete rotation of all the sectors. The elimination of collisions and minimisation of idle time due to backoff reduce the DCF-AB HOL delay by 58% compared to DCF-RB. PCF also suffers no collisions but has a slightly higher delay than DCF-AB arising from its higher overhead.

Fig. 5.14 shows the average energy utilised per packet transmitted for each of the techniques. The power utilised in the Transmit, Receive/Sense and Sleep states is 255 mW, 135 mW and 1.5 mW, respectively [145]. DCF-AB utilises the lowest energy per transmitted packet because it has the highest throughput and STAs sleep outside their group slot. The STAs in DCF-RB sleep for the same duration as STAs in DCF-AB. However, because of the longer delay experienced by the packets, their energy utilisation is higher. PCF has the highest energy utilisation per successful packet because the STAs do not sleep until they are polled, which means that they spend a much longer time in the Receive/Channel Sense state than STAs in the other techniques.

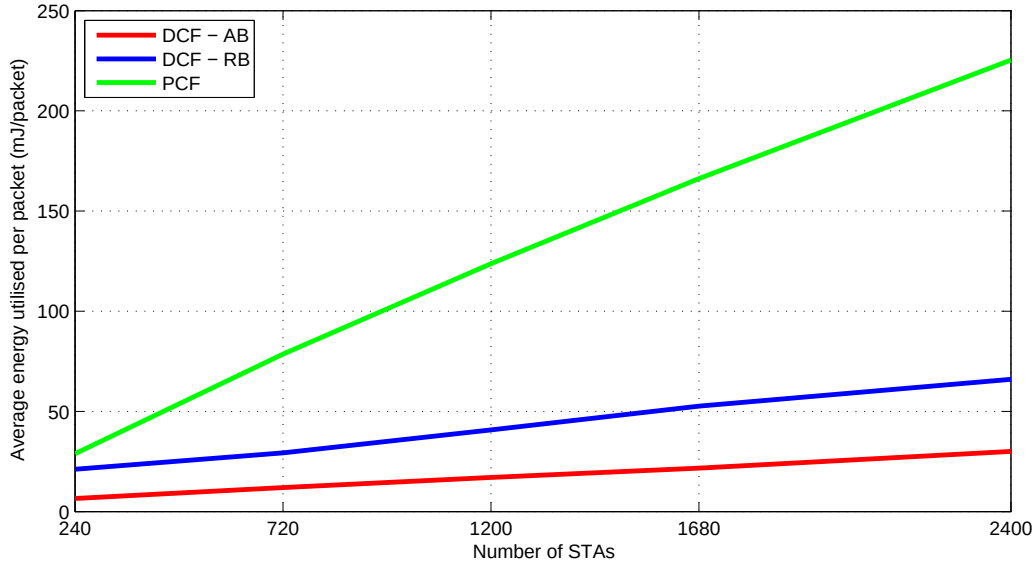


Figure 5.14: Energy utilisation per transmitted packet

5.9 Performance Evaluation of STA Relays

The performance of the proposed technique is compared to three approaches i.e. the reliability set by the standard which is 90% reliability for a 256 byte packet, reduced reliability and reduced packet size [145]. For reduced reliability, the reliability is set at 60% i.e. maximum packet error rate is set at 0.4 while for reduced packet size, the reliability is at 90% but if an STA cannot meet that requirement for 256 Byte packet but meets it for a 128 Byte packet, then it will be allowed to connect to the AP but send a smaller packet. Finally, the performance of a hybrid technique which involves reducing the reliability to 60% and connecting the remaining STAs through the relay STA is also compared.

Figure 5.15 shows the percentage of STAs in outage for all the techniques. The standard's 90% reliability results in the highest outage followed by the reduced packet size and reduced reliability of 60%. The percentage of STAs in outage does not vary with the number of STAs for these three techniques. However for the relay STA and hybrid techniques, the percentage STAs reduce from 0.6 – 0.9% for 240 STAs to 0 for 2400 STAs. The lowest outage is attained by the hybrid technique followed closely by using the relay STA. The small outage for 240 STA is highest for the two techniques

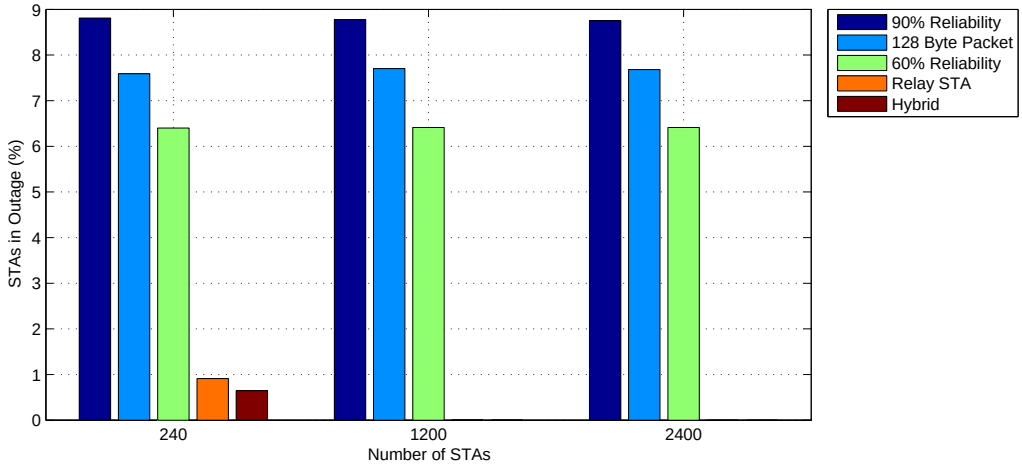


Figure 5.15: Percentage of STAs in outage

because some STAs fail to pair up with neighbouring STAs and hence remain in outage. Notably, once the number of STAs increases and the separation distance between neighbouring STAs reduces, all outage STAs are paired up leaving none in outage.

Figure 5.16 shows the normalised throughput of the techniques. The throughput is highest with 90% reliability and lowest with the relay STAs which result in a throughput reduction of 7%. This is because for a packet to be received by the AP from the outage STA, two STA slots are used as opposed to one. This is an additional overhead which results in reduced throughput. 60% reliability results in higher throughput than the 128 Byte packet. This is because on average an STA transmitting a 256 byte packet transmits more data with 60% reliability than a 128 byte packet with 90% reliability. The hybrid technique which uses both reduced reliability and relay STAs has a higher throughput than relay STAs only. This because fewer STAs will need to use two group slots to transmit one packet through the relay STA. The reduction in throughput could be remedied if the relay STA and outage STA use another channel for communication. Therefore, the results shown here represent the worst case where the same channel is used for all communications.

Figure 5.17 shows the HOL delay for the techniques. The 90% reliability, 60% reliability and the 128 byte packet size all have nearly the same delay for packets transmitted unlike the hybrid and relay STA. This is due to the longer time it takes to transmit a packet coupled with the increase in number of STAs which means that the beacon period becomes longer and it will take longer for the AP to allocate the next group slot.

