Self-Organized Intelligent Distributed Antenna System in LTE

by

Seyed Amin Hejazi

M.A.Sc., Amirkabir University of Technology, IRAN, 2009

B.A.Sc., University of Tehran, IRAN, 2007

Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

Doctor of Philosophy

in the
School of Engineering Science
Faculty of Applied Sciences

SIMON FRASER UNIVERSITY
Spring 2014

APPROVAL

Name: Seyed Amin Hejazi

Degree: Doctor of Philosophy

Title of Thesis: Self-Organized Intelligent Distributed Antenna System in LTE

Examining Committee: Dr. Bonnie L. Gray, Chair
Associate Professor
Chair

Dr. Shawn Patrick Stapleton, Senior Supervisor
Professor

Dr. Jie Liang, Supervisor
Associate Professor

Dr. Paul Ho, Supervisor
Professor

Dr. Ivan V. Bajic, Internal Examiner,
Associate Professor

Dr. Hong-Chuan Yang, External Examiner,
Professor, Electrical and Computer Eng., Univ. of Victoria

Date Approved: April 7th, 2014

Partial Copyright Licence

Abstract

In order to reduce the operational expenditure, while optimizing network efficiency and service quality, self-organizing network is introduced in long term evolution. The SON includes several functions, e.g. self-establishment of new base stations, load balancing, inter-cell interference coordination. Load balancing and inter-cell interference coordination are two of the most important self-organizing functions.

In this thesis, load-balancing solution is investigated in order to optimize quality of service. To enable load balancing among distributed antenna modules, we dynamically allocate the remote antenna modules to the BTS sectors. Self-optimizing intelligent distributed antenna system is formulated as an optimization problem. Three evolutionary algorithms are proposed for optimization: genetic algorithm, estimation distribution algorithm, and particle swarm optimization. Computational results of different traffic scenarios after performing the algorithms, demonstrate that the the algorithms attain excellent key performance indicators.

The downlink performance of cellular networks is known to be strongly limited by inter-cell interference in multi-carrier based systems when full frequency reuse is utilized. In order to mitigate this interference, a number of techniques have recently been proposed, e.g., the soft frequency reuse scheme. In this thesis, DAS is utilized to implement SFR. The central concept of this architecture is to distribute the antennas in a hexagonal cell such that the central antenna transmits the signal using entire frequency band while the remaining antennas utilize only a subset of the frequency bands based on a frequency reuse factor. A throughput-balancing scheme for DAS-SFR that optimizes cellular performance according to the geographic traffic distribution is also investigated in order to provide a high QoS. To enable throughput balancing among antenna modules, we dynamically change the antenna module's carrier power to manage the inter-cell interference. A downlink power self-optimization algorithm is proposed for the DAS-SFR system. The transmit powers are optimized in order to maximize the spectral efficiency of a DAS-SFR and maximize the number of satisfied users under different users' distributions. The results show that proposed algorithm is able to guarantee a high QoS that concentrates on the number of satisfied users as well as the capacity of satisfied users as the two KPIs.

To My Beloved Parents and Brother

“I’d rather be hated for who I am, than loved for who I am not”

— KURT COBAIN

Acknowledgments

First and foremost, I would like to thank my advisor, Prof. Shawn Stapleton, for having given to me such a wonderful opportunity to pursue this PhD study. I am thankful for your patience and critical advices during the course of my research.

I am very grateful to Prof. Jie Liang and Paul Ho for generously sharing their time and knowledge. I would also like to thank Prof. Ivan Bajic, my internal PhD examiner, and Prof. Hong-Chuan Yang, my external PhD examiner, for their invaluable time to review my thesis and make suggestions and comments.

I am sincerely indebted to Prof. Mahmoud Shahabadi for his encouragements and advices during my B.A.Sc and M.A.Sc studies in University of Tehran before starting my PhD. Also, I want to thank all my friends, who made life in Vancouver enjoyable for me.

Lastly, I am deeply thankful to my parents who gave me love, strength and support to achieve success in all stages of my life. I would not have been where I am now without your unconditional support. I would also like to thank my brother, Dr. Seyed Alireza Hejazi, for always being there.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Thesis Motivation.

In the last 15 years, there has been substantial growth in cellular mobile communication systems. It is imperative to provide a high quality of service (QoS) at a minimum cost. With the substantial increase in cellular users, unbalanced throughput and load distributions are common in wireless networks, which decrease the number of satisfied users.

In the next generation wireless communication systems that use the 3GPP LTE (Long Term Evolution [1]) standard, there is tremendous pressure to support a high data rate transmission. These systems are based on orthogonal frequency division multiple access (OFDMA) to support the high data rate service and improve the QoS, even for cell edge users as the main targets in the downlink [2]. Users located at the cell edge largely suffer from inter-cell interference from eNodeB (LTE base station) of neighboring cells [3].

When the traffic loads among cells are not balanced, the satisfaction probability of heavily loaded cells may be lower, since their neighboring cells cause high inter-cell interference on cell edge users. In this case, load balancing can be conducted to alleviate and potentially avoid this problem.

Also, as traffic environments change, the network performance will not be optimum. Therefore, it is necessary to perform inter-cell optimization of the network dynamically according to the traffic environment, especially when cell traffic is not uniformly distributed. This is one of the important optimization issues in Self-Optimizing Network (SON).

The motivation behind this Ph.D. work is to understand how to employ distributed antenna system (DAS) in order to balance load in SON, and combine DAS with different frequency reuse techniques in order to mitigate inter-cell interference and contribute to the SON development for providing high QoS in a time-varying environment.

1.2 Outline and Main Contributions.

In chapter 2, we provide a brief review of important background materials related to the thesis. We first review different load balancing techniques and inter-cell interference mitigation techniques. We then introduce and review the prior works on DAS and frequency reuse techniques.

In chapter 3, we review LTE architecture and introduce its radio interface. In chapter 4, we first introduce small cell and virtual cell and then show the advantages of utilizing virtual cell in order to balance load and provide high QoS.

In chapter 5, we investigate two load-balancing schemes for mobile networks that optimizes cellular performance according to the traffic geographic distribution in order to provide a high QoS. An intelligent distributed antenna system (IDAS) fed by a eNB (eNodeB) has the ability to distribute the cellular capacity over a given geographic area depending on the time-varying traffic. To enable load balancing among distributed antenna modules we dynamically allocate the remote antenna modules to the eNBs using an intelligent algorithm. Self-organized intelligent distributed antenna system (SOIDAS) is formulated as an integer based linear constrained optimization problem, which tries to balance the load among the eNBs. Three evolutionary algorithms are proposed for optimization: genetic algorithm (GA), estimation distribution algorithm (EDA) and discrete particle swarm optimization (DPSO) .

In chapter 6, we study the combination of DAS and soft frequency reuse technique in a unique cell architecture, which is called DAS-SFR. We propose low-complexity bandwidth allocation scenarios which do not require a complex processing system for allocating resources to users. The results show that, by controlling the amount of resources allocated to users located in different areas, we can increase the frequency reuse and also improve the data rate for exterior users (users near the cell edge). If the throughput requirement for the interior users (users near the cell center) is small, more resources are allocated to the exterior user. We also propose a throughput-balancing scheme on DAS-SFR architecture that optimizes cellular performance according to the geographic traffic distribution. To enable throughput balancing among antenna modules, we dynamically change the antenna module's carrier power to manage the inter-cell interference. A downlink power self-optimization (PSO) algorithm is proposed for the DAS-SFR system. The transmit powers are optimized in order to maximize the spectral efficiency of a DAS-SFR and maximize the number of satisfied users under different users' distributions.

1.3 Notations and Acronyms.

In this section we define the notations and acronyms used throughout this proposal.

a	A boldface lowercase letter denotes a vector.
A	A boldface uppercase letter denotes a matrix.
$\bar{(.)}$	Conjugate operation.
$(.)^{\mathrm{T}}$	Transpose operation.
$(.)^{\mathrm{H}}$	Hermitian transpose operation.
$(.)^{-1}$	Inverse operation.
$[.]_{k,l}$	$(k,l)^{th}$ entry of a matrix.
$[.]_k$	k^{th} entry of a vector.
$\mathbf{I}_N$	identity matrix of size N.

Table 1.1: List of notations.

AMC adaptive modulation and coding
ASE average spectral efficiency
BCCH broadcast control channel
BCH broadcast channel
BTS base transceiver station
CA cooperative encoding
CCCH common control channel
CDF cumulative distribution function
CP cyclic prefix
CR convergence rate
CSIR channel state information at receiver
CQI channel quality information
DAS distributed antenna System
DAU digital access unit
DCCH dedicated control channel
DCI downlink control information
DL-SCH downlink shared channel
DTCH dedicated traffic channel
DL downlink
DPSO discrete particle swarm optimization
DRU digital remote antenna module
EA evolutionary algorithm
EDA estimation distribution algorithm
EPC evolved packet core
EPS evolved packet system
ES exhaustive search
E-UTRAN evolved universal terrestrial
eNB eNodeB
FDD frequency division duplexing
FFR full frequency reuse
GA genetic algorithm
GLB geographic load balancing
GSM global system for mobile communications
GTP GPRS tunneling protocol
HARQ hybrid automatic repeat request
HFR hard frequency reuse
HSS home subscriber server

IDAS intelligent distributed antenna system
KPI key performance indicator
LS least square
LTE long-term evolution
MBSFN multicast broadcast single frequency network
MCCH multicast control channel
MIB master information block
MIMO multiple-input/multiple-output
MISO multiple-input/single-output
MME mobile management entity
MMSE minimum mean square error
MCH multicast channel
MTCH multicast traffic channel
MUD multiuser detection
NAS non-access stratum
OFDMA Orthogonal Frequency-Division Multiple Access
PBCH physical broadcast channel
PCCH paging control channel
PCFICH physical control format indicator channel
PCH paging channel
PCRF proxy and charging rules function
PDCCH physical downlink control channel
PDN packet data network
PDSCH physical downlink shared channel
PHICH physical HARQ indicator channel
PMCH physical multicast channel
PRACH physical random-access channel
P-SCH primary-Synchronization Channel
PSO power self-optimizing
PUSCH physical uplink shared channel
QoS quality of service
RB resource block
RA resource allocation
RACH random access channel
ROHC robust header compression
SC-FDMA single carrier frequency-division multiple access
SFR soft frequency reuse

S-GW service gate-way
SINR signal to interference ratio
SISO single-input/single-output
SON self-optimization network
S-SCH secondary -Synchronization Channel
TDD time division duplexing
TF transport format
TTI transmition time interval
UCI uplink control information
UE user equipment
UL uplink
UL-SCH uplink shared channel

List of acronyms.

Chapter 2

Background

In order to reduce the operational expenditure, while optimizing network efficiency and service quality, SON is introduced in LTE [4]. The SON includes several functions, e.g. self-establishment of new eNodeBs (LTE base stations), load balancing, inter-cell interference coordination, and so on [5]-[8]. Load balancing and inter-cell interference coordination are two of the most important self-organizing functions [7]. Load balancing aims to efficiently use the limited spectrum to improve unequal loaded network reliability. Several load balance policies have been studied, e.g., antenna parameters adjustment, transmit power adjustment and handover parameters adjustment, are proposed.

On the other hand, LTE is designed to use the entire frequency band inside each existing cell, thus the inter-cell interference becomes an important concept in SON. Inter-cell interference coordination using SON requires suitable algorithms which coordinate and assign the resources among cells.

2.1 Load Balancing Techniques.

There has been a substantial growth in mobile broadband communication systems with the introduction of smartphones and tablets. With the substantial increase in cellular users, traffic hot spots and unbalanced call distributions are common in wireless networks. This degrades the QoS and increases call blocking and call drops. As traffic environments change, the network performance will be sub-optimum. It is therefore necessary to perform self-optimization of the network dynamically according to the traffic environment, especially when cell traffic loads are not uniformly distributed. This is one of the important optimization issues in SON for 3GPP LTE [16]. When the traffic loads among cells are not balanced, the blocking probability of heavily loaded cells may be higher, while their neighboring cells may have resources not fully utilized. In this case load balancing can be

conducted to alleviate and even avoid this problem. In SON, parameter tuning is done automatically based on measurements. The use of load balancing is meant to deliver extra gain in terms of network performance. For load-balancing, this is achieved by adjusting the network control parameters in such a way that high load eNBs can offload to low load eNBs. A SON enabled network, where the proposed SON algorithm monitors the network and reacts to these changes in load, can achieve better performance by distributing the load amongst the eNBs. All studies about load balancing can be classified into two categories: "block probability-triggered load balancing" ([17, 18, 19]), and "utility based load balancing" ([20, 21, 22]). "Block probability-triggered load balancing" schemes have been proposed for efficient use of limited resources to increase the capacity of hot spots in wireless networks. Decreasing block probability is the main goal of these load balancing schemes, regardless of whether proportional fairness is applied or not. "Utility based load balancing" schemes have been proposed to balance system throughput while serving users in a fair manner, these schemes result a utility maximization problem with network-wide proportional fairness as an objective in a network.

Traffic load balancing in mobile cellular networks has been well studied since the first generation of mobile communication systems. Many methods have been proposed to address this problem, such as cell splitting [23], channel borrowing [23], channel sharing [24], dynamic channel allocation [25, 26], etc.

Geographic load balancing (GLB) using SON is recognized as a new approach for traffic load balancing ([27] and [28]) which provides dynamic load redistribution in real time according to the current geographic traffic conditions. Studies on GLB such as use of tilted antennas [29], and dynamic cell-size control (cell breathing) [30] have shown that the system performance can be improved by balancing non-uniformly distributed traffic.

One of the contributions of this thesis is to introduce DAS as a GLB technique which has this ability to dynamically distribute resources (or capacity) over a given geographic area depending on time-varying traffic, which requires solving an optimization problem. In conventional base station without a complex scheduler, it is impossible to distribute dynamically resources over a given area.

2.2 Inter-Cell Interference Mitigation Techniques.

In order to reduce inter-cell interference, several techniques have been incorporated in cellular systems which fall either in a signal processing techniques category or interference coordination/avoidance category.

In signal processing category solutions, the receiver and/or transmitter are equipped and aided with extra or modified signal processing techniques.

- **Interference Randomization:** The effect of interference is reduced by averaging it spectrally,

which is done by spreading the signals over a distributed set of non-consecutive subcarriers in order to achieve frequency diversity [10], [11]. Although randomization scheme can be easily implemented using scrambling or interleaving [10], it does not reduce the level of interference in the cell.

- **Interference Cancellation:** The interference is suppressed from the received signal in a sequential manner [73]. The regenerated interfering signals are subsequently subtracted from the received signal. Requiring inter-base station synchronization and requiring accurate channel state knowledge add more complexity in the system.
- **Network-level multiple input- multiple output (MIMO):** The interference is alleviated by using cooperative encoding among neighboring base stations. The antenna on each base station is considered as an element of a spatially distributed MIMO array. Cooperative encoding requires accurate channel knowledge at all base stations as well as precise time and phase synchronization of the transmitted signals [72].
- **Maximum Likelihood Multiuser Detection:** The maximal likelihood (ML) multiuser detection (MUD) minimizes the bit error rate by reducing the effect of interference signal. It requires accurate channels' information and also a complex low-power mobile handset.
- **Beamforming:** Beamforming maximizes the signal energy sent to/from intended users and minimizes the interference sent toward/from interfering users. A beamforming requires the channel state knowledge as well as the complete interference statistics.

In the interference coordination category of solutions, certain restrictions in frequency, time and/or power domain are applied to the resource scheduling between cells in order to minimize the inter-cell interference. This process is also known as frequency reuse technique. Hard frequency reuse technique splits the system bandwidth into a number of distinct sub-bands according to a selected reuse factor and lets neighboring cells transmit on different sub-bands. In soft frequency reuse technique, the power on some of the sub-bands are reduced rather than not utilized. Depending on traffic load and interference adaptively level, soft frequency reuse technique can be divided into semi-static and dynamic reuse types. In static reuse type, fixed predetermined configuration can be reconfigured every couple of days, whereas in semi-static reuse type, the reuse factor can be altered in a basis of a fraction of a minute [74]. In dynamic reuse type, frequency reuse is instantaneously changing with the interference level and traffic load with the objective to maintain optimal operation. Dynamic frequency reuse makes the scheduling process too complex where it causes a huge computation burden on the network and requires more signaling. On the other hand, static and semi-static types are considered as the serious options in practical systems [10, 70, 71].

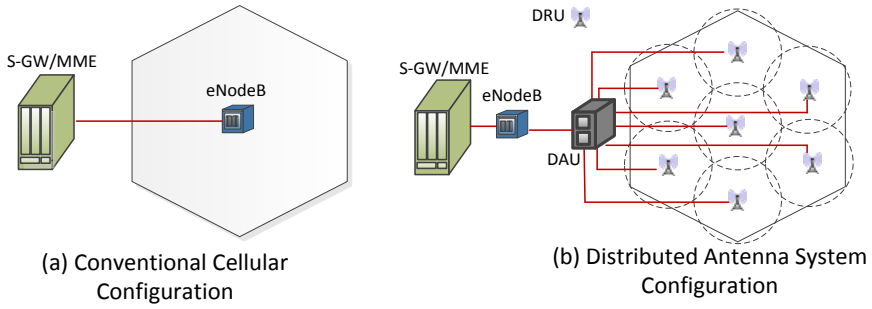


Figure 2.1: Conventional cellular configuration versus DAS.

2.3 Distributed Antenna System (DAS).

DAS have been widely implemented in state-of-the art cellular communication systems to cover dead spots in wireless communications systems [75], [76]. As opposed to a conventional cellular system, where the antenna is centrally located, a DAS network consists of antenna modules that are geographically distributed to reduce access distance. These distributed antennas are connected to an eNB by dedicated wires, fiber optics, or via a radio frequency link. DAS has potential advantages such as: throughput improvement, reducing call blocking rate, coverage improvement, increased cellphone battery life and a reduction in transmitter power [11,12].

A DAS breaks the traditional radio base station architecture into two pieces: a central processing facility and a set of distributed antenna modules connected to the central facility by a high-bandwidth network. The DAS network transports radio signals, in either analog or digital form, to/from the central facility where all the eNBs processing is performed. By replacing a single high-power antenna module with several low-power antennas modules distributed to give the same coverage as the single antenna, a DAS is able to provide more-reliable wireless service within a geographic area while reducing power consumption.

The general architecture of a DAS in a LTE multi-cell environment is shown in Fig. 2.1 (b), where antenna modules named digital remote units (DRUs) are connected to an eNB via an optical fiber and a digital access unit (DAU). The eNBs are linked to a public switched telephone network or a mobile switching center. For the simulcasting operation of DRUs allocated to a given eNB, the access network between each eNB and its DRUs should have a multi-drop bus topology. The DAUs assign the resources of the eNB to the independent DRUs. In contrast, the same cell is covered by only a high-power eNB in a conventional cellular system (Fig. 2.1 (a)).

Several advantages of DAS have been investigated such as, improving coverage indoors [77], [78], increasing both uplink [82] and downlink capacity [83], [84], reducing outages throughout the

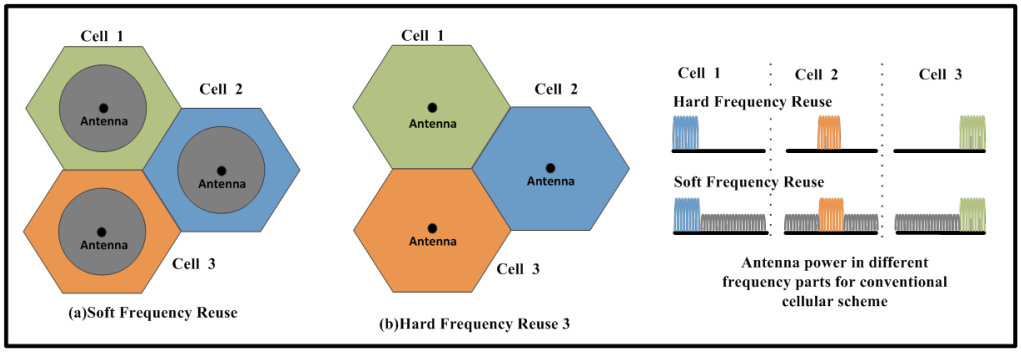


Figure 2.2: Conventional Frequency Reuse Techniques.

cell [79]-[81] and improving fairness among users [85], [86]. Better performances in outage and average capacity are achieved by careful antenna module placement in references [87]-[89]. References [76], [77]-[89] studied blanket transmission in such a way that all DRUs transmit at maximum power or DRU selection in such a way that only one DRU is chosen for transmission/reception. The work in [76], [77]-[89], however, is limited to single antenna per DRU. The work in [90]-[94] is focused on investigating the potential of multi-antenna DRUs in single-cell configuration, where the problem of trading off the number of DRUs with the number of antennas per DRU in cellular networks is considered. References [93], [94] showed that multi-antenna DRUs configuration can achieve large gains over both completely distributed and completely collocated configurations in single user multi-cell case. Multi-user case on DAS is studied on both the uplink and downlink transmission in references [95]-[97] and [98]-[100], respectively. The work in [95]-[97] [99], however does not consider out-of-cell interference.

2.4 Frequency Reuse Techniques.

Frequency reuse techniques have been adopted for inter-cell interference reduction in cellular systems. This benefits users near the cell edges owing to its simplicity and practicality. There are two major frequency reuse patterns for mitigating inter-cell interference: Hard Frequency Reuse and Soft Frequency Reuse

2.4.1 Hard Frequency Reuse (HFR).

HFR splits the system bandwidth into a number of distinct sub-bands according to a chosen reuse factor and lets neighboring cells transmit on different subbands. This inter-cell interference mitigation method is typically seen in GSM (Global System for Mobile Communications) networks, when

it comes to distribution of frequencies among the cells. When applied to LTE the resource blocks (a group of sub-carriers) are divided into 3, 4 or 7 disjoint sets. These sets of resource blocks (RBs) are assigned to the individual eNBs in such a way that neighboring cells do not use the same set of frequencies (Fig. 2.2 (b) with frequency reuse factor 3). This significantly reduces the interference at the cell edge of any pair of cells and can be considered the opposite extreme to full frequency reuse (FFR) or HFR1 (frequency reuse factor 1) in matters of frequency partitioning techniques. However, it may reduce the system capacity and spectrum efficiency [13].

2.4.2 Soft Frequency Reuse (SFR)

SFR has been proposed as an inter-cell interference mitigation technique in OFDM based wireless networks [14,15]. SFR shares the overall bandwidth by all eNBs (i.e. a reuse factor of one is applied), but for transmission on each group of RBs, the eNBs are restricted to a certain power bound. Fig. 2.2 (a) illustrates the power and frequency assignments in the different cells of a system with SFR. It can be noticed in the frequency spectrum of Fig. 2.2 (a) that there is a region of high-power transmissions and some regions of low-power transmissions. The RBs in the high-power region are preferably allocated to users located at the cell edge, while the low-power regions are allocated to user equipments (UEs) located at the cell-center.