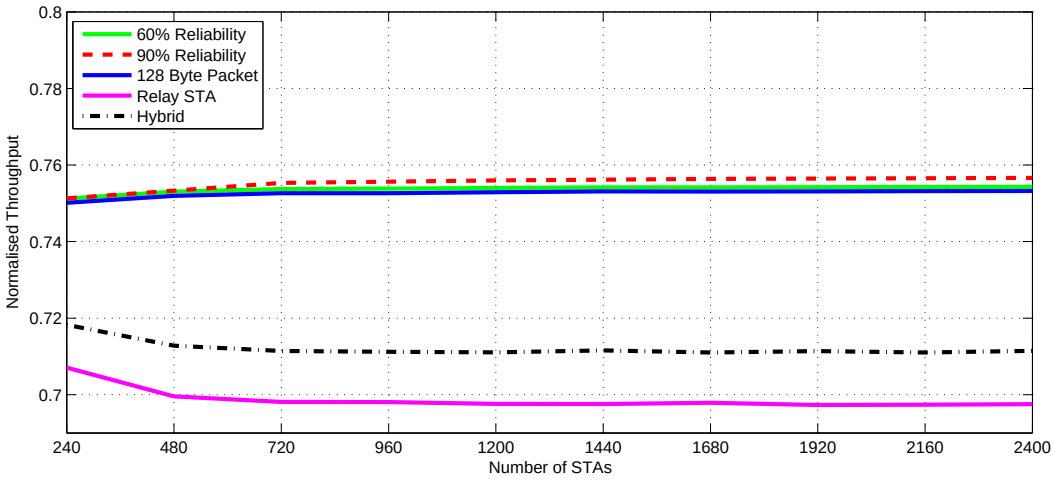


Figure 5.16: Normalised throughput

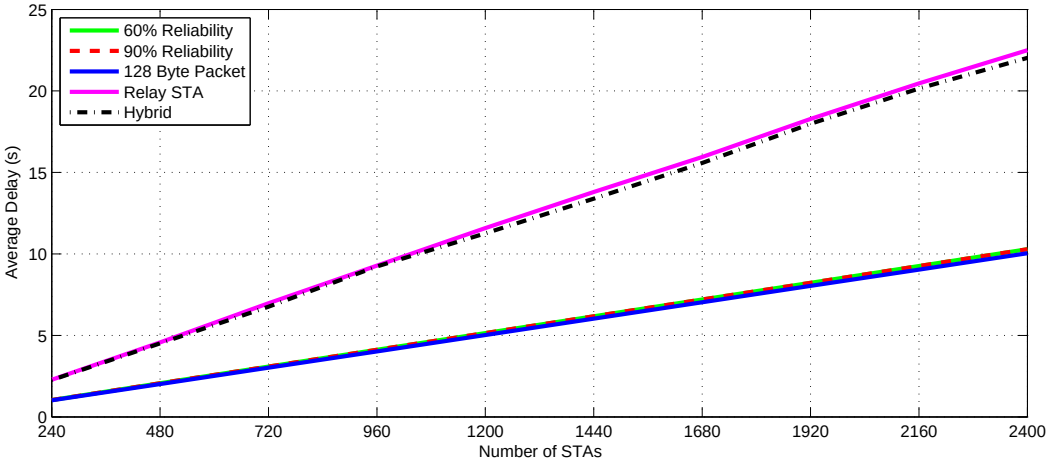


Figure 5.17: Delay

5.10 Summary

This chapter has presented a novel approach to setting the backoff timers of devices in a DCF based 802.11ah network to eliminate collisions in a sectorised network. It was shown that setting the backoff counters based on the AIDs of the group members eliminates collisions in the network. In order to eliminate collisions caused by hidden nodes, the STAs were grouped based on their received power graphs. The results showed that the use of clustering reduces the number of STAs that should be polled per grouping iteration without significantly increasing the number of groups created. The grouping results further show that formation of uniform sized groups of size greater than 2 is generally unattainable for low power devices with the set constraints. A group merging algorithm to merge smaller groups in order to create uniform groups is proposed. Simulation results show that as long as the STAs are aware of the hidden nodes in their group and set their backoff timers according to the AIDs of the STAs they detect in the group, DCF with AID-based backoff timers results in a 40% increment in the normalised throughput and a 58% reduction HOL delay compared to DCF with random backoff timers. Further, the use of STAs as relays for outage STAs in the absence of relays in the network has been proposed for a network using AID-based backoff. In order for the STAs to act as relays, they need to have network virtualisation functionality that enables them to act as both an STA and as relay. Results show that for dense networks, an outage drop from 9% to 0% can be achieved using this technique at the expense of network throughput and delay for in-band communications between the relay STA and the outage STA. The net reduction in throughput is reduced by nearly 2% when a hybrid technique which combines use of relay STAs and reduced reliability of 60% is used. This work assumes that every STA in coverage has the ability to act as a relay and has enough power to sustain the relaying function.

Chapter 6

Conclusions and Future Work

IN this chapter the conclusions drawn from this thesis and the future work are presented.

6.1 Conclusions

This thesis has focused on the optimisation of resource allocation in high user density wireless networks. For 4G networks, we considered an indoor small cell network while for 802.11ah we considered a BSS with up to 2400 STAs.

With regard to 4G networks, this thesis mainly focussed on the power domain. First, the thesis introduced the concept of heterogeneous networks and explained the drivers for the development of heterogeneous networks. The challenges faced by heterogeneous networks were briefly presented and interference management was identified as the key interest in this thesis. The nature of iterference in heterogeneous networks was characterised i.e. they experience both co-tier and cross-tier interference. This thesis went on to describe different approaches to tackling interference management in heterogeneous networks. It showed that one of the key approaches to interference management is interference avoidance which can be achieved through resource allocation. Resource allocation as a means of interference avoidance is quite popular because it does not require any hardware changes. Further, it is shown that resource allocation involves radio resource management (RRM), power allocation or both. A review of

the literature on resource allocation techniques is then presented.

One interesting form of RRM is identified as self organising network (SON) RRM. This type of RRM requires that the base stations have self organisation capabilities which enable the network to make intelligent changes in order to maximise some network performance criteria in response to detected changes in the network [89]. In this thesis SON capability is enabled through the use of an interference map. Changes in the interference map are communicated to the base stations by their UEs which then inform the central controller of the changes so that it can subsequently make changes to the manner in which it allocates resources in order to achieve interference avoidance. Two SON RRM techniques were selected to be used to demonstrate the perfomance of the power allocation techniques presented in this thesis.

A review of pilot power allocation techniques was then presented. Most of the works focussed on load balancing and on reducing interference to macrocell UEs which were given priority over small cell UEs. This thesis then proposed a power allocation technique that maximises the throughput and improves minimum data rate performance of the selected SON RRM techniques by modifying the interference map (referred to as the MoC). The improvement in throughput was achieved through modifying the pilot power using the proposed pilot power allocation technique. The technique is composed of two algorithms whose aim is to modify the MoC in a manner that increases the minimum RBs allocated to UEs in the network while also increasing reuse of RBs by inner UEs. The increased minimum RBs ensure that the minimum data rate performance of the SON RRM technique is never less than that attained by the SON RRM techniques with equal pilot power allocation while the increased reuse by inner UEs is responsible for the increase in throughput. The performance of the proposed technique is shown for both the homogeneous and heterogeneous deployment and it is seen to improve the minimum data rate, throughput as well as the ECR of two SON RRM techniques. The number of iterations required for the algorithms involved to converge is big so it is proposed that instead of sending the MoC, an SIR difference matrix is sent to the central controller which will implement the algorithms and send the transmit power instructions to each SBS. That way the overhead as a result of tens or hundreds of iterations of power reductions steps is reduced significantly.

The thesis then reviewed the literature on power control in heterogeneous networks. Most of the reviewed works either focussed on either interference minimisation or

allocated power minimisation. However, the interest of this thesis is improving of throughput through selective application of power control.

In chapter 4 a power control technique that is designed to improve the throughput attained by MoC based RRM technique is presented. The MoC is used to guide the initial allocation of RBs to all UEs. After the initial allocation, it is used to determine UEs with potentially more RBs to be allocated to them. These are the UEs that are barred from sharing resources with UEs served by neighbouring BSs because their serving BS causes significant interference to the UEs. In other words, the serving SBSs of the interfered UEs cause acceptable interference to the UE that can potentially be allocated extra RBS. From the current MoC, the SBSs with UEs that can potentially access more RBs send the maximum amount by which they can reduce their power in order to meet the SIR threshold. The interfered UE also shares the amount by which each potential UEs serving SBS has to reduce its power in order to share resources with it without reducing its current SIR (ignoring fading). If the latter value is less than the former value, the serving SBS of the potential UE reduces transmit power on the affected RBs and allocates the potential UE. This increases reuse while endeavouring to maintain the current data rate of incumbent UEs on the affected RBs. For further increase in throughput, the technique implements a distributed power adaptation algorithm which reduces transmit power for very high SINR UEs in order to reduce the level of interference in the network. This results in higher SINRs for lower data rate UEs and improves the minimum data rate performance of the network. The technique minimises the required information exchange by restricting it to information in the MR reports. The cost is that some of the UEs will have lower data rates as a result of the increased interference and the power adaptation. However, the reductions are minimal and the minimum data rate of the network is either maintained or improved for 10 or less SBSs.

The final chapter of the thesis focusses on 802.11ah networks. The 802.11ah standard was drafted to support IoT applications and in particular to cater for the large coverage distance when compared to existing 802.11 standards. 802.11ah networks are still plagued by the challenges faced by existing 802.11ah networks including, but not limited to: the hidden node problem and collisions from the uncoordinated nature of access. Grouping has been suggested to mitigate the hidden node problem and this thesis proposes a grouping algorithm which uses both clustering and the Welsh-Powell

algorithm in order to create groups that are hidden-node free. In Chapter 5, an interference map is used to guide grouping of the STAs in an 802.11ah network so that group members can detect each other's transmissions. First, a proximity matrix is developed using the distances between the STAs and it is used to partition the STAs in smaller groups using agglomerative clustering techniques. In order to create groups without hidden nodes, the AP polls STAs in the group and each STA draws up an interference map which shows which STAs in the group it detects. Using the Welsh Powell algorithm, new groups are created and these new groups are free of hidden nodes. This eliminates collisions resulting from the hidden node problem in 802.11ah networks. With the hidden node problem solved through grouping, the setting of backoff timers based on the AIDs of the group members is proposed. This ensures that the backoff timers have unique sequential values which avoid collisions and time wastage due to idle slots. A mathematical model of the performance of the proposed technique developed and the model is seen to match well with simulations. The performance of the proposed grouping and AID-based backoff technique is compared to existing techniques including a centralised one through simulations. The results show that the technique results in higher throughput, lower delay and better energy efficiency when the network is saturated.

Additionally, Chapter 5 builds on the AID based backoff technique to propose the use of connected STAs as relays for outage relays. Using STAs as relays, in the absence of relays, is shown to outperform the benchmark techniques in terms of eliminating outage in the network. However, this comes at the expense of reduction in throughput and increase in delay. The increased delay would still occur if the AP had 100% coverage however, the reduction in throughput is due to the fact that packet transfer between the two STAs is done using the same channel as the AP. A combination of STA relaying with reduced reliability is shown to solve the outage problem while improving the throughput attained by relay STAs alone. This is because of reduction in the number of STAs in outage by the reduced reliability and reduction in time wasted during packet forwarding to the relay STA.

6.2 Future Work

Future work recommendations based on the knowledge acquired in this research are given below:

1. Extending the use of AID-based backoff in a network the OBSS problem. In case of an OBSS, use of AID-based backoff without consideration of the OBSS would result in assured collisions as the STAs in different BSSs would set exactly the same backoff timer values. Therefore, in case of OBSS, there is need to investigate how best to set the AID-based backoff timers in case of the OBSS problem in order to avoid collisions due to setting the same backoff timer values. One approach that needs investigation is extending the spatial orthogonality concept to cater for groups of STAs in addition to one STA with a TXOP [3, 130]. The gain in throughput from this approach need to be investigated.

2. The work presented on 802.11ah assumes that all STAs use the same MCS class however the standard supports many MCS classes. There's a need to investigate how this impacts the performance of the AID-based backoff because it affects the grouping as well as the group slot sizes. In view of this, there's a need to investigate the optimal group slot size given the varying data rates of the STAs.

3. The 802.11ah standard supports three modes of BSS i.e. sensor, non-sensor only and mixed mode. The work presented in this thesis presented a scenario with sensor only BSS. Therefore one area of future research is optimisation of the throughput of 802.11ah networks with mixed mode BSS which deals with the complexities of the IoT [26]. How best to cater for the heterogeneity in terms of power transmitted, QoS, mobility is an open research problem that needs to be solved for the practical use of these networks for fully fledged IoT.

4. Evaluating the performance of MoC-based RRM in CDMA networks. The performance of MoC-based RRM in OFDMA networks is quite impressive however, there is a need to investigate whether similar results can be achieved in CDMA networks. CDMA networks suffer from multiple access interference and self interference which limit the performance of CDMA especially in the asynchronous uplink. An interesting type of codes that are being investigated for this purpose are orthogonal polyphase codes which exhibit cross correlation properties that are favourable for asynchronous communications [146].

5. The work presented on MoC-based RRM that all the UEs have a full buffer and therefore need all the RBs they can get. One future research area requires investigation on use of MoC based allocation where the UEs have specific required data rates and how best to meet the all or the highest percentage of users' data rate demands. Two options of particular interest are: power control for the UEs that do not require high data rates and using different SIR thresholds for different UEs based on the required data rate. In addition the use of sleep-modes or cell-switch of in case under utilisation in order to reduce interference, improve energy efficiency and throughput will be considered as an expansion to the work in [147].

6. A new architectural paradigm in HetNets is the decoupling uplink (UL) and downlink (DL) such that UEs are not constrained to be associated to the same BS in both the uplink and downlink [148]. A key motivation behind this concept is the difference in transmit powers in the UL and DL in HetNets which implies a UE may receive the strongest signal from the MBS in the DL but have a better signal to an SBS in the UL [148]. One area of future research is determining whether decoupling is beneficial to small cell networks in the absence of the MBS. If so, it is imperative to determine how to determine the best BS for a UE to associate with in the UL. Further, in the presence of the MBS, there's a need to determine which MUEs to decouple from the MBS in the uplink. Finally, there's a need to determine how power control would improve the network throughput in case this decoupling is applied in a HetNet scenario.

Bibliography

[1] E. Dahlman, S. Parkvall, and J. Skold, *4G: LTE/LTE-Advanced for Mobile Broadband: LTE/LTE-Advanced for Mobile Broadband*. Academic Press, 2011.

[2] C. So-In, R. Jain, and A.-K. Tamimi, "Scheduling in ieee 802.16e mobile wimax networks: key issues and a survey," *Selected Areas in Communications, IEEE Journal on*, vol. 27, no. 2, pp. 156–171, 2009.