HFR3 (frequency reuse factor 3) though simple in implementation suffers from a reduced spectral efficiency where it does not use the entire spectrum in each cell. On the other hand, the SFR has full spectral efficiency (using the entire spectrum in each cell) and is a strong mechanism for inter-cell interference mitigation. The SFR can be impractical in realistic settings involving a large number of eNBs, random traffic and realistic path-loss models. SFR can be also impractical in realistic terms; since, each base station antenna requires complex and expensive hardware tools to transmit on different restricted power bounds for different bands. However, an encouraging result is that by using these techniques, significant performance benefits can still be obtained over a conventional cellular architecture [2]. By utilizing DAS in order to implement SFR (DAS-SFR), the distance between transmit antenna modules and users is reduced; therefore, low realistic path-loss gain can be attained. Also, since each antenna module transmits on the same power for different frequency bands in DAS-SFR, complex hardware tools are not required to tranmit on different restricted power bands. Therefore, the combinations of DAS and SFR are considered in this thesis and then to enable throughput balancing, a downlink power self-optimization algorithm is proposed for the DAS and SFR combinations.

Chapter 3

LTE overview

LTE is the preferred development path of GSM/W-CDMA/HSPA networks currently deployed, and an option for evolution of CDMA networks. LTE enables networks to offer the higher data throughput in order to deliver advanced mobile broadband services to mobile terminals.

3.1 LTE Network Architecture.

Fig. 3.1 shows the LTE network architecture which is called the evolved packet system (EPS). EPS is a flat IP based architecture and is divided into the evolved universal terrestrial radio access network (E-UTRAN) and evolved packet core (EPC). The five elements of EPS architecture are as follows,

- **E-UTRAN:** The radio network, called the E-UTRAN, is comprised of the eNodeBs that are responsible in scheduling and allocation of the radio resources for the users in the LTE network. The eNodeBs are connected to the core network elements over the S1 interface and interconnected to each other over the X2 interface. The eNodeB terminates the control plane signaling messages as well as the user plane data with the EPC over the S1 interface.
- **EPC:** The core network, called EPC, is comprised of five elements: 1) mobility management entity (MME), 2) serving gateway (S-Gateway), 3) packet data network (PDN) gateway, 4) proxy and charging rules function (PCRF) and 5) home subscriber server (HSS).
- **MME:** The most important element in the EPC is the MME, which terminates the control plane signaling from the user. The MME performs the authentication, mobility management, security and retrieval of subscription information from the HSS.
- **Service Gateway:** Service gateway is responsible to forward the user plane packets from the mobile to the PDN Gateway. When the user moves across different eNodeBs, tunneling

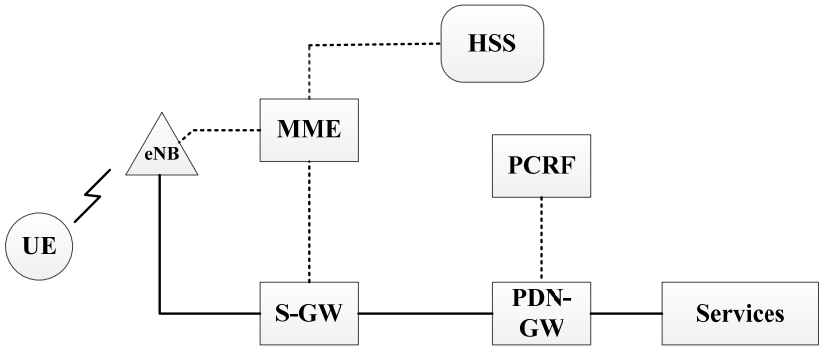


Figure 3.1: LTE Network Architecture.

the user plane IP packets using the GPRS tunneling protocol (GTP) is performed by service gateway.

- **PDN Gateway:** PDN Gateway is the last node in the LTE network. IP address allocation to the user is performed by PDN gateway. It is also responsible to route the user plane IP packets from the mobile nodes to other networks like Internet, IMS etc.
- **PCRF:** PCRF is responsible to execute various operator policies on the network like guaranteed QoS, maximum bit rate provisioned for a user.
- **HSS:** HSS comprises all the subscription information of the user along with the subscription for various services that are offered by the operator. It also comprises of the authentication center which stores all the keys required for ensuring the encryption and integrity of the data in the network.

The functional split between E-UTRAN and EPC is shown in Fig. 3.2. The UMTS RNC functionalities were split between base station and S-GW. It also has the functionalities of SGSN.

In this section, the functions of the different protocol layers and their location in the LTE architecture were described. Figures 3.2 and 3.3 show the control plane and the user plane protocols stack , respectively [68]. In the control-plane, the non-access stratum (NAS) protocol runs between the MME and the UE. It is also used for control-purposes such as network attach, authentication, setting up of bearers, and mobility management. MME and UE cipher and integrity protect all NAS messages. Handover decisions are made by the RRC layer in the eNB based on neighbor cell measurements sent by the UE, pages for the UEs over the air, broadcasts system information, controls UE measurement reporting such as the periodicity of channel quality information (CQI) reports and allocates cell-level temporary identifiers to active UEs. RRC layer also executes transfer of UE

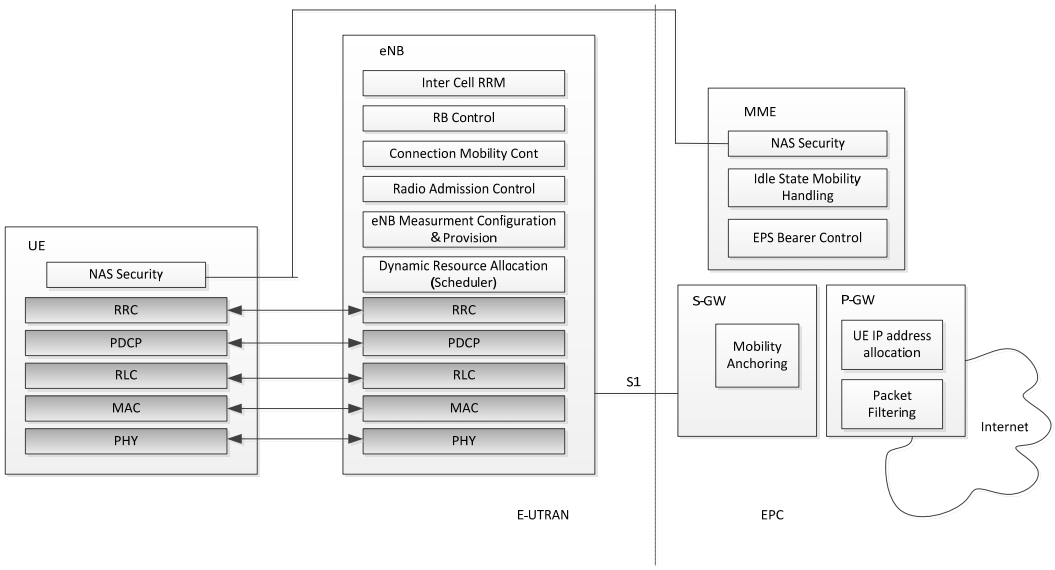


Figure 3.2: Functional split between E-UTRAN and EPC and control plane protocol stack.

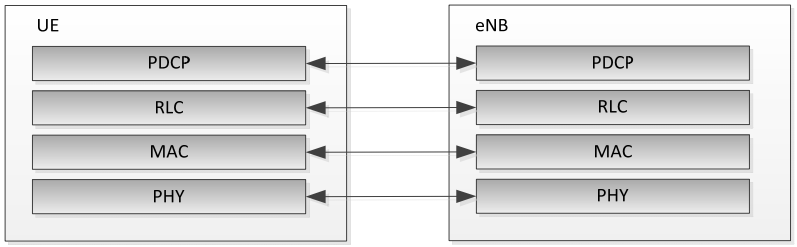


Figure 3.3: User plane protocol stack.

context from the source eNB to the target eNB during handover, and does integrity protection of RRC messages. The setting up and maintenance of radio bearers is performed by the RRC layer. In the user-plane, compressing/decompressing the headers of user plane IP packets is performed by the PDCP layer using robust header compression (ROHC) to enable efficient use of air interface bandwidth. This layer is also responsible to cipher both user plane and control plane data. Because the NAS messages are carried in RRC, they are effectively double ciphered and integrity protected, once at the MME and again at the eNB.

Formatting and transporting the traffic between the UE and the eNB are performed by the RLC layer. Three different reliability modes for data transport- acknowledged mode, unacknowledged mode, or transparent mode are provided by RLC layer. Because transport of real time services are delay sensitive and cannot wait for retransmissions, the unacknowledged mode is suitable for

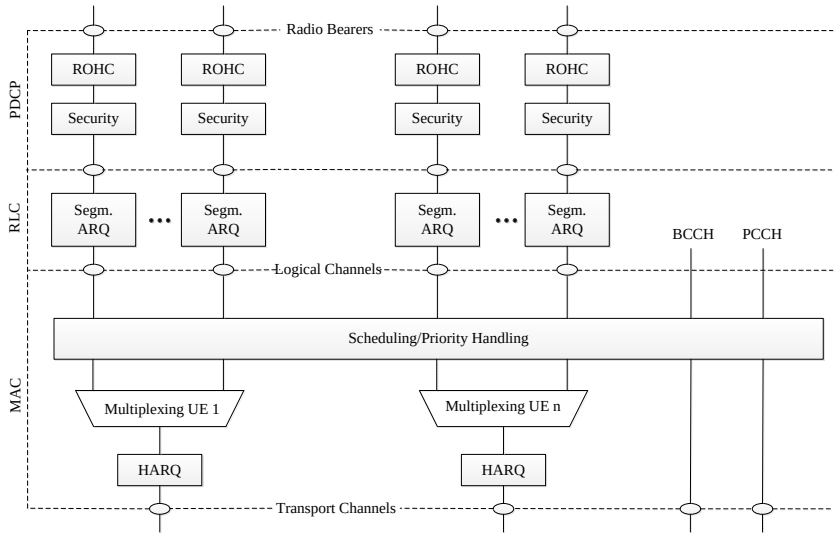


Figure 3.4: Layer 2 structure for DL.

such services. The acknowledged mode, on the other hand, is appropriate for non-RT services such as file downloads. When the PDU sizes are known, a priori such as for broadcasting system information, the transport mode is used.

Furthermore, the hybrid automatic repeat request (HARQ) at the MAC layer and outer ARQ at the RLC layer are two levels of re-transmissions for providing reliability. Handling residual errors are performed by the outer ARQ that are not corrected by HARQ. Asynchronous re-transmissions in the DL causes to the N-process stop-and-wait HARQ and synchronous re-transmissions in the UL. The re-transmissions of HARQ blocks occur at pre-defined periodic intervals in Synchronous HARQ mode. Therefore, no explicit signaling is required to indicate to the receiver the retransmission schedule. Asynchronous HARQ offers the flexibility of scheduling re-transmissions based on air interface conditions. The structure of layer 2 for DL and UL are shown in figures 3.4 and 3.5, respectively. The PDCP, RLC and MAC layers together constitute layer 2.

Significant efforts have been made to simplify the number of logical and transport channels. Figures 3.6 and 3.7 show the different logical and transport channels in LTE, respectively. Characteristics (e.g., adaptive modulation and coding) distinguish the transport channel with which the data are transmitted over radio interface. The mapping between the logical channels and transport channels are performed by the MAC layer. Scheduling the different UEs and their services in both UL and DL is also performed by MAC layer depending on their relative priorities. The logical channels are characterized by the information carried by them.

Fig. 3.8 shows, the mapping of the logical channels to the transport channels [68].

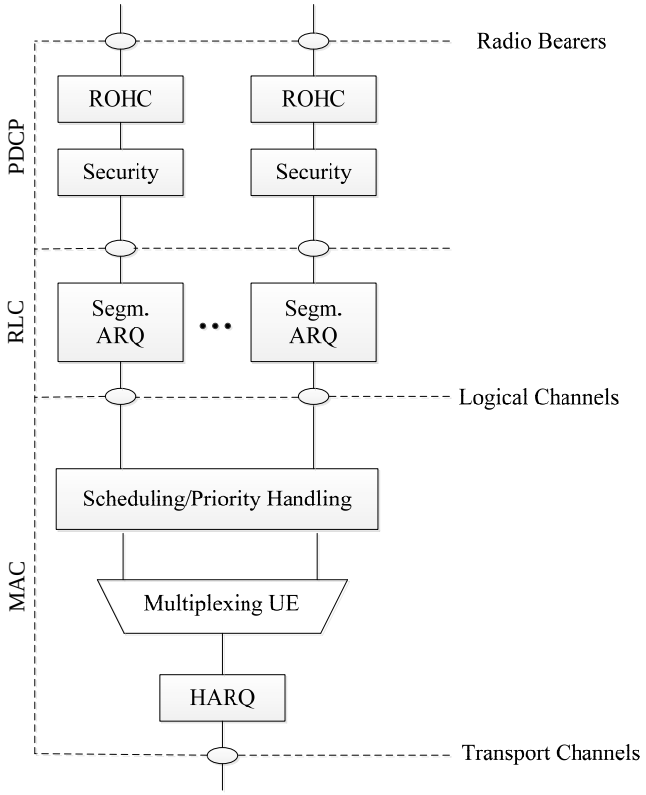


Figure 3.5: Layer 2 structure for UL.

Protecting data against channel errors using adaptive modulation and coding (AMC) schemes are performed by the physical layer at the eNB based on channel conditions. Physical layer also maintains frequency and time synchronization and performs RF processing including modulation and demodulation and processes measurement reports from the UE such as CQI and provides indications to the upper layers.

One time-frequency block corresponding to 12 sub-carriers is the minimum unit of scheduling. MIMO (multiple input multiple output) is supported at the UE with the 2 receive and 1 transmit antenna configuration. MIMO is also supported at the eNB with two transmit antennas being the baseline configuration. Transmission schemes for the DL and UL are orthogonal frequency division multiple access (OFDMA) and single carrier frequency division multiple access (SC-FDMA) with a sub-carrier spacing of 15 kHz, respectively. Each radio frame is 10 ms long containing 10 sub-frames with each sub-frame capable of carrying 14 OFDM symbols.

Services in the form of logical channels to the RLC are provided by the MAC layer. A logical channel is defined by the type of information it carries. It is generally classified as a control channel

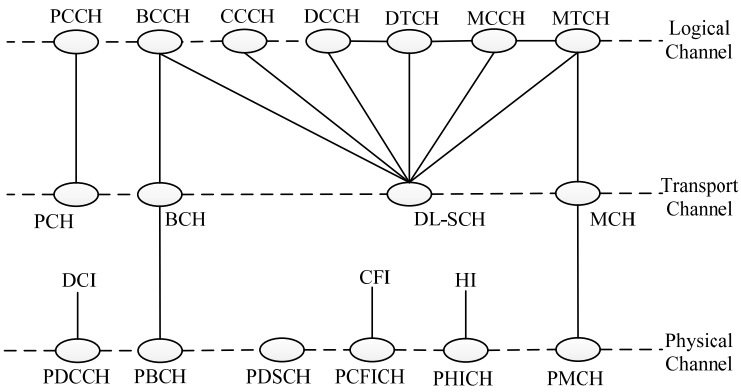


Figure 3.6: Downlink logical, transport and physical channels mapping.

or as a traffic channel. Control channel is utilized for transmission of control and configuration information necessary for operating an LTE system and traffic channel is utilized for the user data. The set of logical-channel types specified for LTE includes [54, section 8.2.2.1]:

- The **Broadcast Control Channel (BCCH)**, used for transmission of system information from the network to all terminals in a cell. A terminal needs to acquire the system information before accessing the system to find out how the system is configured and, in general, how to behave properly within a cell.
- The **Paging Control Channel (PCCH)**, used for paging of terminals whose location on a cell level is not known to the network. The paging message needs to be transmitted in multiple cells.
- The **Common Control Channel (CCCH)**, used for transmission of control information in conjunction with random access.
- The **Dedicated Control Channel (DCCH)**, used for transmission of control information to/from a terminal. This channel is used for individual configuration of terminals such as different handover messages.
- The **Multicast Control Channel (MCCH)**, used for transmission of control information required for reception of the MTCH.
- The **Dedicated Traffic Channel (DTCH)**,used for transmission of user data to/from a terminal. This is the logical channel type used for transmission of all uplink and non-multicast-broadcast single-frequency network (MBSFN) downlink user data.
- The **Multicast Traffic Channel (MTCH)**,used for downlink transmission of MBMS services.

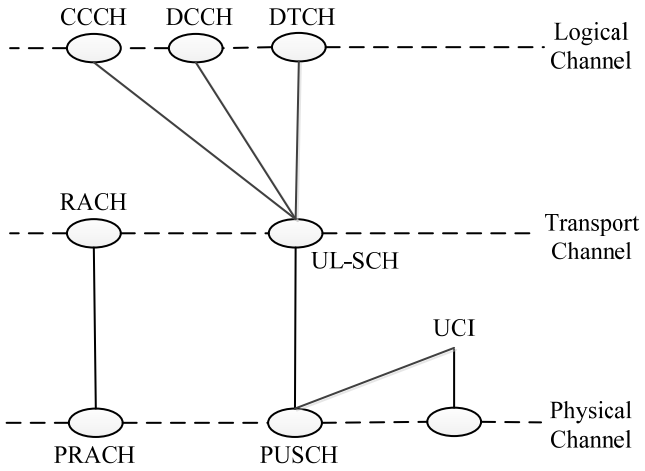


Figure 3.7: Uplink logical, transport and physical channels mapping.

Services in the form of transport channels are used by the MAC layer from the physical layer. How and with what characteristics the information is transmitted over the radio interface define a transport channel. Data on a transport channel is organized into transport blocks. At most one and two transport block is transmitted over the radio interface to/from a terminal in the absence and existence of spatial multiplexing in each transmission time interval (TTI), respectively.

A transport format which is associated with each transport block, specifies how the transport block is to be transmitted over the radio interface. The transport format includes information about the transport-block size, the modulation-and-coding scheme, and the antenna mapping. By varying the transport format, different data rates are realized by the MAC layer. Rate control is therefore also known as transport-format selection.

The following transport-channel types are defined for LTE [54, section 8.2]:

- The **Broadcast Channel (BCH)**, has a fixed transport format, provided by the specifications. It is used for transmission of parts of the BCCH system information, more specically the so-called master information block (MIB).
- The **Paging Channel (PCH)**, is used for transmission of paging information from the PCCH logical channel. The PCH supports discontinuous reception to allow the terminal to save battery power by waking up to receive the PCH only at predefined time instants.
- The **Downlink Shared Channel (DL-SCH)**, is the main transport channel used for transmission of downlink data in LTE. It supports key LTE features such as dynamic rate adaptation and channel dependent scheduling in the time and frequency domains, hybrid ARQ with soft

combining, and spatial multiplexing. It also supports DRX to reduce terminal power consumption while still providing an always-on experience. The DL-SCH is also used for transmission of the parts of the BCCH system information not mapped to the BCH. There can be multiple DL-SCHs in a cell, one per terminal scheduled in this TTI, and, in some subframe, one DL-SCH carrying system information.

- The **Multicast Channel (MCH)**, is used to support multimedia broadcast multicast service. It is characterized by semi-static transport format and semi-static scheduling. In the case of multi-cell transmission using MBSFN, the scheduling and transport format configuration is coordinated among the transmission points involved in the MBSFN transmission.

- The **Uplink Shared Channel (UL-SCH)**, is the uplink counterpart to the DL-SCH - that is, the uplink transport channel used for transmission of uplink data.

In addition, the random access channel (RACH) is also defined as a transport channel, although it does not carry transport blocks.

The physical layer is responsible for coding, physical-layer hybrid-ARQ processing, modulation, multi-antenna processing, and mapping of the signal to the appropriate physical time-frequency resources. It also handles mapping of transport channels to physical channels, as shown in figures 3.6 and 3.7 [54, section 8.2.3] .

Services to the MAC layer in the form of transport channels are provided by the physical layer. The DL-SCH and UL-SCH transport-channel types are used by data transmission in downlink and uplink, respectively. In the case of carrier aggregation, there is one DL-SCH (or UL-SCH) per component carrier. A physical channel corresponds to the set of time-frequency resources used for transmission of a particular transport channel and each transport channel is mapped to a corresponding physical channel, as shown in figures 3.6 and 3.7. In addition to the physical channels with a corresponding transport channel, there are also physical channels without a corresponding transport channel. These channels, known as L1/L2 control channels, are used for downlink control information (DCI), providing the terminal with the necessary information for proper reception and decoding of the downlink data transmission, and uplink control information (UCI) used for providing the scheduler and the hybrid-ARQ protocol with information about the situation at the terminal.

The physical-channel types defined in LTE include the following [54, section 8.2.3]:

- The **Physical Downlink Shared Channel (PDSCH)**, is the main physical channel used for unicast data transmission, but also for transmission of paging information.

- The **Physical Broadcast Channel (PBCH)**, carries part of the system information, required by the terminal in order to access the network.

- The **Physical Multicast Channel (PMCH)**, is used for MBSFN operation.

- The **Physical Downlink Control Channel (PDCCH)**, is used for downlink control information, mainly scheduling decisions, required for reception of PDSCH, and for scheduling grants enabling transmission on the PUSCH.
- The **Physical Hybrid-ARQ Indicator Channel (PHICH)**, carries the hybrid-ARQ acknowledgement to indicate to the terminal whether a transport block should be retransmitted or not.
- The **Physical Control Format Indicator Channel (PCFICH)**, is a channel providing the terminals with information necessary to decode the set of PDCCHs. There is only one PCFICH per component carrier.
- The **Physical Uplink Shared Channel (PUSCH)**, is the uplink counterpart to the PDSCH. There is at most one PUSCH per uplink component carrier per terminal.
- The **Physical Uplink Control Channel (PUCCH)**, is used by the terminal to send hybrid-ARQ acknowledgements, indicating to the eNodeB whether the downlink transport block(s) was successfully received or not, to send channel-state reports aiding downlink channel-dependent scheduling, and for requesting resources to transmit uplink data upon. There is at most one PUCCH per terminal.
- The **Physical Random-Access Channel (PRACH)**, is used for random access.

Note that some of the physical channels, more specically the channels used for downlink control information (PCFICH, PDCCH, and PHICH) and uplink control information (PUCCH), do not have a corresponding transport channel.

The remaining downlink transport channels are based on the same general physical-layer processing as the DL-SCH, although with some restrictions in the set of features used. This is especially true for PCH and MCH transport channels. For the broadcast of system information on the BCH, a terminal must be able to receive this information channel as one of the first steps prior to accessing the system. Consequently, the transmission format must be known to the terminals a priori, and there is no dynamic control of any of the transmission parameters from the MAC layer in this case. The BCH is also mapped to the physical resource (the OFDM timefrequency grid) in a different way.

For transmission of paging messages on the PCH, dynamic adaptation of the transmission parameters can, to some extent, be used. In general, the processing in this case is similar to the generic DL-SCH processing. The MAC can control modulation, the amount of resources, and the antenna mapping. However, as an uplink has not yet been established when a terminal is paged, hybrid ARQ cannot be used as there is no possibility for the terminal to transmit a hybrid-ARQ acknowledgement.

	1.4 MHz	3.0 MHz	5 MHz	10 MHz	15 MHz	20 MHz
Sub-frame (TTI)[ms]	1					
Sub-carrier spacing [kHz]	15					
Sampling [MHz]	1.92	3.84	7.68	15.36	23.04	30.72
FFT	128	256	512	1024	1536	2048
Sub-carriers	72	180	300	600	900	1200
Symbols per frame	4 with short CP and 6 with long CP					
Cyclic perfix	5.21 micro seconds with short CP and 16.67 micro seconds with long CP					

Table 3.1: Key Parameters for different bandwidths.