[3] J. Wang. (2013, 01) Sectorization follow up 2. [Online]. Available: https://mentor.ieee.org/802.11/dcn/13/11-13-0081-01-00ah-sectorization-follow-up-2.pptx

[4] L. Zheng, M. Ni, L. Cai, J. Pan, C. Ghosh, and K. Doppler, "Performance analysis of group-synchronized DCF for dense IEEE 802.11 networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 11, pp. 6180–6192, Nov 2014.

[5] M. G. Kibria, G. P. Villardi, K. Nguyen, K. Ishizu, and F. Kojima, "Heterogeneous networks in shared spectrum access communications," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 1, pp. 145–158, 2017.

[6] C. V. N. Index, "Global mobile data traffic forecast update, 2015–2020 white paper," *link: http://goo. gl/ylTuVx*, 2016.

[7] F. Ghavimi, Y.-W. Lu, and H.-H. Chen, "Uplink scheduling and power allocation for m2m communications in sc-fdma-based lte-a networks with qos guarantees," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 7, pp. 6160–6170, 2017.

[8] A. Ghosh, N. Mangalvedhe, R. Ratasuk, B. Mondal, M. Cudak, E. Visotsky,

T. Thomas, J. Andrews, P. Xia, H.-S. Jo, H. Dhillon, and T. Novlan, "Heterogeneous cellular networks: From theory to practice," *Communications Magazine, IEEE*, vol. 50, no. 6, pp. 54–64, 2012.

[9] J. Andrews, "Seven ways that hetnets are a cellular paradigm shift," *Communications Magazine, IEEE*, vol. 51, no. 3, pp. 136–144, 2013.

[10] R. P. Liu, G. J. Sutton, and I. B. Collings, "Wlan power save with offset listen interval for machine-to-machine communications," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2552–2562, 2014.

[11] R. Abbas, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "Random access for m2m communications with qos guarantees," *IEEE Transactions on Communications*, vol. 65, no. 7, pp. 2889–2903, 2017.

[12] S. Chen, R. Ma, H.-H. Chen, H. Zhang, W. Meng, and J. Liu, "Machine-to-Machine communications in ultra-dense networks–a survey," *IEEE Communications Surveys & Tutorials*, 2017.

[13] L. M. Correia and R. Prasad, "An overview of wireless broadband communications," *IEEE Communications Magazine*, vol. 35, no. 1, pp. 28–33, 1997.

[14] B. P. Crow, I. Widjaja, J. G. Kim, and P. T. Sakai, "Ieee 802.11 wireless local area networks," *IEEE Communications magazine*, vol. 35, no. 9, pp. 116–126, 1997.

[15] D. Roeland and S. Rommer, "Advanced wlan integration with the 3gpp evolved packet core," *IEEE Communications Magazine*, vol. 52, no. 12, pp. 22–27, 2014.

[16] H. Claussen, L. T. Ho, and L. G. Samuel, "An overview of the femtocell concept," *Bell Labs Technical Journal*, vol. 13, no. 1, pp. 221–245, 2008.

[17] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive mimo: Ten myths and one critical question," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 114–123, 2016.

[18] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of

base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, 2010.

[19] H. Claussen, H. R. Karimi, and B. Mulgrew, "Low complexity detection of high-order modulations in multiple antenna systems," *IEE Proceedings-Communications*, vol. 152, no. 6, pp. 789–796, 2005.

[20] V. Jones and H. Sampath, "Emerging technologies for wlan," *IEEE Communications Magazine*, vol. 53, no. 3, pp. 141–149, 2015.

[21] T. Adame, A. Bel, B. Bellalta, J. Barcelo, and M. Oliver, "IEEE 802.11 ah: the WiFi approach for M2M communications," *IEEE Wireless Communications*, vol. 21, no. 6, pp. 144–152, 2014.

[22] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3gpp heterogeneous networks," *Wireless Communications, IEEE*, vol. 18, no. 3, pp. 10–21, 2011.

[23] S. ping Yeh, S. Talwar, G. Wu, N. Himayat, and K. Johnsson, "Capacity and coverage enhancement in heterogeneous networks," *Wireless Communications, IEEE*, vol. 18, no. 3, pp. 32–38, 2011.

[24] Y. Yang and S. Roy, "Grouping-based MAC protocols for EV charging data transmission in smart metering network," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 7, pp. 1328–1343, 2014.

[25] A. Hazmi, J. Rinne, and M. Valkama, "Feasibility study of IEEE 802.11 ah radio technology for IoT and M2M use cases," in *Globecom Workshops (GC Wkshps), 2012 IEEE*. IEEE, 2012, pp. 1687–1692.

[26] E. Khorov, A. Lyakhov, A. Krotov, and A. Guschin, "A survey on IEEE 802.11 ah: An enabling networking technology for smart cities," *Computer Communications*, vol. 58, pp. 53–69, 2015.

[27] M. Park, "IEEE 802.11 ah: Energy efficient MAC protocols for long range wireless LAN," in *Proc. of IEEE International Conference on Communications (ICC)*, 2014, pp. 2388–2393.

[28] F. Khan, *LTE for 4G mobile broadband: air interface technologies and performance*. Cambridge University Press, 2009.

[29] A. Abdelnasser, E. Hossain, and D. I. Kim, "Clustering and resource allocation for dense femtocells in a two-tier cellular ofdma network," *Wireless Communications, IEEE Transactions on*, vol. 13, no. 3, pp. 1628–1641, March 2014.

[30] Q. Song, X. Wang, T. Qiu, and Z. Ning, "An interference coordination-based distributed resource allocation scheme in heterogeneous cellular networks," *IEEE Access*, vol. 5, pp. 2152–2162, 2017.

[31] W. Pramudito and E. Alsusa, "A hybrid resource management technique for energy and qos optimization in fractional frequency reuse based cellular networks," *Communications, IEEE Transactions on*, vol. 61, no. 12, pp. 4948–4960, December 2013.

[32] A. Ebrahim and E. Alsusa, "Interference and resource management through sleep mode selection in heterogeneous networks," *IEEE Transactions on Communications*, vol. 65, no. 1, pp. 257–269, 2017.

[33] A. Dudnikova, A. Mastrosimone, and D. Panno, "An adaptive pilot power control for green heterogenous networks," in *Proc. of European Conference on Networks and Communications (EuCNC)*. IEEE, 2014, pp. 1–5.

[34] I. Ashraf, H. Claussen, and L. T. Ho, "Distributed radio coverage optimization in enterprise femtocell networks," in *Proc. of International Conference on Communications (ICC)*. IEEE, 2010, pp. 1–6.

[35] H. Kpojime and G. A. Safdar, "Efficacy of coverage radius-based power control scheme for interference mitigation in femtocells," *Electronics Letters*, vol. 50, no. 8, pp. 639–641, 2014.

[36] H. Claussen, "Performance of macro-and co-channel femtocells in a hierarchical cell structure," in *Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007. IEEE 18th International Symposium on*. IEEE, 2007, pp. 1–5.

[37] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535–547, 2000.

[38] J. G. Andrews, A. Ghosh, and R. Muhamed, *Fundamentals of WiMAX: understanding broadband wireless networking*. Pearson Education, 2007.

[39] T. S. Rappaport *et al.*, *Wireless communications: principles and practice*. Prentice Hall PTR New Jersey, 1996, vol. 2.

[40] B. Sklar, *Digital communications*. Prentice Hall NJ, 2001, vol. 2.

[41] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[42] I. Glover and P. M. Grant, *Digital communications*. Pearson Education, 2010.

[43] T. Ali-Yahiya, *Understanding LTE and its Performance*. Springer, 2011.