The MCH is used for MBMS transmissions, typically with single-frequency network operation, by transmitting from multiple cells on the same resources with the same format at the same time. Hence, the scheduling of MCH transmissions must be coordinated between the cells involved and dynamic selection of transmission parameters by the MAC is not possible.

3.2 Radio Interface.

The multiple-access is based on the use of SC-FDMA with cyclic prefix (CP) in the UL and OFDMA in the DL [101]. QAM modulator is coupled with the addition of the cyclic prefix for SC-FDMA transmission. The inter symbol interference (ISI) is eliminated by utilizing cyclic prefix, which enables the low complexity equalizer receiver. The fundamental difference to WCDMA is now the use of different bandwidths, from 1.4 up to 20 MHz.

Parameters have been picked up in such a way that FFT lengths and sampling rates are easily obtained for all operation modes and at the same time ensures the easy implementation of dual mode devices with a common clock reference. The parameters for the different bandwidths are shown Table 3.1.

The LTE physical layer is designed in such a way that there are only shared channels to enable dynamic resource utilization for maximum efficiency of packet-based transmission. There are synchronization signals in order to facilitate cell search and reference signals in order to facilitate channel estimation and estimate channel quality.

LTE has two radio frame structures. Frame structure type 1 uses both frequency division duplexing (FDD) and time division duplexing (TDD), and frame structure type 2 uses TDD duplexing. Frame structure type 1 is optimized to co-exist with 3.84 $Mbps$ UMTS. Frame structure type 2 is optimized to co-exist with 1.28 $Mbps$ UMTS TDD, also known as time division-synchronous code division multiple access.

Fig. 3.8 shows frame structure type 1 where the DL radio frame has a duration of 10 ms and consists of 10 sub-frames with a duration of 1 ms. A sub-frame consists of two slots. The physical mapping of DL physical signals for frame structure type 1 is:

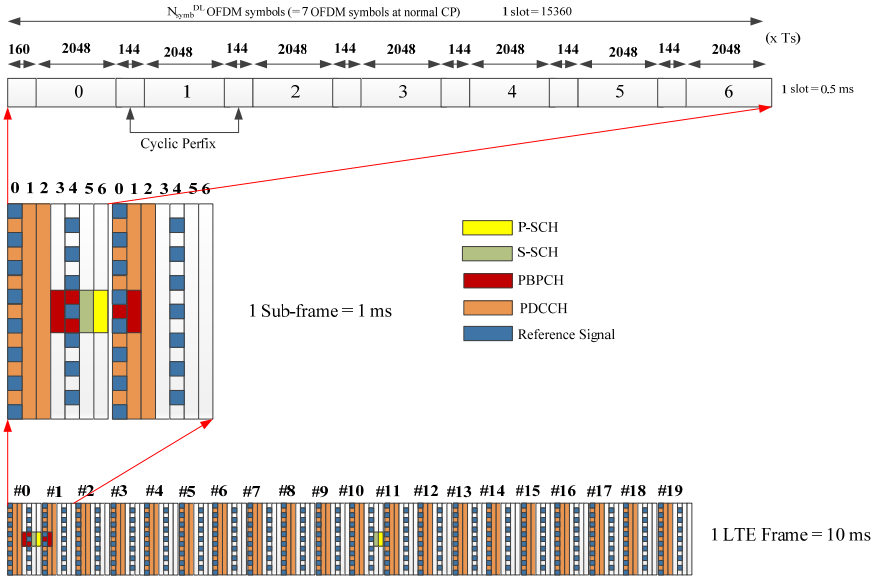


Figure 3.8: DL frame structure type 1.

- **Reference signal**, which is transmitted at OFDM symbol 0 and 4 of each slot. This depends on antenna port number.
- **Primary-Synchronization Channel (P-SCH)**, which is transmitted on symbol 6 of slots 0 and 10 of each radio frame.
- **Secondary-Synchronization channel (S-SCH)**, which is transmitted on symbol 5 of slots 0 and 10 of each radio frame.
- **PBCH physical channel**, which is transmitted on 72 sub-carriers centered around the DC sub-carrier. The smallest time-frequency unit for DL transmission is called a resource element, which is one symbol on one sub-carrier. A group of 12 contiguous sub-carriers in frequency and one slot in time form a resource block (RB) as shown in Figure 3.9. Data is allocated to each UE in units of RB.

For a frame structure type 1, using normal CP, a RB spans 12 consecutive sub-carriers and 7 consecutive OFDMA symbols over a slot duration. For extended CP there are 6 OFDMA symbols per slot. A CP is appended to each symbol as a guard interval. Thus, an RB has 84 resource elements (12 sub-carriers $\times$ 7 symbols) corresponding to one slot in the time domain and 180 kHz (12 sub-carriers $\times$ 15 kHz spacing) in the frequency domain. The size of an RB is the same for

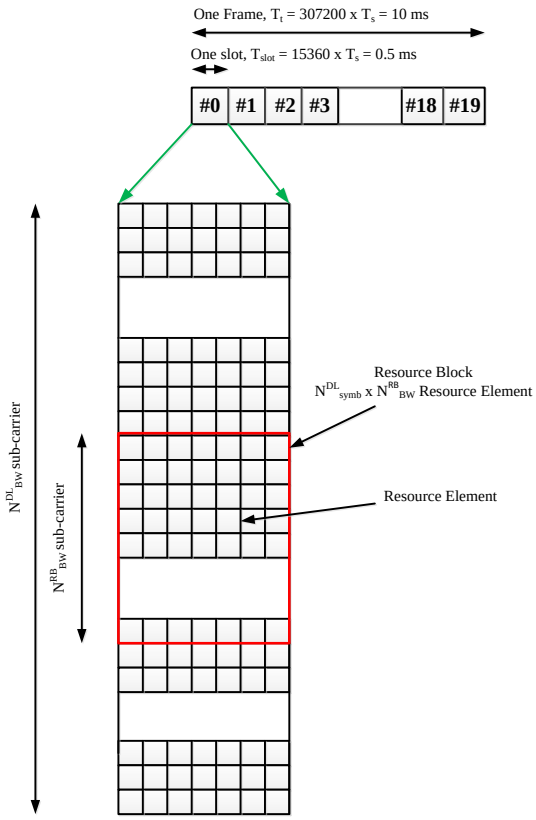


Figure 3.9: DL Resource Grid.

all bandwidths; therefore, the number of available physical RBs depends on the transmission bandwidth. In the frequency domain, the number of available RBs can range from 6, when transmission bandwidth is 1.4 MHz, to 100, when transmission bandwidth is 20 MHz. The UL frame, slot, and sub-frame of structure type 1 is the same as DL's one. Fig. 3.10 shows an UL structure type 1. The number of symbols in a slot depends on the CP length. For a normal CP, there are 7 SC-FDMA symbols per slot. For an extended CP there are 6 SCFDMA symbols per slot. UL demodulation reference signals, which are used for channel estimation for coherent demodulation, are transmitted in the fourth symbol (i.e., symbol number 3) of the slot.

Three potential frequency bands which are used in the 3GPP specified UMTS spectrum are: the 900 MHz band, [890;915] MHz for UL and [935;960] MHz for DL, the 2100 MHz frequency band and the 2600 MHz band, [2500;2570] MHz for UL and [2620;2690] MHz for the DL [69].

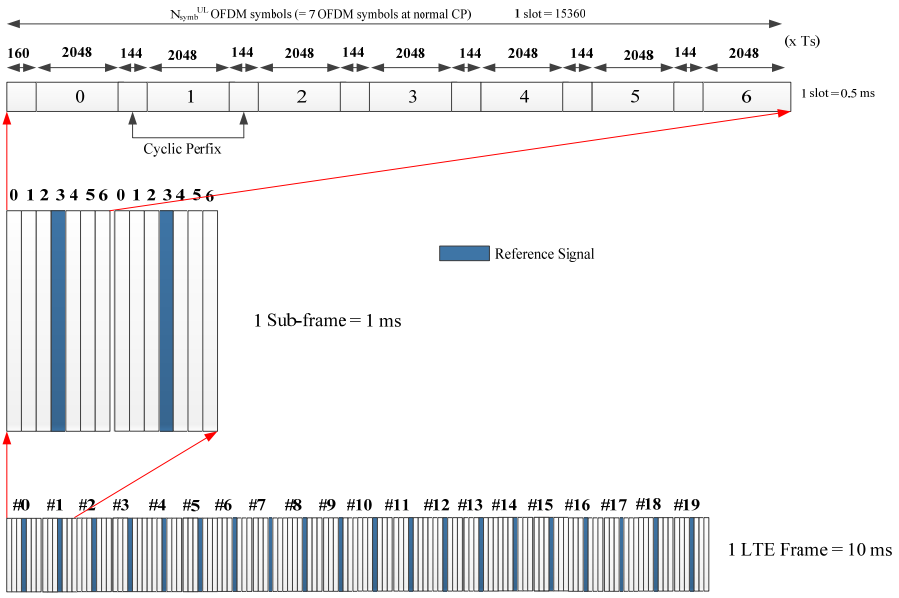


Figure 3.10: UL frame structure type 1.

3.3 Capacity and Coverage.

Data rate for a particular user is dependent on: the number of resource blocks allocated, rate of the channel coding, modulation applied, whether MIMO is used or not and the configuration, amount of overhead, including whether long or short cyclic prefix is used.

Table 3.2 shows the achievable DL peak bit rates. QPSK modulation carries 2 bits per symbol, 16QAM 4bits per symbol and 64QAM 6 bits. And 2$\times$2 MIMO doubles the peak bit rate. The bandwidth is included in the data rate calculation by taking the corresponding to the number of used sub-carriers, i.e., 72 , 180, 300, 600 and 1200 subcarriers per 1.4, 3.0, 5, 10, and 20 MHz bandwidth, respectively. The highest theoretical data rate is approximately 170 $Mbps$ where it is assumed 13 data symbols per 1 ms sub-frame.

Table 3.3 shows the achieved UL peak bit rates. Since single user MIMO is not specified in UL, the peak data rates are lower in UL than in DL. MIMO can be used in UL as well to increase cell data rates, not single-user peak data rates.

		Peak bit rate per sub-carrier/bandwidth combination [Mbps]				
Modulation Coding		72/1.4 MHz	180/3.0 MHz	300/5.0 MHz	600/10 MHz	1200/20 MHz
QPSK 1/2	Single Stream	0.9	2.2	3.6	7.2	14.4
16QAM 1/2	Single Stream	1.7	4.3	7.2	14.4	28.8
16QAM 3/4	Single Stream	2.6	6.5	10.8	21.6	43.2
64QAM 3/4	Single Stream	3.9	9.7	16.2	32.4	64.8
64QAM 4/4	Single Stream	5.2	13.0	21.6	43.2	86.4
64QAM 3/4	2x2 MIMO	7.8	19.4	32.4	64.8	129.6
64QAM 4/4	2x2 MIMO	10.4	25.9	43.2	86.4	172.8

Table 3.2: DL peak bit rates.

		Peak bit rate per sub-carrier/bandwidth combination [Mbps]				
Modulation Coding		72/1.4 MHz	180/3.0 MHz	300/5.0 MHz	600/10 MHz	1200/20 MHz
QPSK 1/2	Single Stream	0.9	2.2	3.6	7.2	14.4
16QAM 1/2	Single Stream	1.7	4.3	7.2	14.4	28.8
16QAM 3/4	Single Stream	2.6	6.5	10.8	21.6	43.2
16QAM 4/4	Single Stream	3.5	8.6	14.4	28.8	57.6
64QAM 3/4	Single Stream	3.9	9.0	16.2	32.4	64.8
64QAM 4/4	Single Stream	5.2	13.0	21.6	43.2	86.4

Table 3.3: UL peak bit rates.

Chapter 4

Virtual Cells Utilization for Self-Organized Network

With the increase of mobile broadband users, traffic hot spots and unbalanced traffic distributions are common in wireless networks.

The traffic load of wireless networks is often unevenly distributed among the eNBs, which results in unfair bandwidth allocation among users. We argue that the load imbalance and consequent unfair bandwidth allocation can be greatly reduced by intelligent association control. As users are, typically, not uniformly distributed, some eNBs tend to suffer from heavy load while adjacent eNBs may carry only light load or be idle. Such load imbalance among eNBs is undesirable as it hampers the network from providing fair services to its users.

The past two decades have witnessed a rapid development of cellular networks. As cellular systems are gaining popularity, the traffic demand has increased signicantly, while the available wireless bandwidth is still scarce. Wireless technology standards are evolving towards higher bandwidth requirements for both peak data rates and cell throughput growth. Higher data-rates require a strong signal strength to interference plus noise (SINR) ratio. To address this, a dedicated, wireless system is preferred for greater coverage and capacity. In order to balance an imbalanced network, an SON enabled network can offload the high load eNBs to low load eNBs.

4.1 Virtual Cells versus Small Cells.

Another impact of the arrival of broadband mobile communications is the increase in data traffic, which requires more capacity from the network. This can be achieved by using more spectral bandwidth, increasing the number of base station cells or access points, increasing the spectral efficiency and using load balancing techniques. In a wireless cellular network, call activity can be

more intensive in some areas than others. These high-traffic areas are called hotspot regions. Virtual Cell and distributed antenna system (DAS) were originally introduced to solve hotspot coverage problem, which are mainly affected by the traffic demands and spectral efficiency [62].

DAS is comprised of many remote antenna ports distributed over a large area and connected to a single base station by fiber, CAT 6, coax cable or microwave links. Without advanced signal processing techniques in the DAS, the same downlink signal is broadcast on all of its antennas, also known as simulcast. Studies show that simulcasting is an effective means to combat shadowing in noise-limited environments due to transmitter macro diversity [61]. DAS can help enhance the coverage and SINR when compared to Small Cells for the same transmit power [62].

Small Cell is a cost-effective alternative to extend coverage and capacity. The number of Small Cells is typically equivalent to the number of remote antennas in DAS. Small Cell systems allow greater spectral reuse, larger capacity, and use of low power hand-held user devices. However, there is also an increase in the number of cell boundaries that a mobile unit crosses. These boundary crossings stimulate hand off calls and location tracking operations, which are very expensive in terms of time delay and communication bandwidth, hence limiting the call handling capacity of a cellular system. Another challenge with Small Cell deployment is the severe inter-cell interference: Small Cell system performance is significantly degraded without any interference management [63]. Since each Small Cell provides a limited capacity, the areas with high user density need to be provisioned to provide sufficient users average busy hour throughput plus some headroom. Over provisioning of the Small Cells will lead to inefficiencies in the deployment of resources and, ultimately, additional costs.

One way of controlling the increase of signaling traffic, while preserving the frequency reuse advantage of smaller cells, is to adopt an intelligent DAS architecture. DAS has a number of advantages: centralization of base station resources, neutral host compatibility, modulation independece, and higher SINR over the coverage area [64]. An IDAS system, which has the ability to alter the simulcast ratio via load balancing, has a high spectral efficiency as well as a data throughput performance equivalent to that of a Small Cell at a hot spot. The terminology used to define an IDAS node is a Virtual Cell. A Virtual Cell is a remote node that has access to a base station with adequate and scalable resources, potentially located in a Base Station Hotel. A Virtual Cell will have the added advantages of scalability of eNB resources. The Base Station Hotel can be viewed as a Local Cloud. When the remote units have access to all of the systems resources, there is no need to add new base stations or bandwidth for higher capacity requirements in hotspot areas. If more capacity is required, additional resources or base stations can simply be added at the head-end central location. Moreover, in a distributed network architecture, where all resources are centralized, multi-band and multi-operator scenarios can easily be accommodated. The Base Station Hotel resources that are centrally available can be routed to the remote Virtual Cells via the distributed network.

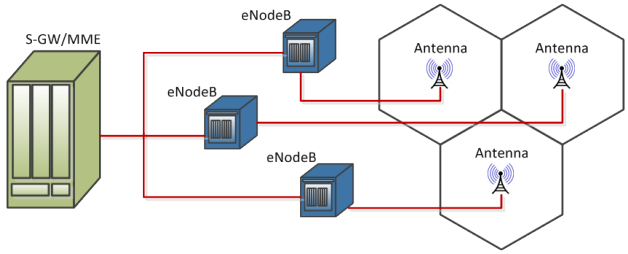


Figure 4.1: Small Cell configuration.

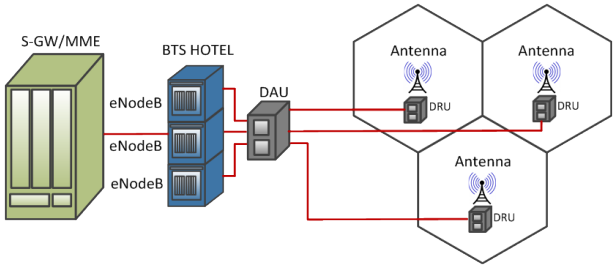


Figure 4.2: Virtual Cell configuration.

4.1.1 Small Cell.

Small Cells are small base stations that deliver capacity to a small coverage area. They can support and provide coverage in hot-spots or in-building as compared to macro cells [65]. With a Small Cell deployment, the total system bandwidth requirements increase proportionally to the number of nodes. However, having more nodes also increases the inter-cell interference, which in turn reduces the achievable spectrum efficiency per user. The placement of the Small Cell nodes has a significant impact on the system performance [66].

This placement includes increased signaling, ports on the MME, Service- gateway and Network Management databases. The basic Small Cell architecture can be seen in Fig. 4.1.

4.1.2 Virtual Cell.

The basic Virtual Cell distributed network architecture consists of a BTS Hotel with multiple remote units as seen in Fig. 4.2. In a traditional DAS architecture the remote units are connected to the centrally located BTS, and the downlink signal is broadcast to all the cells. Similarly, in the uplink direction, the received signals from the different remote units will be combined at the base station [67]. The division of the coverage area into smaller cells results in improved performance through

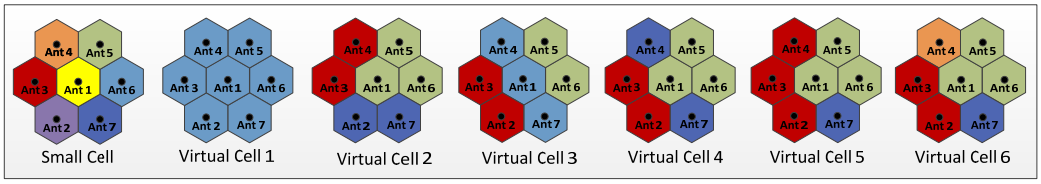


Figure 4.3: Small Cell and 6 different Virtual Cell architectures.

minimal path loss and optimized transmission power. With the use of multiple antennas, the path loss decreases and less downlink power from the base station is required to cover the same area. Similarly, less uplink power from the mobile unit is required to communicate with the DAS remote units, thereby improving the mobile battery life. In DAS, several remote antenna elements are connected to an eNodeB through a fiber optic cable, LAN cabling or microwave link via a DAU, as shown in Fig. 4.2. The remote antenna elements are identified as DRUs in Fig. 4.2.

A Virtual Cell is a remote node that has access to all of the system's resources at the eNB. The eNB resources can be routed to the remote Virtual Cells via the distributed network. As an example, sectors can be routed to a particular Virtual Cell or carrier frequency bands could be activated at a particular cell, independent of the other Virtual Cells.

4.1.3 System Model.

We considered a two-ring hexagonal cellular system with nineteen remote antenna units, wherein the distance between antennas is set at 300 meters for in-building cases. The Small Cell architecture requires an individual eNodeB for each antenna unit. The simulations of the Virtual Cell (VC) architecture are based on six distinct scenarios: 1) VC1, seven central antennas ($\{Ant1, 2, ..., 7\}$) are supported by one eNodeB, 2) VC2, three different groups of antennas ($\{Ant1, 5$ and $6\}$,$\{Ant3$ and $4\}$, $\{Ant2$ and $7\}$) are separately connected to three eNodeBs, 3) VC3, three different groups of antennas ($\{Ant1, 4$ and $7\}$,$\{Ant2$ and $3\}$, $\{Ant5$ and $6\}$) are separately connected to three eNodeBs, 4) VC4, three different groups of antennas ($\{Ant1, 5$ and $6\}$, $\{Ant2$ and $3\}$, $\{Ant4$ and $7\}$) are separately connected to three eNodeBs, 5) VC5, three different groups of antennas ($\{Ant1, 5$ and $6\}$, $\{Ant7\}$, $\{Ant2, 3$ and $4\}$) are separately connected to three eNodeBs, 6) VC6, four different groups of antennas ($\{Ant1, 5$ and $6\}$,$\{Ant2$ and $3\}\{Ant4\}\{Ant7\}$) are separately connected to four eNodeBs. VC1 is a traditional DAS implementation with a 1:7 simulcast ratio. All these architectures are shown in Fig. 4.3.

The performance of the Small Cell and Virtual Cell architectures is analyzed through system level simulations. An eNodeB allocates the available RBs to UEs by estimating the signal and uplink power level of the UEs. The simulation system parameters, as shown in Table 4.1, are chosen to investigate the technical performance of the various architectures.

PARAMETERS	VALUE
Channel Bandwidth	5 *MHz*
Carrier Frequency	2.14 *GHz*
FFT size	1024
Number of Resource Blocks	25
Subcarrier Spacing	15 *kHz*
Cellular Layout	Hexagonal grid, 19 Antennas
Inter-Antenna Distance	300 meters
Propagation loss	128.1+37.6 $\log_{10}(R(km))$
White Noise Power Density	-174 *dBm/Hz*
Scheduling	Proportional Fair,
TTI	1 *ms*
Transmission scheme	SISO
Antenna Transmission Power	1 *W*
Noise Figure	10 *dB*

Table 4.1: Simulation Parameters.

At a given transmission time interval (TTI) for the LTE simulation, the eNodeB in a cell gathers the CQI information of UEs and allocates the frequency RBs to each UE, using various scheduler techniques.

- **Path-loss Model:** The propagation model is used to predict the path loss. The path-loss model is a simple model that calculates the path loss of the indoor environment under ideal conditions. Path loss is usually expressed in dB. In its simplest form, the path loss can be calculated using the formula:

$$L = 10n\log_{10}(d) + 20\log_{10}(4\pi f) + c \tag{4.1}$$

  where L is the path-loss in dB and is represented by the path-loss exponent n=3.76 for the in-building simulations. d is the distance between the transmitter and the receiver, measured in kilometers, c is a constant which takes into account the system losses and f is the carrier frequency.

- **Received Signal Strength (RSS):** Received Signal Strength is usually expressed in dBm. In its simplest form, the RSS can be calculated using the formula:

$$RSS(dBm) = P_{Tx}(dBm) - L(dBm) \tag{4.2}$$

  where RSS is the received signal strength in dBm, P_{Tx} is antenna transmission power and L is the path-loss in dB.

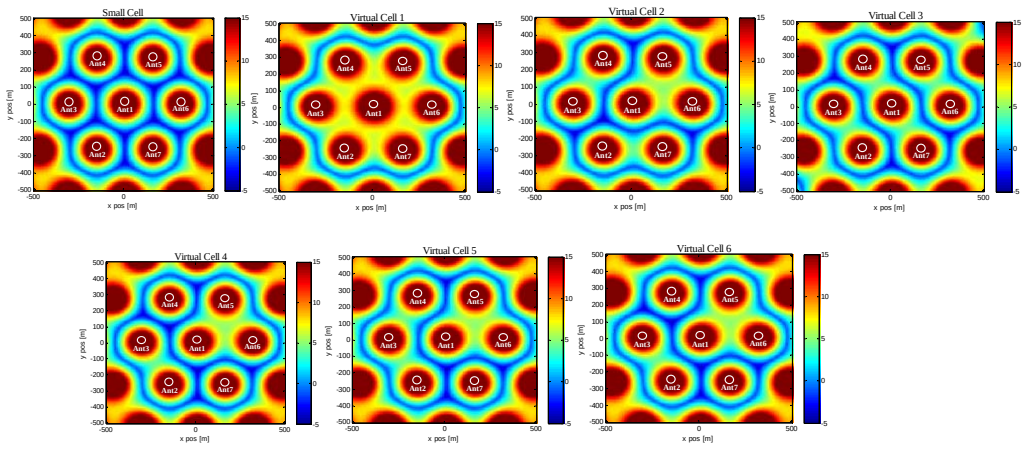


Figure 4.4: SINR distribution of different solutions.

4.1.4 Comparison of Results for Small Cell and Virtual Cell.

SINR Distribution:

Signal to interference plus noise (SINR) ratio is usually expressed in dB. In its simplest form, the SINR can be calculated using the formula:

$$SINR = RSS - RISS - N_{th} \tag{4.3}$$

where SINR is signal to interference plus noise ratio, $RISS$ is received interference signal strength and N_{th} is thermal noise in dB which is calculated as follows:

$$N_{th} = N_{th-density} - 30 + 10\log_{10}(BW) + NF \tag{4.4}$$

where $N_{th-density}$ is white noise power density in dBm, BW is bandwidth in Hz and NF is noise figure in dB. The SINR with respect to the remote antenna configuration is shown in Fig. 4.4. Red represents the largest SINR and blue represents the weakest SINR. The SINR distribution of each scenario is presented to highlight the inter-antenna interference, and it can be seen that the best downlink SINR coverage is achieved with VC1. This is expected because there is no inter-antenna interference amongst the 7 central antennas with VC1. The Small Cell architecture, however, has the worst SINR distribution, because of the inter-antenna interference. No interference coordination between the Small Cells is assumed, as this would impact the useable capacity. In the VC2, 3 and 4 scenarios, the simulcast ratio is smaller than in VC1, and the cells are grouped together and fed by a unique eNodeB. The VC5 and 6 scenarios would be used in a hotspot application whereby an entire eNodeB resource would be allocated to one remote unit.

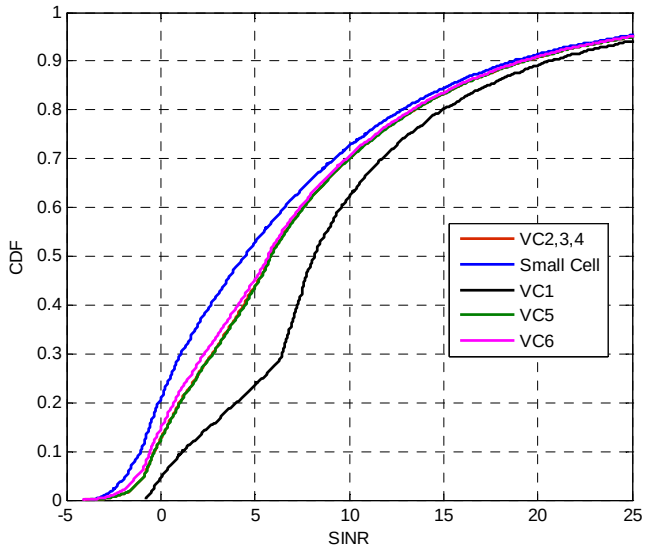


Figure 4.5: SINR distribution of different solutions in terms of CDF.

Fig. 4.5 shows the SINR distribution of each solution in terms of CDF: 24 percent of coverage area in VC1, 45 percent of coverage area in VC2, 3, 4, 5 and 6 and 53 percent of coverage area in Small Cell has a SINR less than 5 dB.

SINR to CQI mapping:

CQI is a measurement of the communication quality of wireless channels. CQI can be a value (or values) representing a measure of channel quality for a given channel. Typically, a high value CQI is indicative of a channel with high quality and vice versa. A CQI for a channel can be computed by making use of performance metric, such as an SINR. These values can be measured for a given channel and then used to compute a CQI for the channel. Fifteen unique modulation and coding scheme (MCS) levels are classified for LTE, which are associated with 15 CQI values. The CQIs use coding rates between 1/13 and 1 when integrated with QPSK, 16-QAM and 64QAM modulations. The SINR-to-CQI mapping is shown in Fig. 4.6. The attained CQIs are later floored to obtain the integer CQI values that are reported back to the eNodeB.

In Fig. 4.7, the CQI is mapped with respect to the remote antenna configuration and the CQI distribution for the Small Cell, VC1, VC2,..., VC6 scenarios is demonstrated. VC1 has the best CQI distribution in the central coverage area, which is expected since there is no inter-antenna interference. The Small Cell system has the worst CQI distribution, since no inter-antenna interference mitigation techniques are employed.

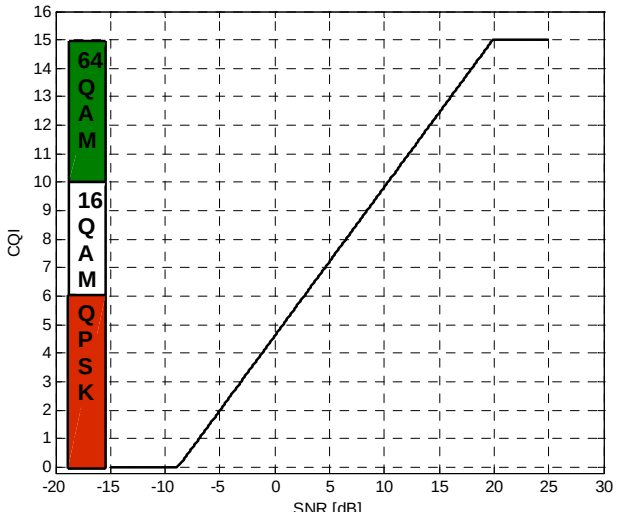


Figure 4.6: SINR-to-CQI mapping.

Shannon Capacity:

Capacity theorem is used to determine the total information that can be transmitted over the communication channel. According to the Shannon theorem, the maximum rate of information C for a given communication system is known as channel capacity in bit per second (bps) and stated as:

$$C = B \log_2(1 + SINR) \tag{4.5}$$

where B is the bandwidth of the AWGN channel. The spectral efficiency is defined by:

$$\text{Spectral Efficiency} = C/B = \log_2(1 + SINR) \tag{4.6}$$

Fig 4.8 displays the spectral efficiency of Small Cell, VC1, VC2,..., VC6 scenarios. The simulation results show that over the central coverage area, the VC1 configuration is more spectrally efficient than VC2, VC3,..., VC6 or the Small Cell configuration.

Although beneficial in terms of spectral efficiency, the configuration with X number of eNodeB may not achieve as high a throughput per user as the configuration with larger X due to the low effective bandwidth per antenna. i.e., VC2 outperforms VC1, VC6 outperforms VC2, and Small Cell outperforms VC6 in terms of throughput per users.

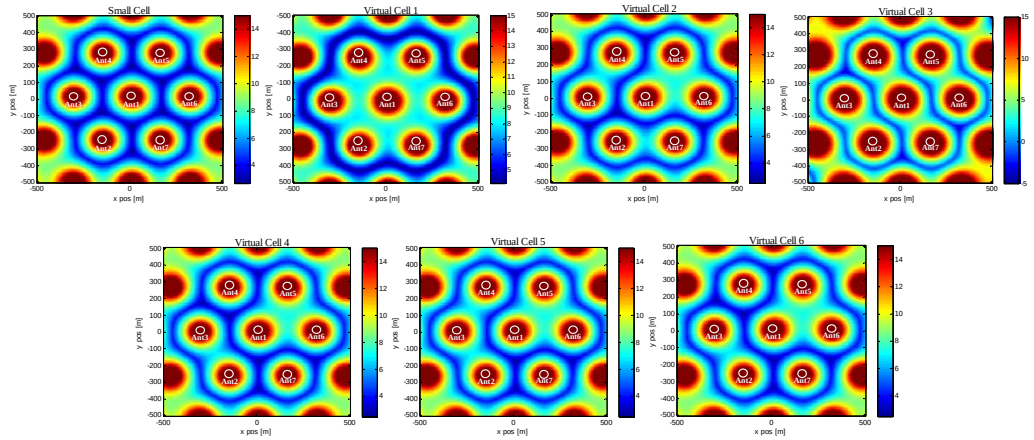


Figure 4.7: CQI coverage of 7 central antennas

		VC1 (1 eNB)	VC2 (3 eNBs)	VC 3 (3 eNBs)	Small cell (7 eNBs)
SINR_Ave		13.421	11.915	12.001	9.512
Useable_Capacity_Ave (Mbps)		14.4889	13.001	13.204	10.8494
System_Capacity_Ave (Mbps)		14.4889	26.002	26.408	32.5482
Spectral_Efficiency_Ave (Mbps/Hz)		3.8501	3.6910	3.7006	3.3940
Modulation percentage	QPSK	2.03%	20.65%	19.67%	42.25%
	16-QAM	35.12%	24.54%	25.32%	12.04%
	64-QAM	62.85%	54.81%	55.01%	45.71%

Table 4.2: Simulation Results of single user scenario.

4.2 Simulation Scenarios.

4.2.1 Single-User.

Fig. 4.9 shows the single-user scenario when UE1 moves from Antenna 5 to Antenna 2.

The results for VC1, VC2, VC3 and Small Cell are shown in Table 4.2 where Usable Capacity is defined as available capacity per eNB and System Capacity is defined as the capacity of all eNBs. Results show that the average SINR for VC1 is around 4 dB better than Small Cell and 1.5 dB better than VC2 and VC3. Figures 4.10 and 4.11 show the CQI and throughput experiences when UE1 moves from Antenna 5 to Antenna 2. Most CQI values used in VC1 were anywhere from 10 to 15 due to better SINR values, which results in a better modulation format. UE1 uses 64-QAM only 45.71 percent of the time in the Small Cell scenario and 54.81 percent in the VC2 scenario. However, in the VC1 deployment, 64-QAM is used 62.85 percent, which shows optimized average

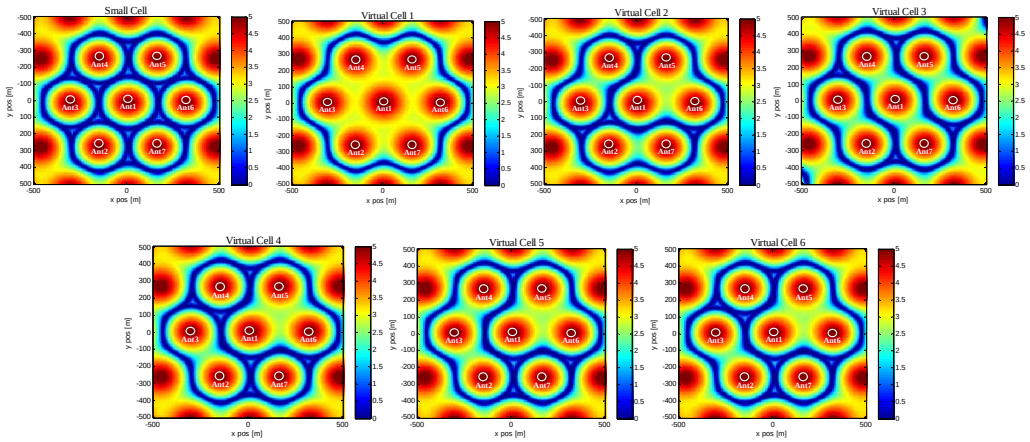


Figure 4.8: Virtual Cell vs. Small Cell Spectral Efficiency.

simulation useable capacity and average cell spectral efficiency for VC1. In terms of simulation system capacity, Small Cell outperforms VC1 and VC2 by using more eNodeBs.

4.2.2 Multi-User

We design three different user distributions to highlight the user throughput with effective bandwidth per user. Two different low and high user load densities are also considered (70, 560 users respectively), which are distributed throughout the area for each different user distributions. Note that the useable capacity is defined as the available capacity per eNB.

- **Uniform Distribution:** In this scenario, the users are uniformly distributed inside a central area containing seven cells (Antenna 1, 2,..., 7) (Fig. 4.12(a)). 10 users and 80 users per cell are distributed for low and high densities, respectively.
- **Uniform Distribution Including One Hot-Spot:** In this scenario, there are large numbers of users inside cell 7, where the other users are distributed uniformly inside the other 6 cells (Fig. 4.12(b)). In low density, cell 7 supports 28 users, where each of the 6 other cells supports 7 users. In high density, cell 7 supports 224 users, where each of the 6 other cells supports 56 users.
- **Uniform Distribution Including Two Hot-Spots:** In this scenario, there are large and equal numbers of users inside cell 7 and 4, where the other users are distributed uniformly inside the other 5 cells (Fig. 4.13(c)). In low density, each two cell 7 and 4 supports 20 users, where each of the 5 other cells supports 6 users. In high density, each two cell 7 and 4 supports 160 users, where each of the 5 other cells supports 48 users. Fig. 4.13 shows the high load density, 560 users are distributed inside 7 antennas area.

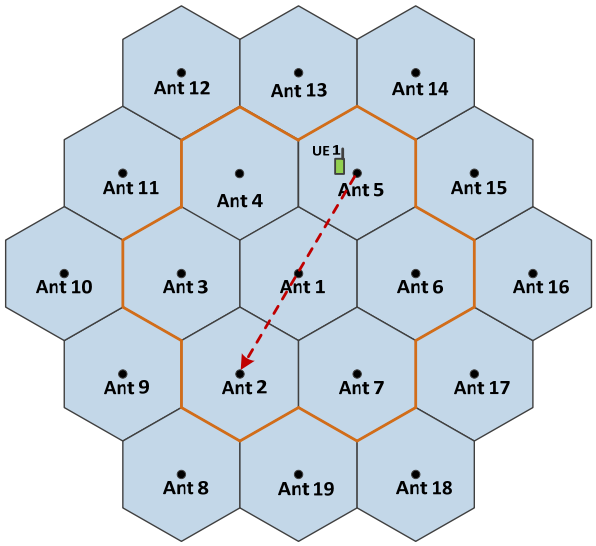


Figure 4.9: Structure of Single User Simulation

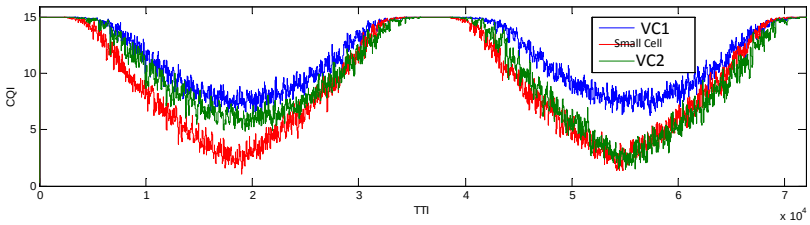


Figure 4.10: CQI report is sent by UE1 in single user simulation

The results of VC1, VC2,..., VC6 and Small Cell are shown in Table 4.3, 4.4, 4.5, 4.6 and 4.7. Note that, the color allocation for each row is done independently from the other row colors and is based on the numbers obtained in each cell. The modulation format percentage results of VC1, VC2,..., VC6 and Small Cell for different user distribution are shown in Table 4.3. The better SINR values in VC1 translate into higher CQIs (between 10 to 15) resulting in a higher modulation format percentage at every load density. Under both low and high load densities in uniform users distribution, only 12 percent of Small Cell users are using 64-QAM samples; 22 percent of VC6 (7 eNBs); 31-32 percent of VC2, VC3, VC4 and VC5(3 eNBs); yet 51 percent of users in VC1 (1 eNB) are using 64-QAM. VC2, VC3, VC4 and VC5 are similar in performance to each other because the users were assumed to be uniformly distributed in the geographic area. By considering only hot-spot users in non-uniform distribution scenarios, due to the different SINR and CQI distributions inside hot-spot areas, 3 eNB configurations (VC2, VC3, VC4 and VC5) are not similar in performance to

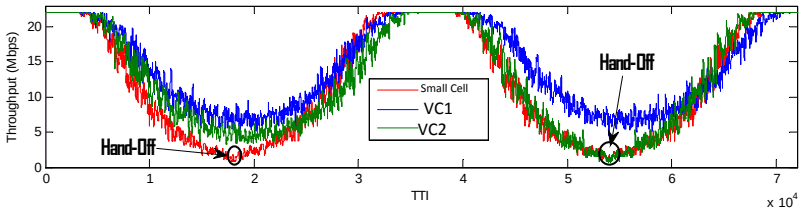


Figure 4.11: UE1 throughput in single user simulation

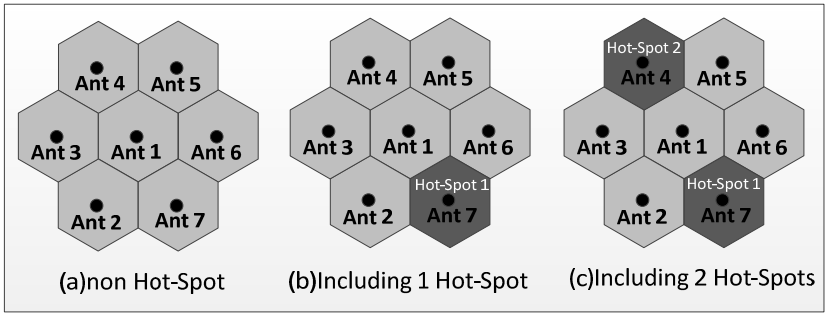


Figure 4.12: different user distributions.

each other. This is because of different SINR and CQI distribution inside the hot-spot area.

Table 4.4 demonstrates the system performance for VC1,...,VC6 and Small Cell using scheduler proportionally fair. Because of the larger number of eNodeBs used in the coverage area, the Small Cell system (7 eNodeBs) has a higher average user throughput than all VCs architecture (1, 3 and 4 eNodeBs). As it shown in Table 4.4, by decreasing the number of eNodeB the average user's throughput is degrading.

Note that, as a result of using two separate eNodeB to assign RBs to both hot-spot 1 and 2 users in VC6, the average throughput of hot-spots' users of VC6 is very close to the highest throughput which are obtained from Small Cell. However, VC6 uses less number of eNodeB than Small Cell. Similarly, the average throughput of hot-spot 1 users of VC5 is very close to the highest throughput which is obtained from Small Cell. However, VC5 uses fewer eNodeB than Small Cell.

When comparing all Virtual Cell configurations with 3 eNodeBs (VC2, 3, 4 and 5), all the average throughput of all users are the same; yet, a better hot-spot users throughput is obtained via VC5. Tables 4.5, 4.6 and 4.7 show the results of the key performance indicators for the different scenarios. The average throughput per user and usable capacity were included as these are good indicators of customer satisfaction. For a given coverage area, (M:1) is defined as "M antennas are supported by one eNodeB". This is commonly referred to as the simulcast ratio per eNodeB. As an example, VC6 can be seen as utilizing three configurations: (3:1), (2:1), (1:1) and (1:1). Note that, Table 4.6

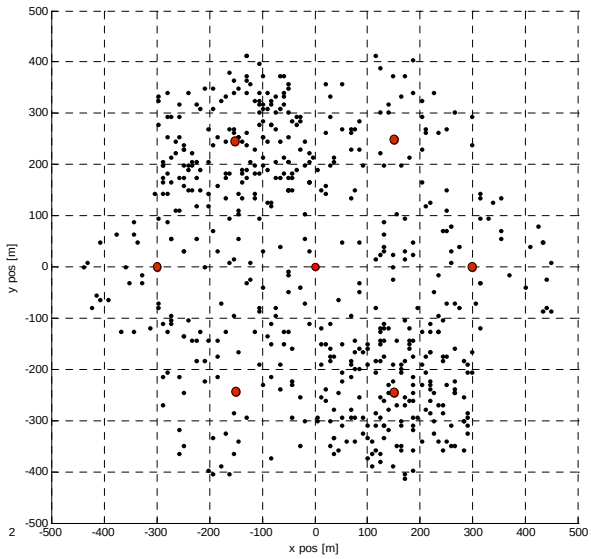


Figure 4.13: High density Uniform distribution including two hot-spots

shows the results when the hot-spot 1 is considered as one of the M antennas (cells), and Table 4.7 shows the results when both the hot-spot 1 and 2 are considered as two of the M antennas (cells). Since, at least 2 antennas (cells) are required to support both hot-spot cells, the (1:1) configuration is not considered in Table 4.7.

The SINR degrades as we reduce M because of the increased inter-cell interference. The VC1 (7:1) configuration outperforms the other configurations in terms of useable capacity and spectral efficiency due to the higher SINR value. The (1:1) configuration for the Small Cell and VC5 antenna 7 outperforms the other configurations in terms of average throughput per user because of a dedicated eNodeB per cell, which results in a larger effective bandwidth per user. The (1:1) configuration supports the lowest number of users per eNodeB in comparison with the other (M:1) configurations.

The results indicate that a Virtual Cell that dynamically alters the simulcast ratio has the ability to achieve throughput performance of a Small Cell at a hot spot while achieving the highest spectral efficiency throughout the coverage area. A Virtual Cell in a distributed network architecture is the most cost effective solution as it routes the capacity on demand. Centralizing the base station resources with a distributed network has the added benefit of capacity scalability and centralizing the backhaul infrastructure. A Small Cell deployment with interference mitigation techniques such as a SON would be required in order to improve the spectral efficiency and reduce the hand-offs. However, this would come at the expense of reduced capacity as the resource blocks available per Small Cell would be restricted.

%	Distribution type		VC1 1 eNB	VC2 3 eNB	VC3 3 eNB	VC4 3 eNB	VC5 3 eNB	VC6 4 eNB	SmallCell 7 eNB
QPSK	Uniform	All cells	9	28	29	28	29	43	55
	Uniform including 1 hot-spot	All cells	12	38	38	41	42	49	54
		Hot-Spot cell 1	11	48	47	54	53	54	53
	Uniform including 2 hot-spots	All cells	11	44	42	47	47	50	54
		Hot-Spot cell 1	9	49	48	51	52	53	52
		Hot-Spot cell 2	10	49	48	52	49	52	52
16QAM	Uniform	All cells	40	41	39	40	40	35	33
	Uniform including 1 hot-spot	All cells	33	37	38	36	34	33	32
		Hot-Spot cell 1	36	33	36	34	34	33	36
	Uniform including 2 hot-spots	All cells	36	27	31	32	33	33	34
		Hot-Spot cell 1	37	34	33	34	35	34	31
		Hot-Spot cell 2	38	33	34	32	33	34	32
64QAM	Uniform	All cells	51	31	32	32	31	22	12
	Uniform including 1 hot-spot	All cells	55	25	24	23	22	18	12
		Hot-Spot cell 1	53	19	17	12	13	13	11
	Uniform including 2 hot-spots	All cells	53	29	27	21	20	17	14
		Hot-Spot cell 1	54	17	19	15	13	13	17
		Hot-Spot cell 2	52	18	18	16	19	14	16

Table 4.3: Modulation Percentage.