[44] S.-T. Cheng and M. Wu, "Contention-polling duality coordination function for ieee 802.11 wlan family," *Communications, IEEE Transactions on*, vol. 57, no. 3, pp. 779–788, 2009.

[45] J. He and H. K. Pung, "Performance modelling and evaluation of ieee 802.11 distributed coordination function in multihop wireless networks," *Computer Communications*, vol. 29, no. 9, pp. 1300–1308, 2006.

[46] A. Kanjanavapastit and B. Landfeldt, "Enhancements of the modified pcf in ieee 802.11 wlans," *Journal of Communications and Networks*, vol. 7, no. 3, pp. 313–324, 2005.

[47] T. Kim, D. J. Love, M. Skoglund, and Z. Y. Jin, "An approach to sensor network throughput enhancement by PHY-aided MAC," *IEEE Transactions on Wireless Communications*, vol. 14, no. 2, pp. 670–684, Feb 2015.

[48] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3gpp heterogeneous networks," *Wireless Communications, IEEE*, vol. 18, no. 3, pp. 10–21, June 2011.

[49] S.-P. Yeh, S. Talwar, G. Wu, N. Himayat, and K. Johnsson, "Capacity and coverage enhancement in heterogeneous networks," *IEEE Wireless Communications*, vol. 18, no. 3, 2011.

[50] J. Andrews, "Seven ways that hetnets are a cellular paradigm shift," *Communications Magazine, IEEE*, vol. 51, no. 3, pp. 136–144, March 2013.

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

[52] W. C. Cheung, T. Q. Quek, and M. Kountouris, "Throughput optimization, spectrum allocation, and access control in two-tier femtocell networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 561–574, 2012.

[53] K. Zheng, Y. Wang, W. Wang, M. Dohler, and J. Wang, "Energy-efficient wireless in-home: the need for interference-controlled femtocells," *Wireless Communications, IEEE*, vol. 18, no. 6, pp. 36–44, December 2011.

[54] L. Li, C. Xu, and M. Tao, "Resource allocation in open access ofdma femtocell networks," *IEEE wireless communications letters*, vol. 1, no. 6, pp. 625–628, 2012.

[55] D. Knisely, T. Yoshizawa, and F. Favichia, "Standardization of femtocells in 3gpp," *Communications Magazine, IEEE*, vol. 47, no. 9, pp. 68–75, 2009.

[56] F. Mhiri, K. Sethom, and R. Bouallegue, "A survey on interference management techniques in femtocell self-organizing networks," *Journal of Network and Computer Applications*, vol. 36, no. 1, pp. 58–65, 2013.

[57] N. Saquib, E. Hossain, L. B. Le, and D. I. Kim, "Interference management in ofdma femtocell networks: Issues and approaches," *Wireless Communications, IEEE*, vol. 19, no. 3, pp. 86–95, 2012.

[58] V. Chandrasekhar, J. Andrews, and A. Gatherer, "Femtocell networks: a survey," *Communications Magazine, IEEE*, vol. 46, no. 9, pp. 59–67, 2008.

[59] R. Kwan and C. Leung, "A survey of scheduling and interference mitigation in lte," *Journal of Electrical and Computer Engineering*, vol. 2010, p. 1, 2010.

[60] A. K. Jain and R. C. Dubes, *Algorithms for clustering data*. Prentice-Hall, Inc., 1988.

[61] R. Xu and D. Wunsch, "Survey of clustering algorithms," *IEEE Transactions on neural networks*, vol. 16, no. 3, pp. 645–678, 2005.

[62] F. Murtagh, "A survey of recent advances in hierarchical clustering algorithms," *The Computer Journal*, vol. 26, no. 4, pp. 354–359, 1983.

[63] F. Murtagh and P. Legendre, "WardâĂŹs hierarchical agglomerative clustering method: which algorithms implement wardâĂŹs criterion?" *Journal of Classification*, vol. 31, no. 3, pp. 274–295, 2014.

[64] I. S. Association *et al.*, "802.11-2012-ieee standard for information technology–telecommunications and information exchange between systems local and metropolitan area networks–specific requirements part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications," *Retrived from http://standards. ieee. org/about/get/802/802.11. html*, 2012.

[65] H. Nabuuma, E. Alsusa, W. Pramudito, and M. W. Baidas, "A power allocation technique for fairness and enhanced energy efficiency in future networks," in *Proc. of International Wireless Communications and Mobile Computing Conference (IWCMC)*. IEEE, 2016, pp. 244–249.

[66] H. Claussen and F. Pivit, "Femtocell coverage optimization using switched multi-element antennas," in *Communications, 2009. ICC'09. IEEE International Conference on*. IEEE, 2009, pp. 1–6.

[67] D. Lopez-Perez, A. Valcarce, G. de la Roche, and J. Zhang, "Ofdma femtocells: A roadmap on interference avoidance," *Communications Magazine, IEEE*, vol. 47, no. 9, pp. 41–48, 2009.

[68] K. Zheng, Y. Wang, W. Wang, M. Dohler, and J. Wang, "Energy-efficient wireless in-home: the need for interference-controlled femtocells," *Wireless Communications, IEEE*, vol. 18, no. 6, pp. 36–44, 2011.

[69] A. Ebrahim, E. Alsusa, and M. W. Baidas, "An uncoordinated frequency allocation scheme for future femtocell networks," in *Wireless Communications and Mobile Computing Conference (IWCMC), 2016 International*. IEEE, 2016, pp. 239–243.

[70] E. E. Tsiropoulou, P. Vamvakas, and S. Papavassiliou, "Supermodular game-based distributed joint uplink power and rate allocation in two-tier femtocell

networks," *IEEE Transactions on Mobile Computing*, vol. 16, no. 9, pp. 2656–2667, 2017.

[71] T. LeAnh, N. H. Tran, W. Saad, L. Le, D. Niyato, T. Ho, and C. S. Hong, "Matching theory for distributed user association and resource allocation in cognitive femtocell network," *IEEE Transactions on Vehicular Technology*, 2017.

[72] Z. Liu, S. Li, H. Yang, K. Y. Chan, and X. Guan, "Approach of power allocation in two-tier femtocell networks based on robust non-cooperativegame," *IET Communications*, 2017.

[73] P. Kulkarni, W. H. Chin, and T. Farnham, "Radio resource management considerations for lte femto cells," *ACM SIGCOMM Computer Communication Review*, vol. 40, no. 1, pp. 26–30, 2010.

[74] V. Chandrasekhar and J. Andrews, "Spectrum allocation in tiered cellular networks," *Communications, IEEE Transactions on*, vol. 57, no. 10, pp. 3059–3068, 2009.

[75] P. Lin, J. Zhang, Y. Chen, and Q. Zhang, "Macro-femto heterogeneous network deployment and management: from business models to technical solutions," *Wireless Communications, IEEE*, vol. 18, no. 3, pp. 64–70, 2011.

[76] T. Zahir, K. Arshad, A. Nakata, and K. Moessner, "Interference management in femtocells," 2012.

[77] S. Guruacharya, D. Niyato, D. I. Kim, and E. Hossain, "Hierarchical competition for downlink power allocation in ofdma femtocell networks," *Wireless Communications, IEEE Transactions on*, vol. 12, no. 4, pp. 1543–1553, 2013.