4.3 Self-Optimizing Network of Virtual Cell.

In the Virtual Small Cell network, it is important to allocate the DRUs to cope with dynamically changing traffic and to balance the traffic for each eNB. Without proper DRU allocation there may be cases of unbalanced traffic where the number of users with a data rate below a specified threshold is increased for a specific eNB, even if the other eNBs traffic resources are not utilized. The networks performance is defined by the number of performance factors from different parts of the network and this, ultimately determines the cost function for optimization. Operators may have different business goals and service levels defined. Depending on these considerations, cost effective and efficient network performance might vary from operator to operator. Consequently, QoS metrics could be defined and mapped to a set of network performance cost factors.

In the following sections, two practical scenarios will be discussed in order to demonstrate the advantage of SON of Virtual Cell. Each DRU area is shown in figures 4.14, 4.15, 4.19 and 4.20 as a cell covered by a thick black line. Each color indicates DRU cells allocated to a specific eNB, and all yellow cells are external interference DRUs.

LOWER	LOW		HIGH						HIGHER
#Active Users	Distribution type		VC1 1eNB	VC2 3eNB	VC3 3eNB	VC4 3eNB	VC5 3eNB	VC6 4eNB	Small Cell 7eNB
70	Uniform	SINR Ave	8.44	7.62	7.64	7.64	7.77	7.59	7.39
		Ave Throughput per user (Mbps)	0.17	0.37	0.39	0.38	0.40	0.47	0.81
	Uniform including 1 hot-spot	SINR Ave	8.66	7.97	7.93	7.83	7.94	7.83	7.72
		Ave Throughput per user (Mbps)	0.18	0.45	0.44	0.45	0.45	0.55	0.93
		Ave-SINR of hot-spot 1 users	8.33	7.78	7.71	7.52	7.58	7.54	7.58
		Ave-Throughput per user of hot-spot 1 users	0.17	0.29	0.23	0.27	0.33	0.33	0.35
	Uniform including 2 hot-spots	SINR Ave	8.41	7.68	7.68	7.58	7.72	7.55	7.46
		Ave Throughput per user (Mbps)	0.17	0.40	0.41	0.41	0.39	0.52	0.84
		Ave-SINR of hot-spot 1 users	8.13	7.15	7.35	7.01	7.07	7.01	6.96
		Ave-Throughput per user of hot-spot 1 users	0.15	0.33	0.19	0.20	0.37	0.37	0.38
		Ave-SINR of hot-spot 2 users	8.59	7.99	7.88	7.71	8.08	7.65	7.72
		Ave-Throughput per user of hot-spot 2 users	0.18	0.42	0.22	0.25	0.35	0.52	0.55
560	Uniform	SINR Ave	8.61	7.93	7.87	7.85	7.94	7.83	7.65
		Ave Throughput per user (Mbps)	0.02	0.05	0.05	0.05	0.05	0.07	0.11
	Uniform including 1 hot-spot	SINR Ave	8.56	7.87	7.81	7.75	7.81	7.73	7.63
		Ave Throughput per user (Mbps)	0.02	0.05	0.05	0.05	0.05	0.07	0.11
		Ave-SINR of hot-spot 1 users	8.49	7.87	7.79	7.63	7.65	7.65	7.66
		Ave-Throughput per user of hot-spot 1 users	0.02	0.04	0.03	0.03	0.05	0.05	0.05
	Uniform including 2 hot-spots	SINR Ave	8.46	7.73	7.70	7.61	7.71	7.58	7.47
		Ave Throughput per user (Mbps)	0.02	0.05	0.05	0.05	0.05	0.06	0.10
		Ave-SINR of hot-spot 1 users	8.58	7.90	7.92	7.70	7.67	7.68	7.69
		Ave-Throughput per user of hot-spot 1 users	0.02	0.05	0.03	0.03	0.06	0.06	0.06
		Ave-SINR of hot-spot 2 users	8.49	7.78	7.76	7.48	7.83	7.49	7.52
		Ave-Throughput per user of hot-spot 2 users	0.02	0.05	0.03	0.03	0.04	0.06	0.06

Table 4.4: Simulation Results for Multi-User.

4.3.1 Conference Room Scenario.

In Fig. 4.14, one floor of the Conference Hall is covered by 5 eNB (5 different colors). At a regular (i.e., non-peak) time, all 90 active users are distributed uniformly over the 18 DRU areas. However, several mobile users decide to move to a small conference room at a specific time and expect the same quality of service they had previously experienced. Fig. 4.16 presents SINR distribution for the initial DRUs distribution. As indicated in Fig. 4.14, the conference room is located next to the interfering cell (yellow), which degrades the QoS. This QoS reduction is more pronounced when several users are in the conference room at the time of the conference. The results reveal that there is a significant degradation in the user throughput, performance, at the time of the conference, which is depicted in Table 4.8.

Two solutions are investigated to overcome throughput degradation: the first is a Traditional Solution, and the second is SON.

- **Traditional Solution:** In order to permanently cover the conference room, this solution requires adding an extra eNB to the eNB hotel. Fig. 4.15 displays the DRU allocation and shows sufficient capacity. Fig. 4.17 displays the SINR distribution; the conference users would experience a lower SINR with this solution.
- **SON of Virtual Cell:** This approach uses a SON algorithm to find the best DRU allocation

HIGHER	HIGH		LOW		LOWER
Uniform	#active users	7:1(VC1) (Ant1,…,7)	3:1(VC5) (Ant2, 3, 4)	2:1(VC2) (Ant2, 7)	1:1(VC5) and Small Cell (Ant 7)
SINR Ave	70	8.44	7.65	7.06	7.65
	560	8.61	7.92	7.83	7.58
Ave Throughput per user (Mbps)	70	0.17	0.30	0.37	0.88
	560	0.02	0.04	0.06	0.11
Useable Capacity(Mbps)	70	12.09	9.26	7.50	8.85
	560	13.32	10.68	10.02	9.18
Spectral efficiency(Mbps/Hz)	70	3.23	3.11	3.01	3.11
	560	3.26	3.15	3.14	3.10

Table 4.5: Simulation Results for Multi-User (Uniform).

HIGHER	HIGH		LOW		LOWER
Uniform including 1 hot-spot	#active users	7:1 (VC1) (Ant1,…,7)	3:1(VC5) (Ant2, 3, 4)	2:1(VC2) (Ant2, 7)	1:1(VC5)) and Small Cell (Ant 7)
SINR Ave	70	8.66	7.81	7.62	7.58
	560	8.56	7.77	7.80	7.65
Ave Throughput per user (Mbps)	70	0.18	0.23	0.28	0.33
	560	0.02	0.03	0.04	0.04
Useable Capacity(Mbps)	70	13.02	10.97	9.86	9.46
	560	13.55	11.75	11.42	11.04
Spectral efficiency(Mbps/Hz)	70	3.27	3.13	3.10	3.10
	560	3.25	3.13	3.13	3.11

Table 4.6: Simulation Results for Multi-User (Uniform including 1 hot-spot).

configuration and does not require adding any additional eNB resources. Fig. 4.15 shows the new DRU allocation after running a SON algorithm. As is depicted, the three adjacent DRUs to the conference room and the conference room DRU are still operating on the same eNB resource. Fig. 4.17 shows the SINR distribution for the SON solution and users do not experience high interference from other eNBs.

Table 4.8 shows that both solutions outperform the initial DRU allocation in terms of the conference cell users throughput. The traditional solution outperforms the SON in terms of system capacity and average total users throughput because of the extra eNB resource added.

However, the SON outperforms the traditional solution in terms of average conference room users throughput as a result of intelligently allocating DRUs to eNB resources. SON also outperforms the traditional solution in terms of SINR and useable capacity because of a lower total number of eNB resources.

Fig. 4.18 shows the CDF of users' throughput for the initial DRU allocation along with the traditional solution and SON solution. 42 percent of users in the initial DRU allocation have less than 0.5 $Mbps$; 39 percent of SON users and 33 percent of users with the traditional solution have less than 0.5 $Mbps$.

HIGHER	HIGH	LOW		LOWER	
Uniform including 2 hot-spots	#active users	7:1(VC1) (Ant1,…,7)	3:1(VC3) (Ant1, 4, 7)	2:1(VC4) (Ant 4, 7)	1:1
SINR Ave	70	8.41	7.58	7.36	N
	560	8.46	7.77	7.59	N
Ave Throughput per user (Mbps)	70	0.17	0.20	0.22	N
	560	0.02	0.02	0.03	N
Useable Capacity(Mbps)	70	12.20	9.50	9.18	N
	560	12.88	10.61	10.33	N
Spectral efficiency(Mbps/Hz)	70	3.23	3.10	3.06	N
	560	3.24	3.13	3.10	N

Table 4.7: Simulation Results for Multi-User (Uniform including 2 hot-spots).

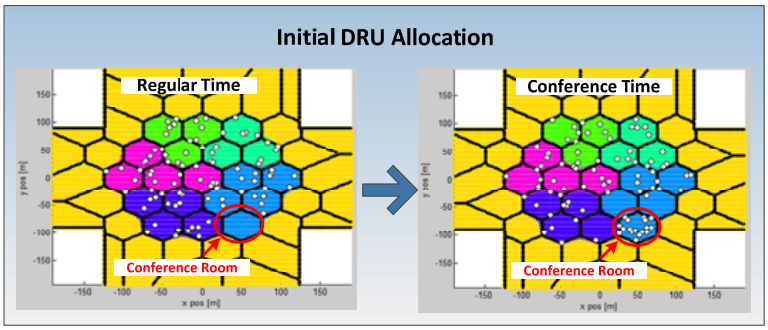


Figure 4.14: Users' distribution at both regular and conference time

4.3.2 Stadium/Parking Lot Scenario.

Let's assume a Stadium/Parking Lot Scenario where over the course of an event, users move from the parking lot to the stadium and vice-versa.

In the initial DRU allocation, 3 eNBS are located in the parking lot and 3 eNBS in the stadium for a total of 6 eNBS as shown in Fig. 4.19. Both distribution cases are considered: during 'parking-time' when most users are located in the parking lot; and during the 'stadium-time' (or 'game-time') when most users are based in the stadium. The initial SINR distribution is shown in Fig. 4.21. Because the same number of eNBs (3) is allocated to the stadium and the parking lot respectively, there is a significant gap between the parking and stadium users' throughput as depicted in Table 4.9. The same results are demonstrated for the initial DRU allocation at the 'stadium-time'.

SON is one approach to address this unbalanced throughput. Fig. 4.20 presents two different DRU allocation results after running the SON algorithm at both the 'parking-time' and 'stadium-time'. Table 4.9 shows the results using SON at both the 'parking-time' and the 'stadium-time' scenarios.

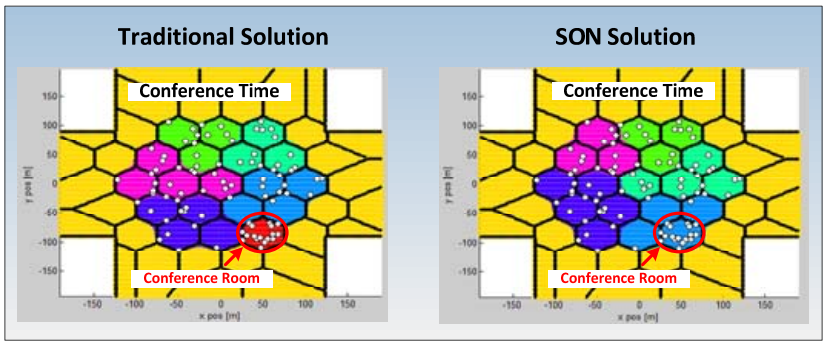


Figure 4.15: DRU allocation for both Traditional and SON solutions

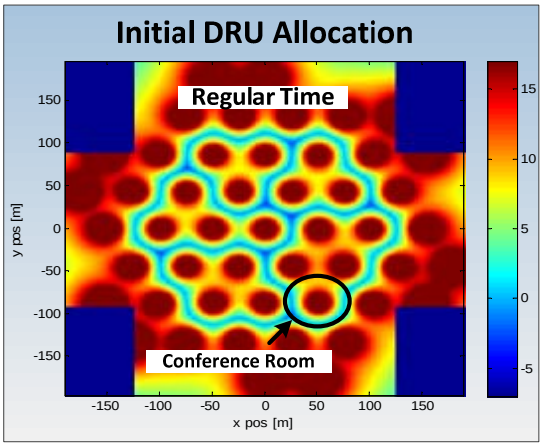


Figure 4.16: SINR distribution for the initial DRU allocation

When a large number of active users (80) are located in the parking lot space, the average parking lot users' throughput has improved 42 percent using SON. Even though the average stadium users' throughput has degraded by 56 percent for the small number of active users (17), the overall average users' throughput is roughly the same after using SON.

After implementing SON at the 'stadium-time', the average stadium users' throughput has improved by 37 percent. On the other hand, the average parking lot users' throughput has degraded 50 percent when a small number of active users (20) are located in the stadium. Again notice that the overall average users' throughput is roughly the same after using SON.

Fig. 22 demonstrates the CDF profile of users' throughput for the initial DRU allocation and after using SON at the 'parking-time'. Lets assume that 0.75 $Mbps$ is a throughput constraint to distinguish satisfied and unsatisfied users.

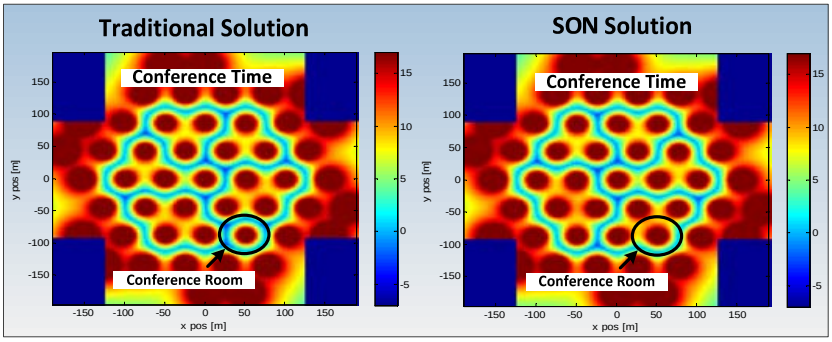


Figure 4.17: SINR distribution for both Traditional and SON solutions

	Initial DRU Allocation-Regular Time	Initial DRU Allocation-Conference Time	Traditional Solution-Conference Time	SON Solution-Conference Time
Ave-User-SINR	7.8698	7.5824	7.2617	7.6875
Ave-Usable Capacity	13.0417	12.3832	11.8318	12.6772
System Capacity	65.2085	61.9161	70.9907	63.3859
Ave-Total Users-Throughput(Mbps)	0.7245	0.6880	0.7888	0.7043
Ave-Conference Users-Throughput(Mbps)	NONE	0.4697	0.6561	0.7407

Table 4.8: Conference Room Scenario Results.

Therefore, at the 'parking-time', 66 percent of users are dissatisfied in the initial DRU allocation whereas only 48 percent of users are dissatisfied using SON. Fig. 23 indicates the CDF profile of the users' throughput. At the 'stadium-time', 57 percent of users are dissatisfied with the initial DRU allocation, whereas only 46 percent of users after implementing SON.

4.4 Summary.

The performance analysis for a Virtual Cell distributed architecture and a Small Cell architecture was undertaken. The results indicate that the inter-cell interference from Small Cells have a significant impact on the user performance. Without any coordination between Small Cells, the SINR over the geographic area is significantly poorer. This results in the inability for the users to operate at the highest spectral efficiency available from LTE. The increased hand-offs from a poor SINR profile will increase the burden on the network and negatively impact the user experience.

Small Cells provide increased capacity to a given geographic area, whereas a traditional DAS architecture has been viewed as a coverage solution. However, a Virtual Cell distributed network that has the ability to provide capacity on demand by altering the simulcast ratio will, ultimately,

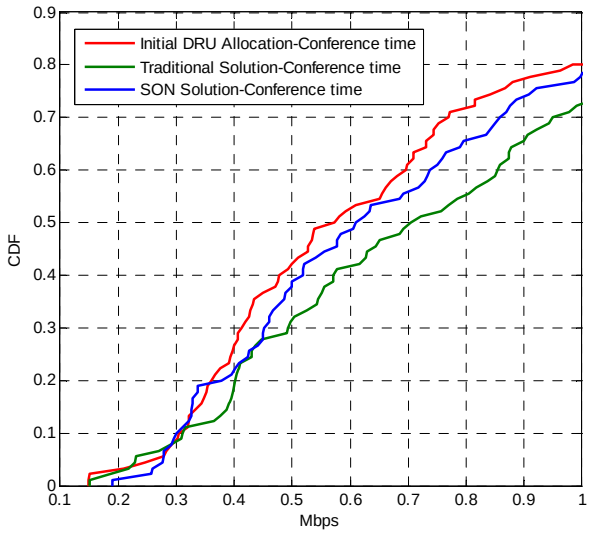


Figure 4.18: CDF of users' throughput for Conference Room Scenario

	Initial DRU Allocation-Parking Time	Parking Lot DRU Allocation-Parking Time	Initial DRU Allocation-Stadium Time	Stadium DRU Allocation-Stadium Time
Ave-Parking Lot Users Throughput(Mbps)	0.547 +42%for 80 users	0.776	1.61 -50% for 20 users	0.80
Ave-Stadium Users Throughput(Mbps)	2.3 -56% for 17 users	1.003	0.6280 +37% for 68 users	0.8546
Ave-Total Users Throughput(Mbps)	0.848	0.821	0.853	0.843

Table 4.9: Parking/Stadium Scenario Results.

provide the most optimal solution. A Virtual Cell distributed network provides the throughput demands at peak loads and maximizes the user experience throughout the entire coverage area. This investigation demonstrates that the throughput data rate of a Virtual Cell distributed network is the same as that of a Small Cell at a hotspot yet at a fraction of the required number of eNodeBs. The Virtual Cell distributed network architecture provides the best spectral efficiency and subsequently the highest useable capacity.

DRU allocation of eNB resources is examined in order to balance the traffic for a Virtual Cell network. DRU allocation of eNB resources is considered an effective technique for load balancing traffic on a network. Compact DRU allocations are proposed to reduce handoffs and interferences. DRU allocation of eNB resources is formulated as an integer linear programming problem, which minimizes a performance based cost function.

SON algorithms are required to solve the DRU allocation of eNB resources. A significant throughput improvement is demonstrated using the SON algorithm.

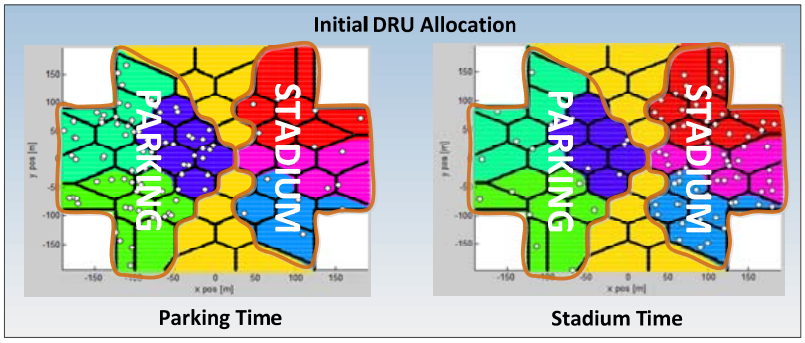


Figure 4.19: Users distribution at both Parking and Stadium time

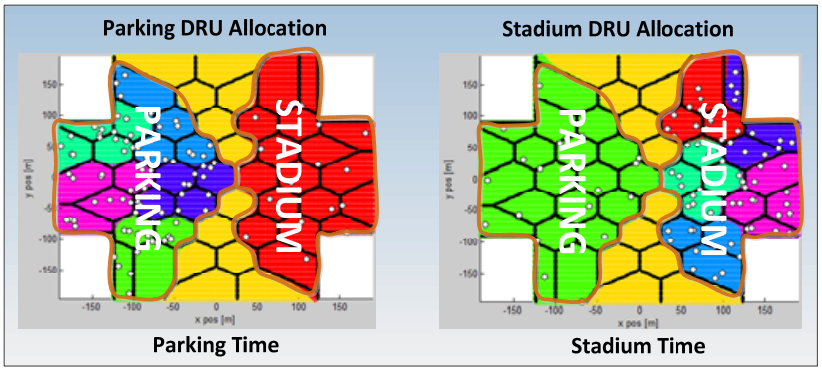


Figure 4.20: SON DRU allocation for both Parking and Stadium time

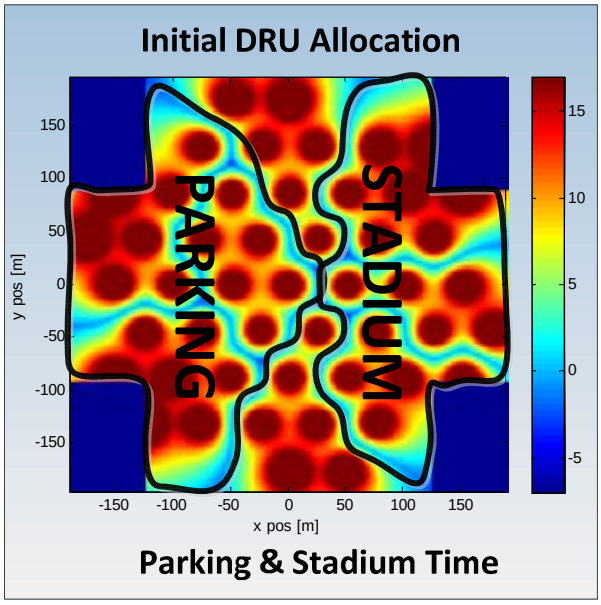


Figure 4.21: SINR distribution for the primarily DRU allocation

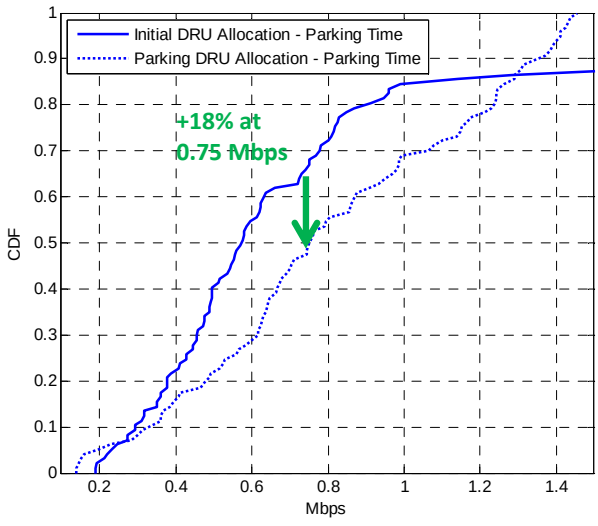


Figure 4.22: CDF of users' throughput for Parking time

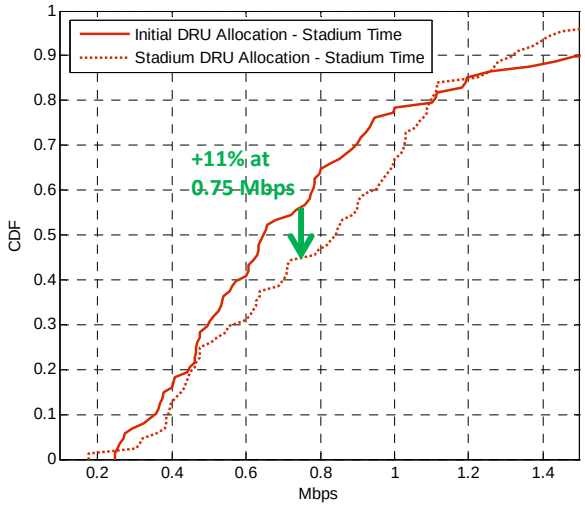


Figure 4.23: CDF of users' throughput for Stadium time

Chapter 5

Self-Organized Intelligent Distributed Antenna System for Geographic Load Balancing

One approach to solve traffic hot spots and unbalanced traffic distributions is to take advantage of the load balancing capability of an IDAS. A Virtual Cell network is an IDAS with capacity routing capability. A Virtual Cell network can assign capacity on demand and improve the link budget of the user. In a Virtual Cell network, the available resources from multiple eNBs, at a base station hotel (BTS hotel), are distributed to multiple remote antenna modules. With respect to a Virtual Cell configuration, multiple remote antenna modules can be dynamically assigned to the different eNB resources depending on the time-varying traffic in order to resolve unbalanced traffic scenarios. Once traffic resources are aggregated into an eNB hotel, the discrete resources of a single eNB are still allocated to a specific set of antennas associated with that eNB and provide coverage to a geographic area. The traffic resources are fixed, i.e., only the resources associated with a specific eNB can be allocated to the antennas modules associated with that eNB. However, because the eNBs are collocated in an eNB hotel, the system can use the aggregated traffic resources of the discrete eNBs as a single pooled traffic resource that can be allocated according to an algorithm. Provisioning assumptions are typically predicated on worst-case traffic assets in all areas. Network design is wasteful at large percentage of the time, inevitably resulting in over or under-provisioning of the fixed resources. Traffic resources either go unused or are under-provisioned and are insufficient to handle the offered traffic. Both circumstances give rise to the same outcome: lost revenue and lost opportunity. When a site's traffic resources are idle and unused, the traffic asset fails to provide an optimal return on investment. But a site that lacks sufficient capacity to support the offered traffic at any point during the day garners dropped calls, lost revenues, and dissatisfied customers.