[78] Y. Wu, D. Zhang, H. Jiang, and Y. Wu, "A novel spectrum arrangement scheme for femto cell deployment in lte macro cells," in *Personal, Indoor and Mobile Radio Communications, 2009 IEEE 20th International Symposium on*. IEEE, 2009, pp. 6–11.

[79] H. Li, X. Xu, D. Hu, X. Tao, P. Zhang, S. Ci, and H. Tang, "Clustering strategy based on graph method and power control for frequency resource management in femtocell and macrocell overlaid system," *Communications and Networks, Journal of*, vol. 13, no. 6, pp. 664–677, 2011.

[80] Y. Sun, R. P. Jover, and X. Wang, "Uplink interference mitigation for ofdma femtocell networks," *Wireless Communications, IEEE Transactions on*, vol. 11, no. 2, pp. 614–625, 2012.

[81] A. Damnjanovic, J. Montojo, J. Cho, H. Ji, J. Yang, and P. Zong, "Ue's role in lte advanced heterogeneous networks," *Communications Magazine, IEEE*, vol. 50, no. 2, pp. 164–176, 2012.

[82] S. Singh and J. G. Andrews, "Joint resource partitioning and offloading in heterogeneous cellular networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 888–901, 2014.

[83] Y. Wang, B. Soret, and K. I. Pedersen, "Sensitivity study of optimal eicic configurations in different heterogeneous network scenarios," in *Communications (ICC), 2012 IEEE International Conference on*. IEEE, 2012, pp. 6792–6796.

[84] D. Lopez-Perez, I. Guvenc, and X. Chu, "Mobility management challenges in 3gpp heterogeneous networks," *Communications Magazine, IEEE*, vol. 50, no. 12, pp. 70–78, 2012.

[85] P. Lee, T. Lee, J. Jeong, and J. Shin, "Interference management in lte femtocell systems using fractional frequency reuse," in *Advanced Communication Technology (ICACT), 2010 The 12th International Conference on*, vol. 2. IEEE, 2010, pp. 1047–1051.

[86] T.-H. Kim and T.-J. Lee, "Throughput enhancement of macro and femto networks by frequency reuse and pilot sensing," in *Performance, Computing and Communications Conference, 2008. IPCCC 2008. IEEE International*. IEEE, 2008, pp. 390–394.

[87] R.-T. Juang, P. Ting, H.-P. Lin, and D.-B. Lin, "Interference management of femtocell in macro-cellular networks," in *Wireless Telecommunications Symposium (WTS), 2010.* IEEE, 2010, pp. 1–4.

[88] L. Zhang, T. Jiang, and K. Luo, "Dynamic spectrum allocation for the downlink of ofdma-based hybrid-access cognitive femtocell networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1772–1781, 2016.

[89] O. Aliu, A. Imran, M. Imran, and B. Evans, "A survey of self organisation in

future cellular networks," *Communications Surveys Tutorials, IEEE*, vol. 15, no. 1, pp. 336–361, 2013.

[90] A. Spilling, A. Nix, M. Beach, and T. Harrold, "Self-organisation in future mobile communications," *Electronics Communication Engineering Journal*, vol. 12, no. 3, pp. 133–147, 2000.

[91] Y.-S. Liang, W.-H. Chung, G.-K. Ni, Y. Chen, H. Zhang, and S.-Y. Kuo, "Resource allocation with interference avoidance in ofdma femtocell networks," *Vehicular Technology, IEEE Transactions on*, vol. 61, no. 5, pp. 2243–2255, 2012.

[92] S. Uygungelen, G. Auer, and Z. Bharucha, "Graph-based dynamic frequency reuse in femtocell networks," in *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, 2011, pp. 1–6.

[93] S.-F. Chou, H.-L. Chao, and C.-L. Liu, "An efficient measurement report mechanism for long term evolution networks," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2011 IEEE 22nd International Symposium on*. IEEE, 2011, pp. 197–201.

[94] K. Zheng, Y. Wang, C. Lin, X. Shen, and J. Wang, "Graph-based interference coordination scheme in orthogonal frequency-division multiplexing access femtocell networks," *IET communications*, vol. 5, no. 17, pp. 2533–2541, 2011.

[95] D. López-Pérez, A. Ladányi, A. Juttner, and J. Zhang, "Ofdma femtocells: A self-organizing approach for frequency assignment," in *Personal, Indoor and Mobile Radio Communications, 2009 IEEE 20th International Symposium on*. IEEE, 2009, pp. 2202–2207.

[96] W. Pramudito and E. Alsusa, "Joint dynamic frequency allocation and routing strategy for optimizing the power consumption and data rate of ofdma based femtocell networks," in *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*. IEEE, 2013, pp. 655–660.

[97] X. Chu, Y. Wu, L. Benmesbah, and W.-K. Ling, "Resource allocation in hybrid macro/femto networks," in *Wireless Communications and Networking Conference Workshops (WCNCW), 2010 IEEE*, 2010, pp. 1–5.

[98] W. Pramudito and E. Alsusa, "A hybrid resource management technique for

energy and qos optimization in fractional frequency reuse based cellular networks," *Communications, IEEE Transactions on*, vol. 61, no. 12, pp. 4948–4960, December 2013.

[99] N. Kolehmainen, J. Puttonen, P. Kela, T. Ristaniemi, T. Henttonen, and M. Moisio, "Channel quality indication reporting schemes for utran long term evolution downlink," in *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE*. IEEE, 2008, pp. 2522–2526.

[100] X. Chu, Y. Wu, D. Lopez-Perez, and X. Tao, "On providing downlink services in collocated spectrum-sharing macro and femto networks," *Wireless Communications, IEEE Transactions on*, vol. 10, no. 12, pp. 4306–4315, 2011.

[101] A. Attar, V. Krishnamurthy, and O. N. Gharehshiran, "Interference management using cognitive base-stations for umts lte," *Communications Magazine, IEEE*, vol. 49, no. 8, pp. 152–159, 2011.

[102] E. J. Hong, S. Y. Yun, and D.-H. Cho, "Decentralized power control scheme in femtocell networks: A game theoretic approach," in *Personal, Indoor and Mobile Radio Communications, 2009 IEEE 20th International Symposium on*. IEEE, 2009, pp. 415–419.

[103] S. Barbarossa, S. Sardellitti, A. Carfagna, and P. Vecchiarelli, "Decentralized interference management in femtocells: A game-theoretic approach," in *Cognitive Radio Oriented Wireless Networks & Communications (CROWNCOM), 2010 Proceedings of the Fifth International Conference on*. IEEE, 2010, pp. 1–5.

[104] R. Xie, F. Yu, H. Ji, and Y. Li, "Energy-efficient resource allocation for heterogeneous cognitive radio networks with femtocells," 2012.

[105] X. Kang, Y.-C. Liang, and H. K. Garg, "Distributed power control for spectrum-sharing femtocell networks using stackelberg game," in *Communications (ICC), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1–5.

[106] S. Guruacharya, D. Niyato, E. Hossain, and D. I. Kim, "Hierarchical competition in femtocell-based cellular networks," in *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*. IEEE, 2010, pp. 1–5.

[107] X. Kang, R. Zhang, and M. Motani, "Price-based resource allocation for spectrum-sharing femtocell networks: A stackelberg game approach," *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 3, pp. 538–549, 2012.

[108] N. D. Duong, A. Madhukumar, and D. Niyato, "Stackelberg bayesian game for power allocation in two-tier networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2341–2354, 2016.

[109] Y.-Y. Li and E. S. Sousa, "Base station pilot management for user-deployed cellular networks," in *Proc. of International Conference on Communications*. IEEE, 2009, pp. 1–5.