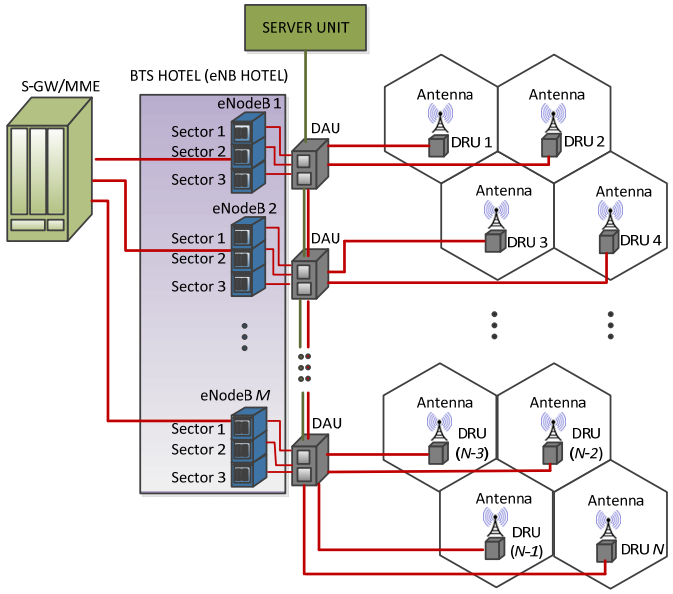


Figure 5.1: Structure of a Virtual Cell Network.

The capability of dynamic load balancing using distributed antenna was introduced by [41] and [42], but they did not propose any methods and algorithms to perform load balancing. Reference [43] investigated a load balancing scheme by contracting the antenna pattern around the traffic hot spot and expanding adjacent cells coverage.

In order to attain load balancing using a Virtual Cell network, a SON algorithm is proposed. Geographic load balancing using SON is recognized as a new approach for traffic load balancing which provides dynamic load redistribution in real time according to the current geographic traffic conditions. It can be used to improve the performance for any distributed systems containing non-uniformly distributed traffic, especially for resolving traffic hot spots.

We examine delivering the eNB traffic to independent antenna modules with the objective to dynamically balance the traffic.

The main contribution of this work is to introduce IDAS as SON network which has this ability to distribute the resources (or capacity) over a given geographic area depending on time-varying traffic, which requires solving an optimization problem. Whereas in conventional base station without a complex scheduler, it is impossible to distribute dynamically resources over a given area. evolutionary algorithms (EAs) such as GA and EDA are proposed in this paper to solve optimization load balancing problem.

5.1 System Model.

In a Virtual Cell architecture as shown in Fig. 5.1, DRUs are connected to an eNB hotel (BTS Hotel) via optical fiber and one or more DAU. The DAUs are interconnected and connected to multiple eNBs and a server unit (SU). Note that, each eNB contains multiple sectors such that each sector utilizes entire resources available. This capability enables the virtualization of the eNB/sector resources at the independent DRUs. The eNB hotel is linked to a service gate-way (S-GW) or a mobile management entity (MME). DRUs are allocated such that each DRU allocated to a given eNB is simulcast. For the simulcasting operation, the access network between the eNB hotel and DRUs should have a multidrop bus topology. The DAUs dynamically assign the radio resources of the various eNBs to the independent DRUs according to the server commands. Server commands determine the proper DRU configuration and how to share eNB resources between independent DRUs. The traffic information at the DRUs will be used at SU in order to generate server commands for dynamically allocating the traffic resources to the required geographical areas [44].

5.2 Dynamic DRU Allocation and Formulation.

In the Virtual Cell network, it is important to allocate the DRUs to cope with dynamically changing traffic and to balance the traffic for each sector.

At the DRU coverage area the traffic is increased or decreased depending on the period of time. Thus it is necessary to dynamically allocate the DRUs such that the allocated DRUs in a sector satisfy the network limitations.

The network's performance determines the QoS values which can be expressed by only one key performance indicator (KPI) or two KPIs or some KPIs. The SU is responsible to find proper DRU allocation for IDAS network using the collected data from DRUs without accessing to the eNBs' KPIs. Each vendor indicates different KPIs to evaluate network performance. i.e. Call blocking, hand-offs. QoS metrics could be defined and mapped to a set of KPIs. When a set of KPIs is used, then the mapping needs to be represented by a weighted normalized function [45]. Given the DRU allocation configuration at time period t, our problem is to obtain new DRU allocation at time period $t+1$ that adaptively balances the change in traffic demands.

Further, we will propose SON algorithms that are based on a simple and decentralized algorithm which is performed at SU. In the SON algorithms, the expected network gain, which is comprised of system KPIs, is used in order to perform DRU allocation.

5.3 Metric Definition and Formulation.

In order to formulate load balancing problems, two different "Utility Based Load Balancing" and "Block Probability-triggered Load Balancing" schemes are investigated. The $load$ is defined as a function of available resources in "Utility Based Load Balancing" scheme and as a function of number of users in "BlockProbability-triggered Load Balancing" scheme.

5.3.1 Block Probability-triggered Load Balancing Scheme.

Unlike traditional cellular architecture where all KPIs are evaluated at eNB by performing user by user base band processing, defining the new KPIs based on collected information from DRUs is necessary in Virtual Cell architecture. Note that, no baseband processing is performed over DRUs in IDAS architecture.

The load per eNB works in tandem with the scheduler and admission control. The admission control, which defines the maximum number of users per cell, is set by the operator. Hard Capacity is defined as the maximum number of simultaneous active users, which are connected to eNB's sector, i.e. Femto-cell is approximately 4-5 users, Pico-cell is approximately 16-30 users, Micro-cell is 60-200 users and Macro-cell is 500-700 users. This is either a hardware limitation or a software licensing limitation. In order to formulate the problem, we consider a service coverage area with *N* DRUs. Assume that an eNB hotel has *M* eNBs and *K* sectors. Let SOS_m and SOD_k be the set of sectors in eNB_m and the set of DRUs allocated to $Sector_k$, respectively, such that $|SOS_m| = 3$ (if each eNB has three sectors). The main binary variable x_{ik}^t is 1, when DRU_i belongs to $Sector_k$ at the period *t*. Therefore, given the allocation of DRUs at time period *t* ($\mathbf{X}_k^t = [x_{1k}^t, x_{2k}^t, ..., x_{Nk}^t], k = 1, ..., K$), finding new allocation at time period *t+1* ($\mathbf{X}_k^{t+1} = [x_{1k}^{t+1}, x_{2k}^{t+1}, ..., x_{Nk}^{t+1}], k = 1, ..., K$) is our main problem. Each DRU is assumed to have active users T_i ($i = 1, ..., N$) at time period *t* and *t+1*. Note that, UE_A(User Equipment) belongs to DRU_B if the received uplink power from UE_A at DRU_B is greater than the other DRUs. The transition probability of UEs from DRU_i to DRU_j is assumed p_{ij}. Then, the hand-offs from DRU_i to DRU_j becomes $h_{ij} = p_{ij}T_i$. The real time estimation for p_{ij} is not the primary issue in this thesis. However, many papers use terminal mobility modeling to determine p_{ij}. The p_{ij} models introduced by other authors can also be used in our optimization problem. Nevertheless, by considering that users are distributed uniformly inside each antenna module area (DRU area), we can claim that p_{ij} can be modeled as being inversely proportional to the distance between DRU_i and DRU_j where the distance between two different DRUs is fixed. Note that, a uniform distribution of users inside each DRU with a different number of users per DRU will produce a non-uniform distribution of users around the entire coverage area.

Unlike traditional cellular architecture where all KPIs are evaluated at eNB by performing user by user base band processing, defining the new KPIs based on collected information from DRUs is necessary in Virtual Cell architecture. In our "Block Probability-triggered Load Balancing" scheme,

we define the KPIs which reqiure no user by user base band processing over DRUs in IDAS architecture .

The following three KPIs are considered in the DRU allocation problem. Fig. 5.2 indicates an example allocation at time t and $t+1$ where the 6 DRUs are fed by two eNBs. Note that eNB_2 has two separate sectors ($Sector_2$ and $Sector_3$) and eNB_1 has only one sector ($Sector_1$). All 3 KPIs are calculated for both DRU allocations at time t and $t+1$.

1. $\mathbf{KPI_{BC}}$ ***(Inverse of the number of Blocked Calls):*** The metric for the blocked calls caused by hard capacity. Let HC_k be the hard capacity of $Sector_k$. The blocked calls occur in $Sector_k$, only when the number of $Sector_k$'s active users is greater than the hard capacity of $Sector_k$ (HC_k). Since the number of $Sector_k$'s active users is $\sum_{i=1}^{N} T_i x_{ik}^{t+1}$, therefore KPI_{BC} is,

$$KPI_{BC} = \left[\sum_{k=1}^{K} BC_k^{t+1}\right]^{-1} \tag{5.1}$$

where the $Sector_k$'s blocked call number BC_k^{t+1} in period *t+1* is $max\left\{0, \left[\sum_{i=1}^{N} T_i x_{ik}^{t+1}\right] - HC_k\right\}$. In the case where there are no blocked calls in the system ($\sum_{k=1}^{K} BC_k^{t+1} = 0$), the KPI_{BC} is assumed to be equal to 1.1 in order to avoid calculating $\frac{1}{0}$.

In the Fig. 5.2 example, by assuming a value of 20 for the hard capacity in the each sector ($HC_k = 20, k = 1, ..., K$)(Pico-cell is assumed), 11 blocked calls (KPI_{BC}=0.09) are produced at time t, where the other two sectors have idle resources. However at $t+1$, the traffic is well balanced and only one call is blocked (KPI_{BC}=1).

2. $\mathbf{KPI_{HO}}$ ***(Inverse of the number of Hand-offs):*** Three different types of Hand-offs are:

A. Inter-eNodeB Hand-off: The UE with an ongoing call needs an inter-eNodeB hand-off when it moves from one eNB to another eNB. As long as the UE does not leave the LTE coverage area, inter-eNodeB hand-off is performed using the X2 interface (eNB uses the X2 interface for making connection with other eNB). Otherwise, when the UE leaves the serving cell inter-eNodeB hand-off is performed using the S1 interface (eNB uses the S1 interface for making connection with the MME and S-GW). During X2 hand-off, the new geographic position of the UE is updated at the MME by the target eNB. To enable this procedure, the communication between MME and S-GW is needed [51]. During the S1 hand-off, the relocation preparation request of the source eNB is received by the MME, and then MME starts requesting the necessary radio resource from the target eNB. The MME triggers the main procedure and S-GW executes it [51].

Let $y_{im}^{t+1} = \sum_{k \in SOS_m} x_{ik}^{t+1}$ then $y_{im}^{t+1} = 1$ when DRU_i belongs to eNB_m at time period *t+1*. The binary variable z_{ijm}^{t+1} is 1, when DRU_i and DRU_j belong to eNB_m at period time *t+1* ($y_{im}^{t+1} = y_{jm}^{t+1} = 1$). The inter-eNodeB hand-off cost variable is computed by using the variable $z_{ij}^{t+1} = 1 - \sum_{m} z_{ijm}^{t+1}$. $z_{ij}^{t+1} = 1$ when DRU_i and DRU_j belong to different eNBs at period time *t+1*. Therefore, the cost

variable becomes $\sum_i \sum_{j \neq i} h_{ij} z_{ij}^{t+1}$ where $h_{ij} = p_{ij} T_i$ is the hand-off from DRU_i to DRU_j and p_{ij} can be modeled as being inversely proportional to the distance between DRU_i and DRU_j. In the Fig. 5.2 example, by assuming $p_{ij} = (1/D_{ij})$, the inter-eNB hand-off is calculated at t and $t+1$ where D_{ij} is the distance between DRU_i and DRU_j.

Note that $z_{ij}^{t+1} = f_1(z_{ijm}^{t+1})$, $z_{ijm}^{t+1} = f_2(y_{im}^{t+1}, y_{jm}^{t+1})$, $y_{im}^{t+1} = f_3(x_{ik}^{t+1})$ and $y_{jm}^{t+1} = f_4(x_{ik}^{t+1})$. Therefore, the variable z_{ij}^{t+1} is a function of the main variables $x_{ik}^{t+1}, i = 1,...,N (z_{ij}^{t+1} = f_1(f_2(f_3(x_{ik}^{t+1}), f_4(x_{ik}^{t+1}))))$.

B. Intra-eNodeB Hand-off: The UE with an ongoing call needs an intra-eNodeB hand-off when it moves from one sector to another sector in an eNB. This procedure does not need to involve the MME or S-GW because it can be handled entirely within that eNB. The binary variable w_{ijk}^{t+1} is 1 when DRU_i and DRU_j belong to $Sector_k$ at time period *t+1* ($x_{ik}^{t+1} = x_{jk}^{t+1} = 1$). Then the intra-eNodeB hand-off cost is computed by using two variables w_{ij}^{t+1} and z_{ij}^{t+1} where $w_{ij}^{t+1} = 1 - \sum_k w_{ijk}^{t+1}$. w_{ij}^{t+1}=1 when DRU_i and DRU_j belong to different sector at time period *t+1*. By considering to previous part that $z_{ij} = 0$ when DRU_i and DRU_j belong to same eNB, therefore, the intra-eNB hand-off cost becomes $\sum_i \sum_{j \neq i} h_{ij}(w_{ij}^{t+1} - z_{ij}^{t+1})$. Note that $w_{ij}^{t+1} = f_5(w_{ijk}^{t+1})$, $w_{ijk}^{t+1} = f_6(x_{ik}^{t+1}, x_{jk}^{t+1})$. Therefore, the variable w_{ij}^{t+1} is a function of the main variables $x_{ik}^{t+1}, i = 1,...,N (w_{ij}^{t+1} = f_5(f_6(x_{ik}^{t+1}, x_{jk}^{t+1})))$.

An intra-eNodeB hand-off occurs when DRU_i and DRU_j belong to different sectors of the same eNB. In the Fig. 5.2 example, by assuming $p_{ij} = (1/D_{ij})$, the intra-eNB is calculated for DRU allocations at time t and $t+1$ where D_{ij} is distance between DRU_i and DRU_j.

C. Forced Hand-off: By changing, DRU's sector, all ongoing calls in that DRU have to change their sector. Dropped calls will not be happen during the allocation procedure in LTE system owing to soft hand-off technique utilization. In soft hand-off technique, while the connection with source eNB sector is maintained, a new wireless link connection is established with the target eNB sector [56]. Lets assume the binary variable $a_{ik} = 1$ when $x_{ik}^t = 0$ and $x_{ik}^{t+1} = 1$. A forced hand-off occurs when DRU_i currently in another sector at time period *t* moves into $Sector_k$ at time period *t+1*, then the forced hand-off cost becomes $\sum_i \sum_k a_{ik} T_i$.

Note that the variable a_{ik} is a function of the main variables x_{ik}^t and x_{ik}^{t+1}, $i = 1, ..., N$ ($a_{ik} = f_7(x_{ik}^t, x_{ik}^{t+1})$). In the Fig. 5.2 example, the number of forced hand-offs is calculated at t and $t+1$. In DRU allocation at time $t+1$, DRU_1's users (5 users) and DRU_2's users (15 users) change their sectors, so 20 forced hand-offs occur.

The weighted combination of these three hand-off cost variables would be:

$$\begin{aligned} PI_{HO} &= [r_1 \sum_i \sum_{j \neq i} h_{ij} z_{ij}^{t+1} + r_2 \sum_i \sum_{j \neq i} h_{ij}(w_{ij}^{t+1} - z_{ij}^{t+1}) + r_3 \sum_i \sum_k a_{ik} T_i]^{-1} \\ &= [r_1 \sum_i \sum_{j \neq i} h_{ij} f_1(f_2(f_3(x_{ik}^{t+1}), f_4(x_{ik}^{t+1}))) + r_2 \sum_i \sum_{j \neq i} [h_{ij}(f_5(f_6(x_{ik}^{t+1}, x_{jk}^{t+1})) \\ &\quad - f_1(f_2(f_3(x_{ik}^{t+1}), f_4(x_{ik}^{t+1}))))] + r_3 \sum_i \sum_k f_7(x_{ik}^t, x_{ik}^{t+1}) T_i]^{-1} \end{aligned} \tag{5.2}$$

where r_1, r_2 and r_3 indicate the priority level of the described hand-off types. In the Fig. 5.2

example, the KPI_{HO} is calculated for DRU allocations at t and $t+1$ where r_1=2 and r_2=r_3=1.

3. $\mathbf{KPI_{CI}}$ ***(Inverse of the Compactness Index):*** Note that all DRUs allocated to a given sector broadcast the radio signals in simulcast. A disconnected DRU allocation pattern generates unnecessary hand-offs between sectors. Therefore, the DRUs allocated to a given sector need to be connected. The compactness of sector allocation is considered in order to minimize the hand-offs and interference among sectors. The DRUs of compact sector have fewer common edges with the neighbor sectors than the DRUs of an incompact sector. In the Fig. 5.2 example, isolated DRU_1 is surrounded by DRUs allocated to other sectors, higher interference is experienced by DRU_1 than the other DRUs in a compact allocation pattern. Compact allocation also reduces the number of hand-offs by decreasing the length of the border between two different eNBs. Therefore, we define the ratio of the number of hand-off cell sides to the total number of cell sides as the compactness index (*CI*) factor in the network. In equation (5.3), number of hand-off DRU sides between two different sectors is represented by the numerator term.

$$KPI_{CI} = [CI]^{-1} = \left[\frac{\sum_i \sum_{j\neq i} B_{ij} w_{ij}^{t+1}}{\sum_i \sum_{j\neq i} B_{ij}}\right]^{-1} = \left[\frac{\sum_i \sum_{j\neq i} B_{ij} f_5(f_6(x_{ik}^{t+1}, x_{jk}^{t+1}))}{\sum_i \sum_{j\neq i} B_{ij}}\right]^{-1} \tag{5.3}$$

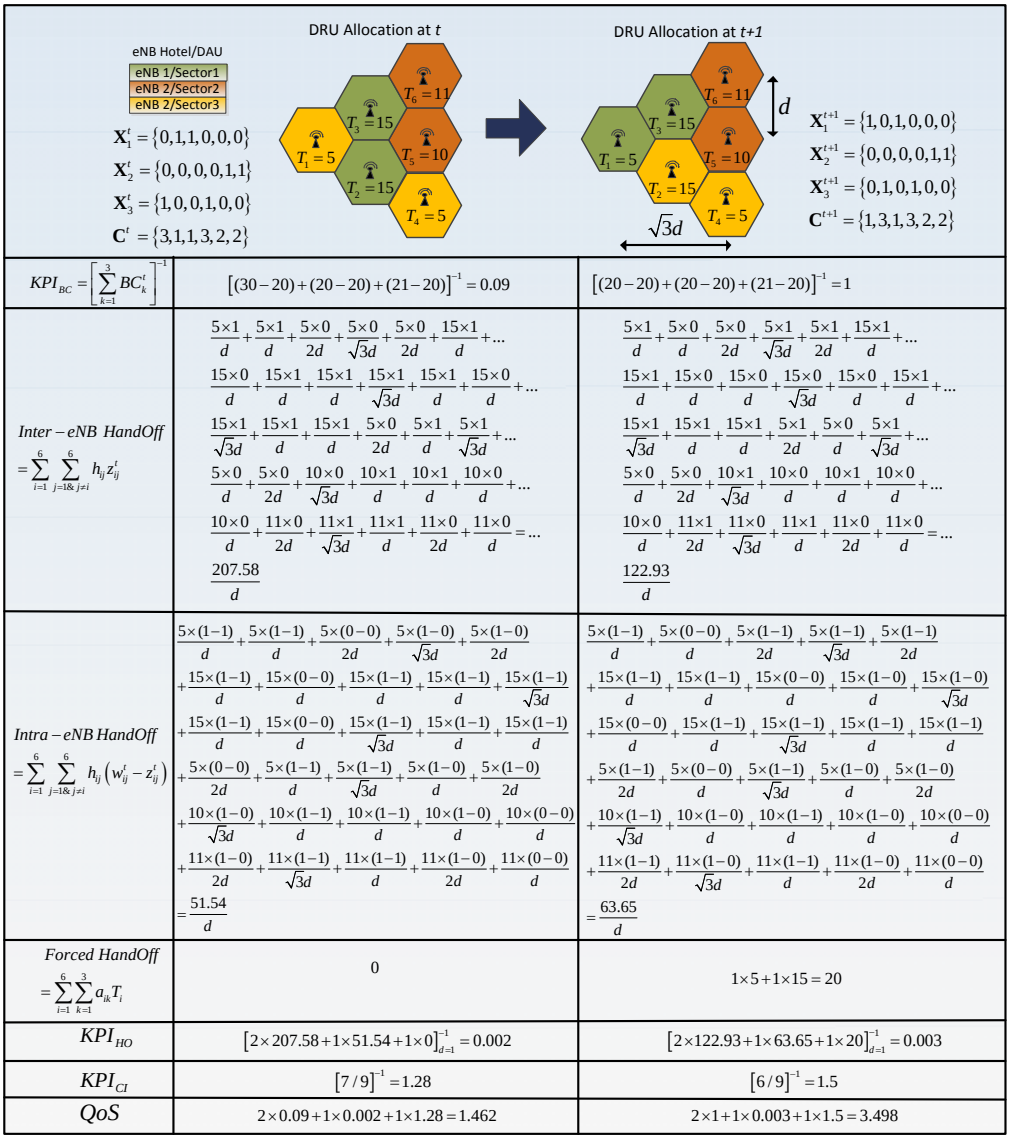
where $B_{ij} = 1$, if DRU_i and DRU_j are adjacent. In the Fig. 5.2. example, the *CI* index and KPI_{CI} are calculated at time t and $t+1$.