[110] L. Mohjazi, M. Al-Qutayri, H. Barada, and K. F. Poon, "Clustering based self-optimization of pilot power in dense femtocell deployments using genetic algorithms," in *Electronics, Circuits, and Systems (ICECS), 2013 IEEE 20th International Conference on*. IEEE, 2013, pp. 686–690.

[111] L. Giupponi and C. Ibars, "Distributed interference control in ofdma-based femtocells," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2010 IEEE 21st International Symposium on*, Sept 2010, pp. 1201–1206.

[112] *LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer - Measurements*, 3rd Generation Partnership Project (3GPP), Apr. 2010, version 9.1.0.

[113] D. T. Ngo, S. Khakurel, and T. Le-Ngoc, "Joint subchannel assignment and power allocation for ofdma femtocell networks," *Wireless Communications, IEEE Transactions on*, vol. 13, no. 1, pp. 342–355, 2014.

[114] A. Abdelnasser and E. Hossain, "Subchannel and power allocation schemes for clustered femtocells in two-tier ofdma hetnets," in *Communications Workshops (ICC), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1129–1133.

[115] E. Dahlman, S. Parkvall, and J. Skold, *4G: LTE/LTE-advanced for mobile broadband*. Academic press, 2013.

[116] L. B. Le, D. Niyato, E. Hossain, D. I. Kim, and D. T. Hoang, "Qos-aware and energy-efficient resource management in ofdma femtocells," *IEEE Transactions on Wireless Communications*, vol. 12, no. 1, pp. 180–194, 2013.

[117] X. Chu, D. López-Pérez, Y. Yang, and F. Gunnarsson, *Heterogeneous Cellular Networks: Theory, Simulation and Deployment*. Cambridge University Press, 2013.

[118] S. S. Jeng, J. M. Chen, C. W. Tsung, and Y. F. Lu, "Coverage probability analysis of ieee 802.16 system with smart antenna system over stanford university interim fading channels," *IET Communications*, vol. 4, no. 1, pp. 91–101, January 2010.

[119] H. Nabuuma and E. Alsusa, "Throughput maximisation in small cell networks using power control," in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*, Submitted.

[120] D. Lopez-Perez, X. Chu, A. Vasilakos, and H. Claussen, "Power minimization based resource allocation for interference mitigation in ofdma femtocell networks," *Selected Areas in Communications, IEEE Journal on*, vol. 32, no. 2, pp. 333–344, February 2014.

[121] A. Hatoum, R. Langar, N. Aitsaadi, R. Boutaba, and G. Pujolle, "Qos-based power control and resource allocation in ofdma femtocell networks," in *Global Communications Conference (GLOBECOM), 2012 IEEE*. IEEE, 2012, pp. 5116–5122.

[122] G. Cao, D. Yang, X. Zhu, and X. Zhang, "A joint resource allocation and power control algorithm for heterogeneous network," in *Telecommunications (ICT), 2012 19th International Conference on*. IEEE, 2012, pp. 1–5.

[123] G. Cao, D. Yang, and X. Zhang, "A distributed algorithm combining power control and scheduling for femtocell networks," in *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*. IEEE, 2012, pp. 2282–2287.

[124] T. Akbudak and A. Czylwik, "Distributed power control and scheduling for decentralized ofdma networks," in *Smart Antennas (WSA), 2010 International ITG Workshop on*. IEEE, 2010, pp. 59–65.

[125] A. Ebrahim and E. Alsusa, "A multi-level interference mapping technique for resource management in cellular networks," in *Vehicular Technology Conference*

(VTC Spring), 2015 IEEE 81st. IEEE, 2015, pp. 1–5.

[126] H. Nabuuma and E. Alsusa, "Enhancing the throughput of 802.11ah sectorized networks using AID-based backoff counters," in *Proc. of IEEE 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, June 2017, pp. 1921–1926.

[127] W. Sun, M. Choi, and S. Choi, "IEEE 802.11 ah: A long range 802.11 WLAN at Sub 1 GHz," *Journal of ICT Standardization*, vol. 1, no. 1, pp. 83–108, 2013.

[128] S. Aust, R. V. Prasad, and I. G. M. M. Niemegeers, "Outdoor long-range WLANs: A lesson for IEEE 802.11ah," *IEEE Communications Surveys Tutorials*, vol. 17, no. 3, pp. 1761–1775, thirdquarter 2015.

[129] O. Raeesi, J. Pirskanen, A. Hazmi, J. Talvitie, and M. Valkama, "Performance enhancement and evaluation of IEEE 802.11 ah multi-access point network using restricted access window mechanism," in *Proc. of IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS)*, 2014, pp. 287–293.

[130] J. Wang. (2013, 01) Sectorization scheme, IEEE 802.11-11/0081/r0. [Online]. Available: https://mentor.ieee.org/802.11/documents

[131] D.-Y. Yang, T.-J. Lee, K. Jang, J.-B. Chang, and S. Choi, "Performance enhancement of multirate IEEE 802.11 WLANs with geographically scattered stations," *IEEE Transactions on Mobile Computing*, vol. 5, no. 7, pp. 906–919, 2006.

[132] Y. Gao, X. Sun, and L. Dai, "Throughput optimization of heterogeneous IEEE 802.11 DCF networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 1, pp. 398–411, 2013.

[133] K.-C. Ting, M.-Y. Jan, S.-H. Hsieh, H.-H. Lee, and F. Lai, "Design and analysis of grouping-based DCF (GB-DCF) scheme for the MAC layer enhancement of 802.11 and 802.11 n," in *Proc. of the 9th ACM international symposium on Modeling analysis and simulation of wireless and mobile systems*. ACM, 2006, pp. 255–264.

[134] S. Aust, R. V. Prasad, and I. G. Niemegeers, "Sector-based RTS/CTS access scheme for high density WLAN sensor networks," in *Proc. of the 39th IEEE*

Conference on Local Computer Networks Workshops (LCN Workshops), 2014, pp. 697–701.

[135] M. Ghasemiahmadi, Y. Li, and L. Cai, "RSS-based grouping strategy for avoiding hidden terminals with GS-DCF MAC protocol," in *Proc. of Wireless Communications and Networking Conference (WCNC)*. IEEE, 2017, pp. 1–6.

[136] V. Loginov, E. Khorov, and A. Lyakhov, "On throughput estimation with TXOP sharing in IEEE 802.11 ah networks," in *Proc. of International Black Sea Conference on Communications and Networking (BlackSeaCom)*. IEEE, 2016, pp. 1–5.

[137] S. Kumar, H. Lim, and H. Kim, "Hierarchical MAC protocol with multi-channel allocation for enhancing IEEE 802.11 ah relay networks," in *Proc. of International Wireless Communications and Mobile Computing Conference (IWCMC)*. IEEE, 2015, pp. 1458–1463.

[138] N. Ahmed and M. I. Hussain, "Relay-based IEEE 802.11 ah network: A smart city solution," in *Proc. of Cloudification of the Internet of Things (CIoT)*. IEEE, 2016, pp. 1–6.

[139] S. Kandula, K. C.-J. Lin, T. Badirkhanli, and D. Katabi, "Fatvap: Aggregating ap backhaul capacity to maximize throughput." in *Proc. of the 5th USENIX Symposium on Networked Systems Design and Implementation*, vol. 8, 2008, pp. 89–104.