Note that, the formulation constraints are: 1. Each DRU has to belong to a sector at time period *t+1* ($\sum_k x_{ik}^{t+1} = 1$ for all i). 2. The sector's DRUs have to be connected if that sector has more than one DRU. In order to formulate connected sectors the cut theorem is employed [46] on SOD_k. If $Sector_k$ is connected, then at least one common side of the hexagonal cells exists in any cut that separates cells in SOD_k. Let $S1_k$ be a proper subset of SOD_k, so, $S1_k \subset SOD_k$,$S1_k \neq \phi$, and $S1_k \neq SOD_k$. Also let $S2_k$ be the complement set of $S1_k$, so, $S2_k = SOD_k - S1_k$. There exists at least one common side of the DRUs separated by the subsets, this is because two subsets are connected. Thus $\sum_{i\in S1_k} \sum_{j\in S2_k} B_{ij} \geq 1$.

Now, our QoS function is obtained by the weighted combination of the three KPIs and our objective function is to maximize the QoS function.

$$\begin{aligned} &\underset{\mathbf{X}_k^T}{Maximize} \quad QoS = q_1 KPI_{BC} + q_2 KPI_{HO} + q_3 KPI_{CI} \\ &\text{Subjent to:} \\ &\sum_k x_{ik}^{t+1} = 1 \qquad \text{for all } \ i \end{aligned} \tag{5.4}$$

where q_1, q_2 and q_3 indicate the priority level of described KPIs. Note that, all binary variables y_{im}^{t+1}, z_{ij}^{t+1}, w_{ij}^{t+1} and a_{ik} ($i,j \in \{1,...,N\}, k \in \{1,...,K\}, m \in \{1,...,M\}$) which are previously defined, are functions of the main binary variable x_{ik}^{t+1}. Since our problem is to obtain a new DRU allocation at time period *t+1* ($\mathbf{X}_k^{t+1}$=$\left\{x_{1k}^{t+1}, x_{2k}^{t+1}, ..., x_{Nk}^{t+1}\right\}, k = 1,...,K$) by varying the main binary variable



$KPI_{BC} = \left[\sum_{k=1}^{3} BC_k^t\right]^{-1}$	$\left[(30-20)+(20-20)+(21-20)\right]^{-1} = 0.09$	$\left[(20-20)+(20-20)+(21-20)\right]^{-1} = 1$
$Inter-eNB\ HandOff = \sum_{i=1}^{6}\sum_{j=1\&\, j\neq i}^{6} h_{ij} z_{ij}^t$	$\frac{5\times1}{d}+\frac{5\times1}{d}+\frac{5\times0}{2d}+\frac{5\times0}{\sqrt{3}d}+\frac{5\times0}{2d}+\frac{15\times1}{d}+\ldots$ $\frac{15\times0}{d}+\frac{15\times1}{d}+\frac{15\times1}{d}+\frac{15\times1}{\sqrt{3}d}+\frac{15\times1}{d}+\frac{15\times0}{d}+\ldots$ $\frac{15\times1}{\sqrt{3}d}+\frac{15\times1}{d}+\frac{15\times1}{d}+\frac{5\times0}{2d}+\frac{5\times1}{d}+\frac{5\times1}{\sqrt{3}d}+\ldots$ $\frac{5\times0}{d}+\frac{5\times0}{2d}+\frac{10\times0}{\sqrt{3}d}+\frac{10\times1}{d}+\frac{10\times1}{d}+\frac{10\times0}{d}+\ldots$ $\frac{10\times0}{d}+\frac{11\times0}{2d}+\frac{11\times1}{\sqrt{3}d}+\frac{11\times1}{d}+\frac{11\times0}{2d}+\frac{11\times0}{d}=\ldots$ $\frac{207.58}{d}$	$\frac{5\times1}{d}+\frac{5\times0}{d}+\frac{5\times0}{2d}+\frac{5\times1}{\sqrt{3}d}+\frac{5\times1}{2d}+\frac{15\times1}{d}+\ldots$ $\frac{15\times1}{d}+\frac{15\times0}{d}+\frac{15\times0}{d}+\frac{15\times0}{\sqrt{3}d}+\frac{15\times0}{d}+\frac{15\times1}{d}+\ldots$ $\frac{15\times1}{\sqrt{3}d}+\frac{15\times1}{d}+\frac{15\times1}{d}+\frac{5\times1}{2d}+\frac{5\times0}{d}+\frac{5\times1}{\sqrt{3}d}+\ldots$ $\frac{5\times0}{d}+\frac{5\times0}{2d}+\frac{10\times1}{\sqrt{3}d}+\frac{10\times0}{d}+\frac{10\times1}{d}+\frac{10\times0}{d}+\ldots$ $\frac{10\times0}{d}+\frac{11\times1}{2d}+\frac{11\times0}{\sqrt{3}d}+\frac{11\times1}{d}+\frac{11\times0}{2d}+\frac{11\times0}{d}=\ldots$ $\frac{122.93}{d}$
$Intra-eNB\ HandOff = \sum_{i=1}^{6}\sum_{j=1\&\, j\neq i}^{6} h_{ij}\left(w_{ij}^t - z_{ij}^t\right)$	$\frac{5\times(1-1)}{d}+\frac{5\times(1-1)}{d}+\frac{5\times(0-0)}{2d}+\frac{5\times(1-0)}{\sqrt{3}d}+\frac{5\times(1-0)}{2d}$ $+\frac{15\times(1-1)}{d}+\frac{15\times(0-0)}{d}+\frac{15\times(1-1)}{d}+\frac{15\times(1-1)}{d}+\frac{15\times(1-1)}{\sqrt{3}d}$ $+\frac{15\times(1-1)}{d}+\frac{15\times(0-0)}{d}+\frac{15\times(1-1)}{\sqrt{3}d}+\frac{15\times(1-1)}{d}+\frac{15\times(1-1)}{d}$ $+\frac{5\times(0-0)}{2d}+\frac{5\times(1-1)}{d}+\frac{5\times(1-1)}{\sqrt{3}d}+\frac{5\times(1-0)}{d}+\frac{5\times(1-0)}{2d}$ $+\frac{10\times(1-0)}{\sqrt{3}d}+\frac{10\times(1-1)}{d}+\frac{10\times(1-1)}{d}+\frac{10\times(1-0)}{d}+\frac{10\times(0-0)}{d}$ $+\frac{11\times(1-0)}{2d}+\frac{11\times(1-1)}{\sqrt{3}d}+\frac{11\times(1-1)}{d}+\frac{11\times(1-0)}{2d}+\frac{11\times(0-0)}{d}$ $=\frac{51.54}{d}$	$\frac{5\times(1-1)}{d}+\frac{5\times(0-0)}{d}+\frac{5\times(1-1)}{2d}+\frac{5\times(1-1)}{\sqrt{3}d}+\frac{5\times(1-1)}{2d}$ $+\frac{15\times(1-1)}{d}+\frac{15\times(1-1)}{d}+\frac{15\times(0-0)}{d}+\frac{15\times(1-0)}{d}+\frac{15\times(1-0)}{\sqrt{3}d}$ $+\frac{15\times(0-0)}{d}+\frac{15\times(1-1)}{d}+\frac{15\times(1-1)}{\sqrt{3}d}+\frac{15\times(1-1)}{d}+\frac{15\times(1-1)}{d}$ $+\frac{5\times(1-1)}{2d}+\frac{5\times(0-0)}{d}+\frac{5\times(1-1)}{\sqrt{3}d}+\frac{5\times(1-0)}{d}+\frac{5\times(1-0)}{2d}$ $+\frac{10\times(1-1)}{\sqrt{3}d}+\frac{10\times(1-0)}{d}+\frac{10\times(1-1)}{d}+\frac{10\times(1-0)}{d}+\frac{10\times(0-0)}{d}$ $+\frac{11\times(1-1)}{2d}+\frac{11\times(1-0)}{\sqrt{3}d}+\frac{11\times(1-1)}{d}+\frac{11\times(1-0)}{2d}+\frac{11\times(0-0)}{d}$ $=\frac{63.65}{d}$
$Forced\ HandOff = \sum_{i=1}^{6}\sum_{k=1}^{3} a_{ik} T_i$	0	$1\times5+1\times15=20$
KPI_{HO}	$\left[2\times207.58+1\times51.54+1\times0\right]_{d=1}^{-1} = 0.002$	$\left[2\times122.93+1\times63.65+1\times20\right]_{d=1}^{-1} = 0.003$
KPI_{CI}	$\left[7/9\right]^{-1} = 1.28$	$\left[6/9\right]^{-1} = 1.5$
QoS	$2\times0.09+1\times0.002+1\times1.28=1.462$	$2\times1+1\times0.003+1\times1.5=3.498$

Figure 5.2: Example of DRU allocation at time t and $t+1$.

x_{ik}^{t+1}, the search space size is 2^{NK}. Since each DRU has to belong to only one sector (based on the constrained $\sum_k x_{ik}^{t+1} = 1$), we convert the constrained to an unconstrained problem and decrease the search space size from 2^{NK} to K^N by introducing a new district variable $c_i^{t+1} \in \{1, ..., K\}, i = 1, ..., N$. So the new DRU allocation at period time *t+1* is identified by C^{t+1}=$\{c_1^{t+1}, c_2^{t+1}, ..., c_N^{t+1}\}$, where only one sector can be assigned to the district variable c_i^{t+1}. Note that, c_i^{t+1}=k represents the DRU_i belongs to $Sector_k$ (x_{ik}^{t+1}=1). Therefore the new unconstrained problem is,

$$\underset{C^T}{Maximize} \quad QoS = q_1 KPI_{BC} + q_2 KPI_{HO} + q_3 KPI_{CI} \tag{5.5}$$

In the Fig. 5.2. example, the QoS is calculated for both DRU allocations at t and $t+1$ where q_1=2, q_2=q_3=1. Both DRU allocation vector variables $\mathbf{X}_k^T$ and C^T are represented in the Fig. 5.2 at time $T = t$ and $T = t+1$.

The optimal solution will be found by exhaustively searching all the search space size. Since the possible number of solutions increases exponentially with the number of sectors and DRUs, the time it takes to execute the algorithm is exponentially increasing. Therefore, we investigate evolutionary algorithms (Appendix B) to solve the optimization DRU allocation problem.

5.3.2 Utility Based Load Balancing scheme.

For this load balancing scheme, the load per eNB is defined on maximum available resources.

In order to formulate the problem, we consider a service coverage area with *M* DRUs and assume that an eNB hotel has *N* eNBs. Since the downlink path of a IDAS can be considered as an equivalent MISO system with additive interference and noise, the received signal of the user k can be expressed as,

$$\begin{aligned} y_k(\mathbf{S}) &= \text{desired signal} + \text{interference signal} + \text{noise} \\ &= \sum_{i=1}^{M} s_k^{(i)} h_k^{(i)} x_k^{(i)} + \sum_{j=1}^{M} (1 - s_k^{(j)}) h_k^{(j)} x_k^{(j)} + \text{noise} \end{aligned} \tag{5.6}$$

where $h_k^{(i)}$ and $x_k^{(i)}$ denote the channel between DRU i to user k and the transmitted signal of DRU i to user k, respectively. The distribution of the additive white noise is $CN(0, \sigma_{noise}^2)$. We define an assignment indicator $\mathbf{S} = \{s_k^{(i)} | k = 1, ..., K \, and \, i = 1, ..., M\}$, where $s_k^{(i)}$ equals to 1 when user k receives the desired signal from i-th DRU and equals to 0 when user k receives an interference signal from i-th DRU. Note that $h_k^{(i)} = h_{k,w}^{(i)} l_k^{(i)}$, where the composite fading channel $h_k^{(i)}$, encompasses not only small-scale fading ($h_{k,w}^{(i)}$) but also large-scale fading ($l_k^{(i)}$).

Then the average spectral efficiency (ASE) (bit/Hz) of user k for MISO vector channel can be obtained in a simple form by (Appendix A.3)

$$\vartheta_k(\mathbf{S}) = -\frac{1}{\ln 2}\sum_{i=1}^{M} s_k^{(i)}\pi_k^{(i)} \exp\left(-\frac{\sigma_k^2}{\left[l_k^{(i)}\right]^2 P^{(i)}}\right) Ei\left(-\frac{\sigma_k^2}{\left[l_k^{(i)}\right]^2 P^{(i)}}\right) \tag{5.7}$$

where, $E_i(t)$ is the exponential integral function ($Ei(t) = -\int_{-x}^{\infty} e^{-t}/t\, dt$) and can be easily calculated with popular numerical tools such as MATLAB.

Since the resulting user ASE is in fact the achievable throughput when multiplied by the assigned bandwidth to that user, R_k number of resource block (RB) is required to assign to user k in order to achieve a requirement throughput ϕ_k.

$$R_k(\mathbf{S}) = \frac{\phi_k}{R_{BW}\vartheta_k(\mathbf{S})} \tag{5.8}$$

where R_{BW} is the bandwidth of RB.

In this section, the DRU allocation problem is formulated as integer programming to balance the load amongst the eNBs and to minimize average load of the network as well as minimizing the number of handovers.

1. Load Balancing Index and Average Load of the Network: We assume that η_n represents the load of eNB n, which is,

$$\eta_n(\mathbf{I},\mathbf{S}) = \frac{\sum_{k=1}^{K} I_{n,k}R_k(\mathbf{S})}{R} \tag{5.9}$$

where, R is the total number RBs for each eNB and $\mathbf{I} = \{I_{n,k}\,|n = 1, ..., N\, and\, k = 1, ..., K\}$ is an assignment indicator. $I_{n,k}$ equals 1 when user k is served by eNB n. We assume all eNBs use the same bandwidth (same number of RB). The constraint $\sum_{n=1}^{N} I_{n,k} = 1, \forall k \in \{1, 2, ..., K\}$ requires that user k can be served by only one eNB. And the constraint $\sum_{k=1}^{K} I_{n,k}R_k(\mathbf{S}) \leq R, \forall n \in \{1, 2, ..., N\}$requires that the RBs occupied by all users in eNB n could not exceed the existing RBs of that eNB. We use $block_n$ to represent the blocking rate of eNB n which is,

$$block_n(\mathbf{I},\mathbf{S}) = \max\left[\eta_n(\mathbf{I},\mathbf{S}) - 1, 0\right]$$

(5.10)

Therefore the blocking rate of the network is $100 \times \sum_{n=1}^{N} block_n(\mathbf{I},\mathbf{S})$. The level of load balancing is evaluated using Jains fairness index [57] as follows,

$$\psi(\mathbf{I},\mathbf{S}) = \frac{\left[\sum_{n=1}^{N}\eta_n(\mathbf{I},\mathbf{S})\right]^2}{N\sum_{n=1}^{N}\eta_n^2(\mathbf{I},\mathbf{S})} \tag{5.11}$$

The value of the load balance index ψ is in interval $[1/N, 1]$. A larger ψ represents a more balanced load distribution amongst the eNBs i.e., $\psi = 1$ denotes that all eNBs have equal load. Therefore, as a first objective, we try to maximize ψ in order to achieve a balanced load. We also use η to represent the average load of the network, which is,

$$\eta(\mathbf{I}, \mathbf{S}) = \frac{1}{N} \sum_{n=1}^{N} \eta_n(\mathbf{I}, \mathbf{S}) \tag{5.12}$$

As a second objective, we minimize the average load of the network in order to prevent performing handovers of improper users with poor channel conditions.

2. Hand-off Index: Allocating DRU to new eNB causes enforced handover. Since DRU handovers of eNBs involves not only those eNBs, but also the MME and S-GW, all these require a significant amount of signal messaging. Therefore, it is important to consider the number of handovers as a third objective for load balancing problems. We use $h(\mathbf{I})$ to represent the average number of handovers required to be performed for an assignment indicator $\mathbf{I}$,

$$h(\mathbf{I}) = \frac{\sum_{n=1}^{N} \sum_{k=1}^{K} (I_{n,k} - I^0_{n,k})}{2K} \tag{5.13}$$

where $\mathbf{I}^0 = \{I^0_{n,k} \,|n = 1, ..., N \, and \, k = 1, ..., K\}$ is an assignment indicator for previous DRU allocation configuration. $I^0_{n,k}$ equals 1 when user k is served by eNB n at previous DRU allocation configuration. Then, we try to minimize $h(\mathbf{I})$ as an objective of the problem,

Now, our final objective function is a weighted combination of the three objectives which we have introduced. We try to maximize this final objective function F.

$$\begin{aligned} \underset{I,S}{max} \quad & F(\mathbf{I}, \mathbf{S}) = \alpha\psi(\mathbf{I}, \mathbf{S}) \, - \beta\eta(\mathbf{I}, \mathbf{S}) - (1 - \alpha - \beta) \, h(\mathbf{I}) \\ s.t. \quad & \sum_{k=1}^{K} I_{n,k} R_k(\mathbf{S}) \leq R, \, \forall n \in \{1, 2, ..., N\} \\ & \sum_{n=1}^{N} I_{n,k} = 1, \, \forall k \in \{1, 2, ..., K\} \end{aligned} \tag{5.14}$$

where α and β are control parameters of the optimization function ($\alpha + \beta \leq 1$).

We define vector $\mathbf{b} = [b_1, b_2, ..., b_K]$ as the location indicator of the users where $b_k \in \{1, 2, ..., M\}$, b_k equals to m when the received uplink power of user k at DRU m is greater than the other DRUs. This implies that user k is located in the DRU m coverage area [44]. We also define vector $\mathbf{u} = u_1, u_2, ..., u_M$ as a DRU allocation indicator where $u_m \in \{1, 2, ..., N\}$, u_m equals n when DRU m is allocated to eNB n. Therefore, $u_{b_k} \in \{1, 2, ..., N\}$, $b_k \in \{1, 2, ..., M\}$, $k \in \{1, 2, ..., K\}$. Note that two matrix variables $\mathbf{S}$ and $\mathbf{I}$ are functions of two vector variables $\mathbf{u}$ and $\mathbf{b}$ in such a way that,

$$\underset{\substack{i \in \{1,2,...,M\} \\ k \in \{1,2,...,K\}}}{s^{(i)}_k(\mathbf{u}, \mathbf{b})} = \begin{cases} 1 & if \;\; u_{b_k} = u_i \\ 0 & otherwise \end{cases} \tag{5.15}$$

and

$$I_{\substack{n,k \\ n\in\{1,2,..,N\} \\ k\in\{1,2,..,K\}}}(\mathbf{u},\mathbf{b}) = \begin{cases} 1 & if \ u_{b_k} = n \\ 0 & otherwise \end{cases} \tag{5.16}$$

Since our problem is to obtain a new DRU allocation ($\mathbf{u}$) for a given user location indicator ($\mathbf{b}$), we define $\mathbf{u}$ as our main vector variable of the problem. As an example, when 3 eNBs are used to support 5 DRUs, $\mathbf{u} = \{3, 2, 1, 1, 3\}$ indicates that DRU 1, 2, 3, 4 and 5 are allocated to eNBs 3, 2, 1, 1 and 3, respectively and the number of possible solutions ($\mathbf{u}$, DRU allocation) is 3^5.

Each DRU allocation will affect the load balancing index and the network average load simultaneously. An exhaustive search (ES) is the only effective algorithm to find the optimal solution (optimal DRU allocation). The large number of possible solutions (DRU allocations) increases exponentially with the number of eNBs (N) and DRUs (M). The time it takes to execute the algorithm increases exponentially with the number of eNBs and DRUs, and it is not feasible to determine the optimum solution. The number of possible solutions for ES is $|N|^{|M|}$. Therefore, the use of an evolutionary algorithm (Appendix B) is proposed in the next section to solve the DRU allocation optimization problem.

5.4 Self-Optimizing Network of virtual Cells.

In this section, we propose a Self-Optimizing IDAS (SOIDAS) according to the above intuitive analysis. The SOIDAS algorithm uses the expected network gain, in order to obtain the proper DRU allocation. Fig. 5.3 represents the block diagram of the SOIDAS algorithm. The information collected at first step will be used to analyze QoS (equation 5.5 or 5.14) for the current DRU allocation configuration. This information includes the “number of users per DRU”[44], “DRUs distance from each other”, “current DRU allocation configuration”, “the channel state information (CSI) of all pairs of DRUs and users” [44] and “user location indicator ($\mathbf{b}$)” [44] .

Observing the block diagram in Fig. 5.3, the DRU allocation configuration is adjusted by comparing the optimized QoS value, calculated at the end of the optimization step, with the current QoS value. The SON algorithm seeks to maximize the objectives (e.g., KPIs) in order to produce a better QoS by using EAs. EAs are inspired by the theory of biological evolution. In EAs, candidate solutions to an optimization problem are represented as individuals in the population, and the cost function value of a candidate solution in the optimization problem indicates the fitness of the individual for the concept of natural selection [47]. Three different genetic algorithm (GA), estimation distributed algorithm (EDA) and particle swarm optimization (PSO) algorithms are used to optimize the objective functions. The details of theses algorithms are discussed at Appendix B.

The EAs terminate when a certain stopping criterion is reached, e.g. after a predefined number of generations. The fitness value of EAs is based on one of the introduced objective function

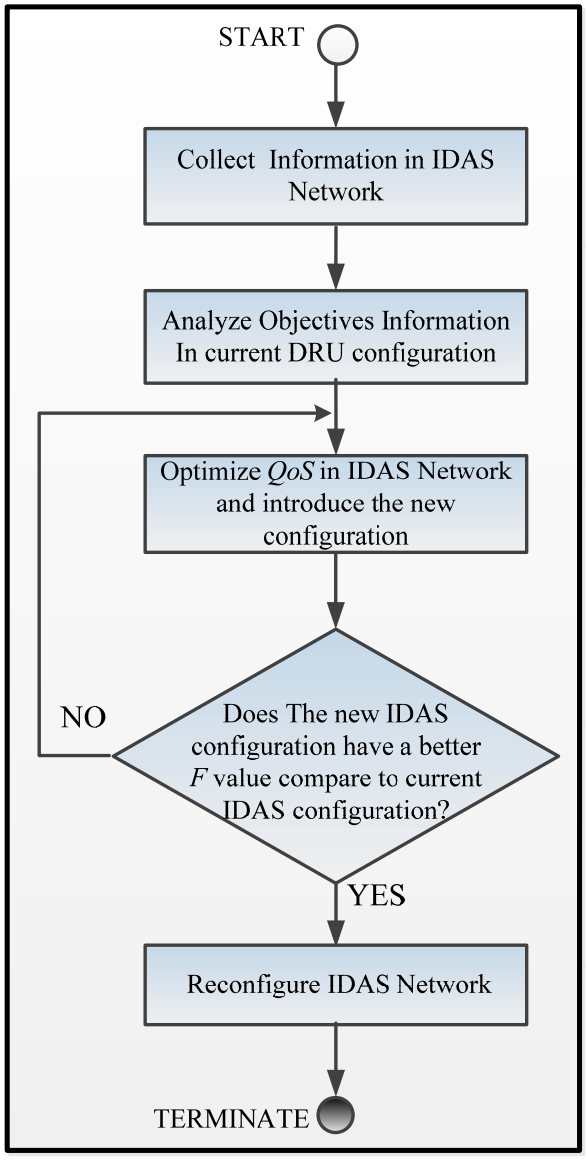


Figure 5.3: Block diagram of the SOIDAS algorithm.

equations (5.5) or (5.14).