[140] H. Wirtz, T. Heer, R. Backhaus, and K. Wehrle, "Establishing mobile ad-hoc networks in 802.11 infrastructure mode," in *Proc. of the 6th ACM workshop on Challenged networks*. ACM, 2011, pp. 49–52.

[141] M. Qutab-ud din, A. Hazmi, L. F. Del Carpio, A. Goekceoglu, B. Badihi, P. Amin, A. Larmo, and M. Valkama, "Duty cycle challenges of IEEE 802.11 ah networks in M2M and IoT applications," *European Wireless Journal*, 2016.

[142] T. C. Chang, C. H. Lin, K. C. J. Lin, and W. T. Chen, "Load-balanced sensor grouping for IEEE 802.11ah networks," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*, Dec 2015, pp. 1–6.

[143] ETSI, "Electromagnetic compatibility and Radio spectrum Matters (ERM)," ETSI, Tech. Rep., 11 2014.

[144] D. J. Welsh and M. B. Powell, "An upper bound for the chromatic number of a graph and its application to timetabling problems," *The Computer Journal*, vol. 10, no. 1, pp. 85–86, 1967.

[145] A. Hazmi, B. Badihi, A. Larmo, J. Torsner, M. Valkama *et al.*, "Performance analysis of IoT-enabling IEEE 802.11 ah technology and its RAW mechanism with non-cross slot boundary holding schemes," in *Proc. of IEEE 16th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2015, pp. 1–6.

[146] E. AlSusa and A. ElKalagy, "On the performance of orthogonal polyphase sequences in asynchronous cdma systems," in *Proc. of the 17th International Conference on Software, Telecommunications & Computer Networks (SoftCOM)*. IEEE, 2009, pp. 196–200.

[147] H. Nabuuma, E. Alsusa, and W. Pramudito, "A load-aware base station switch-off technique for enhanced energy efficiency and relatively identical outage probability," in *Vehicular Technology Conference (VTC Spring), 2015 IEEE 81st*. IEEE, 2015, pp. 1–5.

[148] F. Boccardi, J. Andrews, H. Elshaer, M. Dohler, S. Parkvall, P. Popovski, and S. Singh, "Why to decouple the uplink and downlink in cellular networks and how to do it," *IEEE Communications Magazine*, vol. 54, no. 3, pp. 110–117, 2016.

[149] A. Mahmud and K. A. Hamdi, "A unified framework for the analysis of fractional frequency reuse techniques," *IEEE Transactions on Communications*, vol. 62, no. 10, pp. 3692–3705, Oct 2014.

[150] D. Tsolkas, E. Liotou, N. Passas, and L. Merakos, "A graph-coloring secondary resource allocation for d2d communications in lte networks," in *Proc. of 17th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, Sept 2012, pp. 56–60.

[151] Q. Zhang, X. Zhu, L. Wu, and K. Sandrasegaran, "A coloring-based resource

allocation for ofdma femtocell networks," in *Proc. of Wireless Communications and Networking Conference (WCNC*. IEEE, 2013, pp. 673–678.

[152] D. Wood, "A technique for colouring a graph applicable to large scale timetabling problems," *The Computer Journal*, vol. 12, no. 4, pp. 317–319, 1969.

Appendix A

Average SNR

The average E_b/N_0, $\mathbb{E}(E_b/N_0)$, can be determined from average SNR, $\mathbb{E}(\Upsilon_0)$, where Υ_0 is the SNR and is given by

$$\Upsilon_0 = \frac{P_t \cdot h \cdot d^{-\alpha}}{N_0 \mathcal{B}} \tag{A.1}$$

where, P_t denotes the transmit power, h is the Rayleigh fading channel, d denotes the distance between the AP and the reference device and α is the pathloss exponent. The $\mathbb{E}(\Upsilon_0)$ is dependent on $\mathbb{E}(d^{-\alpha})$ which is derived as follows. From [149], the distribution of the reference user in a circular cell, with the AP at the centre of the cell, without interference from neighbouring cells is given by

$$f_D(d) = \begin{cases} \frac{2d}{D_0^2}, & 0 \leq d \leq D_0, \end{cases} \tag{A.2}$$

where, D_0 is the cell radius. Using (A.2), $\mathbb{E}(d)$ is calculated by

$$\mathbb{E}(d) = \int_0^{D_0} d \cdot \frac{2d}{D_0^2} \cdot \partial d = \frac{2}{3} \cdot D_0. \tag{A.3}$$

Therefore, $\mathbb{E}(d^{-\eta})$ is given by

$$\mathbb{E}(d^{-\eta}) = \frac{2}{3} \cdot D_0^{-\eta}, \tag{A.4}$$

from which, $\mathbb{E}(\Upsilon_0)$ is given by

$$\mathbb{E}(\Upsilon_0) = \frac{P_t \cdot \mathbb{E}(d^{-\eta})}{N_0 \mathcal{B}}. \tag{A.5}$$

Using (5.5), $\mathbb{E}(E_b/N_0)$ is calculated by

$$\mathbb{E}(E_b/N_0) = \mathbb{E}(\Upsilon_0) - 10 \cdot \log_{10}(\mathcal{R}/\mathcal{B}). \tag{A.6}$$

Appendix B

Welsh-Powell Algorithm

In a resource allocation problem where resources need to be partitioned to minimise interference among users, the grouping problem may be defined as follows. Given K users, with an interference matrix $\zeta = \{\zeta_{ij}\}$ where $\zeta_{ij} = 0$ when users i and j only interfere each other to an acceptable degree and $\zeta_{ij} = 1$ where they either cause each other unacceptable interference [150, 151]. The problem is to find the minimum number of groups that can be formed. The smaller the number of groups, the better because each group will get a decent number of resources to share [151].

By respresenting each user in ζ by a point A_i and joining it to A_j by an undirected edge e_{ij} if and only if ζ_{ij}=1, the above problem is clearly equivalent to finding the minimum colouring of vertices of graph G with vertex set $V(G) = \left\{A_j\right\}_{j=1}^{K}$ and edge set $E(G) = \left\{e_{ij}\right\}$. The minimum number of groups, is the chromatic number of G denoted by $k(G)$ [144]. The problem of solving $k(G)$ for an arbitrary graph is a well known unsolved problem [144, 152]. It can be solved through exhaustive search however this is not a viable option for most applications because of the complexity and time requirements [144, 152].

A number of heuristics have been devised to solve this problem. In this regard, the simplest approach is to arbitrarily order the users and then start to group them. The process starts with one group which and if the next user in the queue interferes with any of the members of the group, a new group is formed. The process continues until all the users belong to a group. This is the basic greedy algorithm. Although the basic greedy algorithm is simple, it is known to result in higher than necessary number of groups.

The Welsh-Powell algorithm is a greedy graph colouring algorithm that improves the performance of the basic greedy algorithm [144]. The Welsh-Powell algorithm works as follows.

The degree of a vertex A_i is the number of edges with A_i as an end point and it is denoted by $\tilde{d}_i$. The vetices are sorted in descending order such that

$$\tilde{d}_1 \geq \tilde{d}_2 \geq \cdots \geq \tilde{d}_K \tag{B.1}$$

Given the ordering, the first vertex in the list forms the first group. The other vertices are then inspected in order: any vertex that is not connected to the vertices in the formed group, is added to the group. The second vertex is started by the first vertex that is not yet grouped. The other vertices are then inspected in order once again and only those not connected to the group members, are added to the second group. This process continues until all vertices belong to a group.