5.5 Computational Results and Complexity Comparison.

The EAs are used for optimizing the QoS in SOIDAS algorithm. The EAs terminate when a certain stopping criterion is reached, e.g. after a predefined number of generations. The fitness value of EAs is based on one of the introduced objective function equations (5.5) (introduced for block probability-triggered load balancing scheme) or (5.14)(introduced for block probability-triggered load balancing scheme). We specify a main-vector C^t (introduced for block probability-triggered load balancing scheme) or $\mathbf{u}$ (utility based load balancing scheme) as a solution (string) for EAs.

Generating the Initial Population: Generally in both algorithms, the initial population is chosen randomly from the set $\{1, ..., K\}$ in "Block Probability-triggered Load Balancing" and the set $\{1, ..., N\}$ in "Utility Based Load Balancing" with uniform distribution. It should be noted that it is possible to get all infeasible solutions or very small number of feasible solutions when the size of the population is much smaller than the size of the whole search space. To overcome this problem, it is necessary to have some percentage (e.g. $x\%$) of the initial population to be exactly in the form of DRU allocation configuration, on which the algorithm will be applied. In our algorithms we consider *x*=30.

5.5.1 Block Probability-triggered Load Balancing.

The GA and EDA algorithm (Appendix B.1) are utilized to solve this optimization problem. The main-vector C^t is introduced in section 5.3.1 as a solution (string) for our GA and EDA approach, and also our fitness function is the objective function (5.5).

We denote by C^{t+1}=$\left\{c_1^{t+1}, c_2^{t+1}, ..., c_N^{t+1}\right\}$ as an individual (or a chromosome). Each member (c_i^{t+1}) in an individual (or a chromosome) represents the sector to which the corresponding DRUs belong where $c_i^{t+1} = k$ means DRU_i belongs to $Sector_k$. $|\Delta_l|$=500 is population size of GA and EDA algorithms at the l_{th} generation. p_s=0.5 (section B.2) is the selection probability.

In GA, P_c=0.9 is the crossover probability and exchanged with a rate C_R and P_m=0.1 is the mutation probability (section B.2 step 4).

The cost coefficients employed for the objective function in equation (5.5) are q_1=q_3=10, q_2=1, r_1=2, and r_2=r_3=1. Larger weights are given to q_1 and q_3, in order to minimize the blocked calls by hard capacity and *CI* is the first priority in DRU allocation. The weight of the inter-eNodeB hand-off r_1 is twice of that by the intra-eNodeB hand-off r_2 because a hand-off between two sectors on the same eNB can be handled entirely within that eNB, while a hand-off between two eNBs involves not only those eNBs, but also the MME and S-GW as well; this latter case requires a lot more signaling.

However, since the forced hand-off only occurs at the beginning of DRU allocation period, equal weights are given to r_2 and r_3.

We consider three benchmarking problems. The number of DRUs, eNBs and sectors for these three problems are assumed as: P_1 : 19 DRUs, 2 eNBs and 6 Sectors, P_2 : 37 DRUs, 3 eNBs and 9 Sectors, P_3 : 61 DRUs, 4 eNBs and 12 Sectors.

The Monte Carlo analysis is performed in such a way that each algorithm is repeated 30 times over 50 different improper DRU allocations for each benchmarking problem and then takes an average of these results. Each improper DRU allocation is generated in such a way that 80, 177 and 128 blocked calls are considered for each DRU allocation of P_1, P_2, and P_3, respectively.

Fig. 5.4 demonstrates 3 examples of DRU allocation of the three benchmarking problems at time t and $t+1$. Each number inside each cell demonstrates the number of active users. We assumed that all users are treated equally in this paper in terms of traffic model.

In order to test the algorithms' performance, the optimal solution (proper DRU allocation) is obtained by exhaustively searching all possible solutions (K^N=6^{19} solutions for 19-DRU, 9^{37} solutions for 37-DRU and 12^{61} solutions for 61-DRU).

We assumed the hard capacity=200 for each sector (HC_k=$200, k = 1, ..., K$)(Macro-cell is assumed) in all three benchmark problems. We terminated the algorithms after 200 generations in the bench mark problems. The Convergence Rate (*CR*) is defined as the number of times an optimal (or best found) solution is obtained over the entire number of generations performed.

Performances of GA and EDA algorithms:

Results of GA and EDA algorithms and exhaustive search are shown and compared with improper allocation in Table 5.1 and Table 5.2. Fig. 5.5 shows the QoS average convergence to the best found QoS value for all three bench mark problems. Note that super powerful computer is used in order to find the optimal solution for larg-scale scenario (61-DRU) and it took several days to obtained the final optimal solution.

For the 19-DRU problem, The optimal solution is found by using both evolutionary algorithms (EDA and GA) and exhaustive search (ES). The EDA and GA algorithms are converging to the proper optimal DRU allocation (solution) with an evaluation value of QoS=3.488×10^{-3} where the QoS of improper DRU allocation is 1.1×10^{-3}. The CR_{EDA} is 0.945 with 0.082 CPU seconds where the optimal solution is obtained 189 times over 200 generations. But the CR_{GA} is 0.935 with 0.09 CPU seconds where the optimal solution is obtained 187 times over 200 generations. The EDA converges to the optimal solution faster than the GA (Table 5.2 and Fig. 5.5). Optimal solution is found by exhaustively searching all possible solutions (K^N=6^{19} solutions) whereas EDA and GA obtained the optimal solution after 11 and 13 (after evaluating 11×500 and 13×500 solutions ($|\Delta|$=500)), respectively. Therefore, EDA and GA converge to the optimal solution faster than ES.

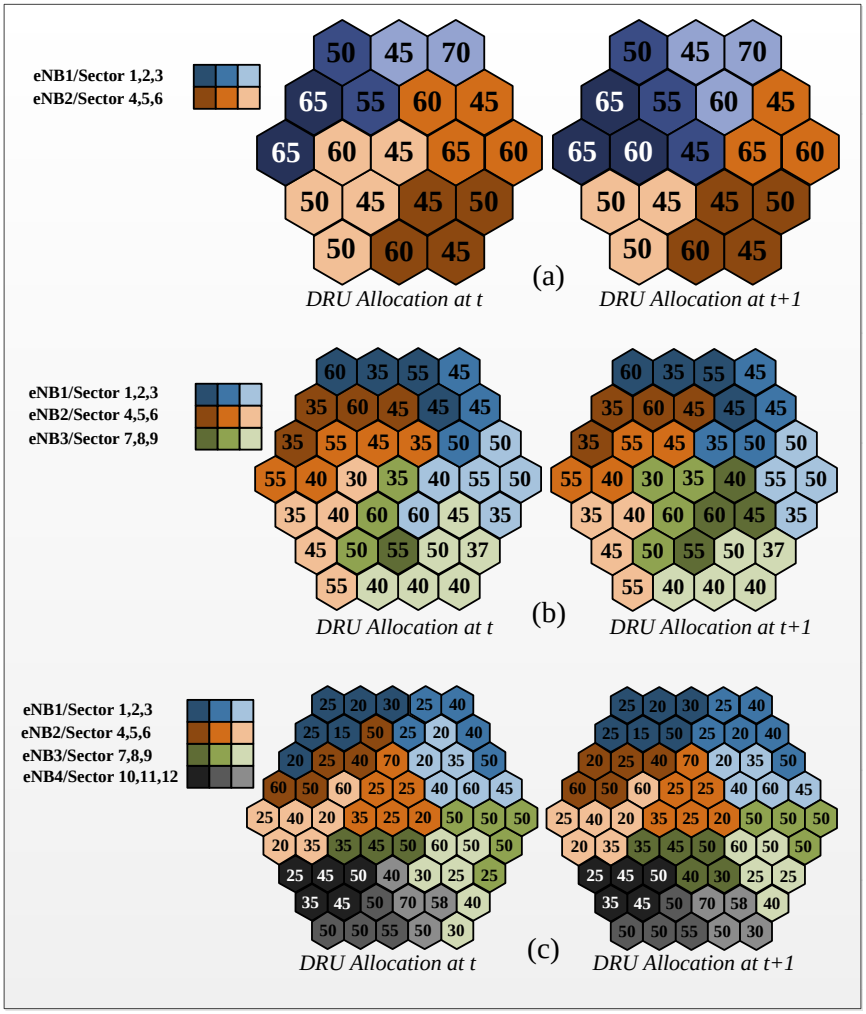


Figure 5.4: Three examples of benchmarking problems: (a)19-DRU, (b)37-DRU and (c) 61-DRU at time *t* and *t+1*.(Each number inside each cell demonstrates the number of active users)

Table 5.1: Computational Results.

		19-DRU	37-DRU	61-DRU
QoS ($\times 10^{-3}$)	Improper Allocation	1.10	0.50	0.77
	GA Result	3.48	2.00	0.87
	EDA Result	3.48	2.47	2.88
	Exhaustive Search	3.48	2.52	2.91
# Blocked Calls	Improper Allocation	80	177	128
	GA Result	0	7.1	14.5
	EDA Result	0	6.9	13.3
	Exhaustive Search	0	5.2	6.3
# Forced Hand-offs	Improper Allocation	0	0	0
	GA Result	152	173.2	114.9
	EDA Result	152	172.1	111.4
	Exhaustive Search	152	153.1	102.8
CI	Improper Allocation	0.52	0.60	0.55
	GA Result	0.48	0.59	0.54
	EDA Result	0.48	0.56	0.49
	Exhaustive Search	0.48	0.54	0.48

Table 5.2: Algorithms comparison results.

	QoS %		CR		CPU time		# generation	
	GA	EDA	GA	EDA	GA	EDA	GA	EDA
19-DRU	2.16	2.16	0.93	0.94	0.09	0.08	13	11
37-DRU	3.00	3.94	0.57	0.74	0.17	0.15	85	30
61-DRU	0.14	2.74	0.62	0.88	0.30	0.28	75	25

Table 5.3: Complexity comparison.

	Exhaustive Search	Evolutionary Algorithm (GA & EDA)
19-DRU	$6^{19}\times 17806$	$500\times 200\times 17806$
37-DRU	$9^{37}\times 137236$	$500\times 200\times 137236$
61-DRU	$12^{61}\times 629584$	$500\times 200\times 629584$

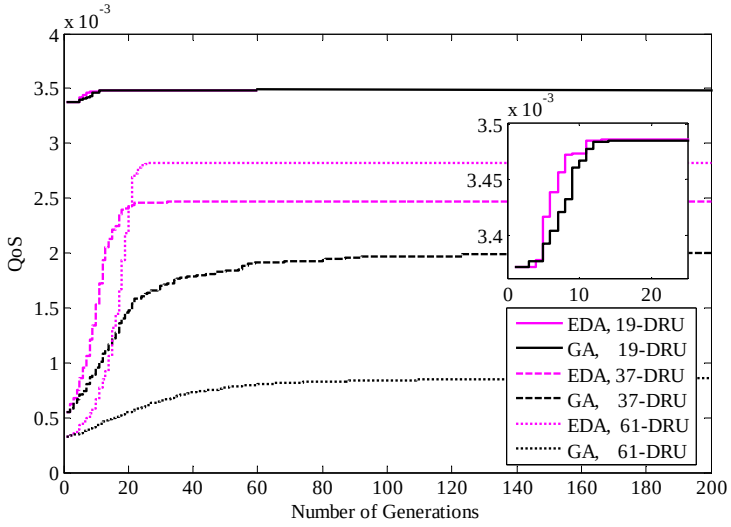


Figure 5.5: QoS value for EDA and GA algorithms in 19-DRU, 37-DRU and 61-DRU.

For the 37-DRU and 61-DRU problems, The evaluation QoS values of the improper and optimal DRU allocation are 0.5×10^{-3} and 2.47×10^{-3}, respectively for 37- DRU and, are 0.77×10^{-3} and 2.8×10^{-3}, respectively for 61-DRU. Note that the optimal DRU allocations are found by exhaustive search method. Since the GA and EDA are not able to find an optimal solution because of the large number of possible solutions and limited number of generations, CR is a possible way to compare the performance of the EDA and the GA. Table 5.1 shows, in both 37 and 61-DRU problems, against GA, the QoSs obtained by EDA are very close to optimal QoS value found by exhaustive search. Note that, both algorithms improve the system performance by finding a better DRU allocation compared to the improper DRU allocation. In 37-DRU, the CR_{EDA} is 0.74 with 0.1532 CPU seconds where the best found solution is obtained 170 times over 200 generations, but the CR_{GA} is 0.57 with 0.169 CPU seconds where the best found solution is obtained 115 times over 200 generations. In 61-DRU, the CR_{EDA} is 0.88 with 0.2815 CPU seconds where the best found solution is obtained 175 times over 200 generations, but the CR_{GA} is 0.62 with 0.3012 CPU seconds where the best found solution is obtained 125 times over 200 generations (Table 5.2 and Fig. 5.5).

Figures 5.6 and 5.7 and Table 5.1 show number of block calls and number of forced hand-offs of both GA and EDA algorithms. Figures 5.5, 5.6 and 5.7 prove that the EDA significantly outperforms the GA not only in terms of the convergence rate, but also the EDA solutions are much closer to the proper solution as compared to the GA. This is due to the way EDA inherently solves this specific problem is completely different from how GA solves the problem. EDA estimates the probability distribution for the best allocation in each generation and uses this probability distribution to generate new allocation population. However, GA generates a new allocation population by

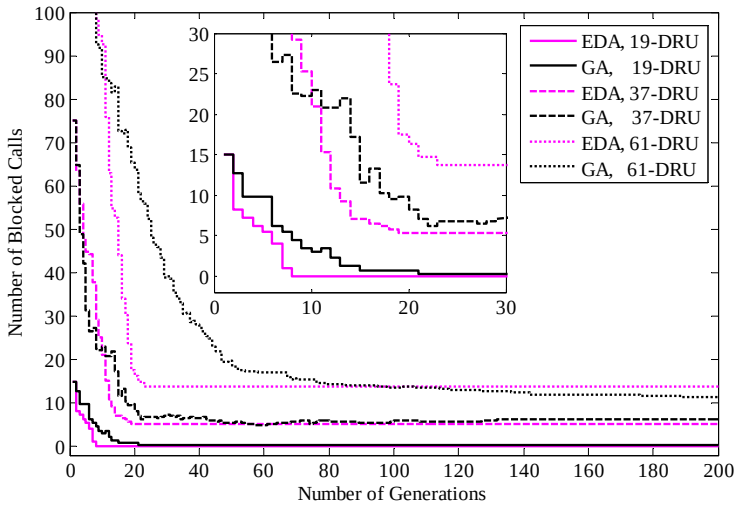


Figure 5.6: Number of Blocked Calls (KPI_{BC}^{-1}) for EDA and GA algorithms in 19-DRU, 37-DRU and 61-DRU.

performing crossovers and mutations. Crossovers and mutations can disrupt good solution and converge in local optimum instead of global solution.

To present improvement level of each algorithm, we define QoS$\%$ $=$ ($QoS_L - QoS_I$)/ QoS_I where QoS_I and QoS_L are the evaluation QoS values of the improper allocation and last generation, respectively. In 19-DRU, QoS$\%$ is 2.16 for both GA and EDA algorithms. In 37-DRU, QoS$\%$ is 3.94 and 3 for EDA and GA algorithms, respectively. In 61-DRU, QoS$\%$ is 2.74 and 0.14 for EDA and GA algorithms, respectively. By considering convergence of 3 different proposed examples, it is clear that the EDA algorithm converges to the close-optimal solution for a larger example in the further algorithm generations (in higher iterations).

These two algorithms have the capability to find a better solution by utilizing the higher number of generations. Also, having fixed number of generations, the algorithms are able to find a better solution using larger population size in each generation. There is a tradeoff between the population size ($|\Delta_l|$) and the number of generations required to achieve a specified performance. Fig. 5.8 shows the tradeoff between the population size ($|\Delta_l|$) and the number of generations in the 61-DRU scenario solved using the EDA based on QoS^{-1}. This tradeoff can be used in order to configure the algorithms to the available hardware resources; e.g., a large population size can be utilized for hardware with a large memory size. Fig. 5.8 shows the optimal solution can be obtained faster using a large population size in comparison to a small population size. On the other hand, with limited hardware resources, we need to spend more time to perform more generations in order to obtain the optimal solution.

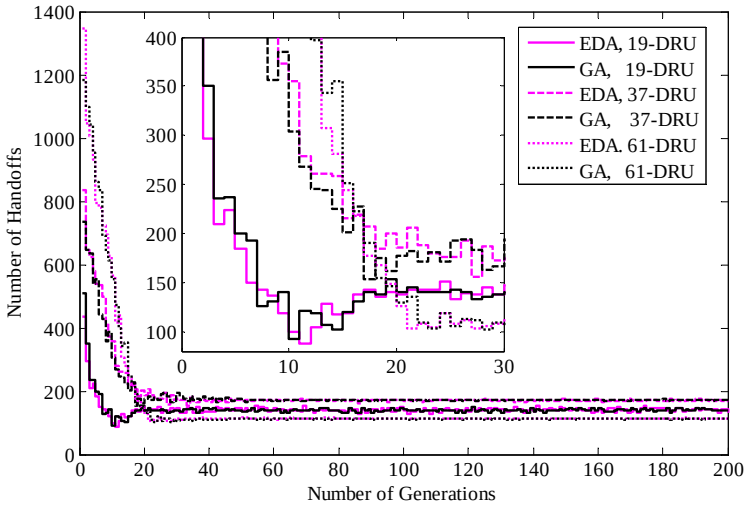


Figure 5.7: Number of Hand-offs for EDA and GA algorithms in 19-DRU, 37-DRU and 61-DRU.

Complexity Comparison:

To measure the computational complexity of the EA algorithms as compared to the exhaustive search algorithm, we consider the number of additions involved in QoS function (fitness function). The number of additions for each possible solution Q^j (N_{add}) to obtain the fitness value is $[NK] + \left[N^2 M\left|SOS_M\right|(K+1)\right] + \left[N^2(K+1)\right] + 3$ where the first, second and third terms are number of additions needed for KPI_{BC}, KPI_{HO} and KPI_{CI}, respectively. Since number of sectors per eNB is 3 ($|SOS_M| = 3$) for all three benchmarking problems, the N_{add}s are 17806, 137236 and 629584 for three P_1, P_2 and P_3 problems, respectively.

Table 5.3 presents the complexity of exhaustive search and the evolutionary algorithms. The exhaustive search algorithm needs to compute $|I_s| \times N_{add}$ additions. Note that, $|I_s|$=K^N is number of all potential solutions which are 6^{19}, 9^{37} and 12^{61} for three P_1, P_2 and P_3 problems, respectively. So the total number of additions needed for P_1 is $6^{19} \times 17806$, for P_2 is $9^{37} \times 137236$, and for P_3 is $12^{61} \times 629584$. The EA algorithms, on the other hand, need to compute $|\Delta|\, I_{Ter} \times N_{add}$ additions. Following these explanations, as an example for P_1, exhaustive search algorithm requires $6^{19} \times 17806$ additions where the EA algorithms require only $500 \times 200 \times 17806$ additions. Note that for EAs, we have assumed I_{Ter}=200 which was illustrated in Fig. 5.5. For larger network, since $|I_s| \,/\, |\Delta|$ is converging to infinity by increasing the DRU number (K) and Sector number (N), the complexity of proposed EA algorithms ($|\Delta|\, I_{Ter} \times N_{add}$) is always ultra-lower than the complexity of exhaustive search ($|I_s| \times N_{add}$).

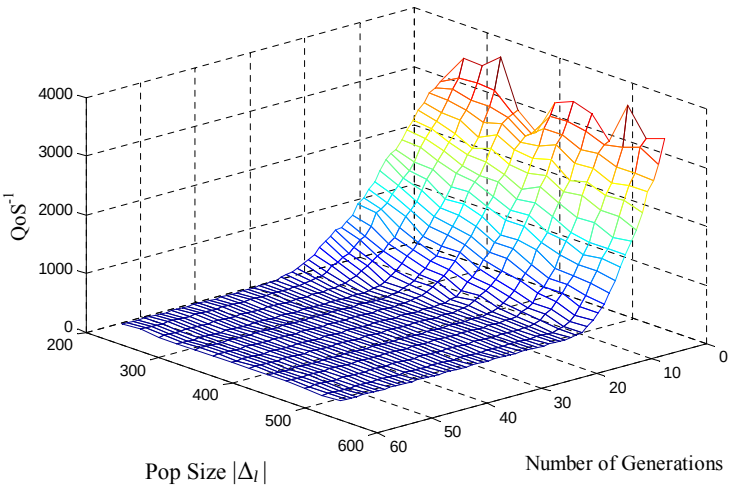


Figure 5.8: The tradeoff between number of individuals/chromosomes (pop size) and number of generations of EDA algorithm in 61-DRU scenario.

5.5.2 Utility Based Load Balancing.

The DPSO algorithm (Appendix B.2) is utilized to solve this optimization problem. The main-vector $\mathbf{u}$ is introduced in section 5.3.2 as a solution (string) for our DPSO approach, and also our fitness function is the objective function (5.14).

We denote by $\mathbf{u}$=$\{u_1, u_2, ..., u_M\}$ as an individual (or a chromosome). Each member (u_m) in a particle represents the eNB to which the corresponding DRUs belong where $u_m = n$ means DRU_m belongs to eNBn.

Two benchmarking problems are considered. In the first benchmark problem which is called 19-DRU (Fig. 5.9 (a)), 6 eNBs (N=6) are used to support 19 uniformly distributed DRUs (M=19). In the second benchmark problem which is called 61-DRU (Fig. 5.9 (b)), 12 eNBs (N=12) are used to support 61 uniformly distributed DRUs (M=61). Each cell represents a different DRU area in Fig. 5.9 where the different colors indicate different DRU groups allocated to a given eNB. Fig. 5.9 shows an example of the initial DRU allocation for the DPSO algorithm. Therefore, based on generating an initial population for algorithm, $x\%$ of the DPSO initial population will be exactly in the form of Fig. 5.9.

The population size ($|\Delta|$) of DPSO for 19-DRU and 61-DRU problems are 6×10^3 and 6×10^7, respectively, Which is less than 1% of the entire search space size. We consider a high level crossover ($c_1 = c_2 = 0.8$) and a low level mutation ($w = 0.2$) for the DPSO algorithm (section B.3). We terminated the algorithms after 2000 generations in both benchmark problems (I_{gen}=2000). The Convergence Rate (CR) is defined as the number of times an optimal solution or the best solution found is obtained over the entire number of generations performed.

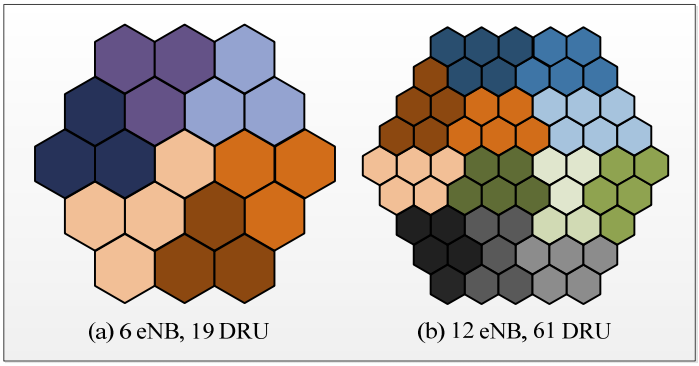


Figure 5.9: Two benchmark problems for DPSO algorithm, (a) 19-DRU, (b) 61-DRU.

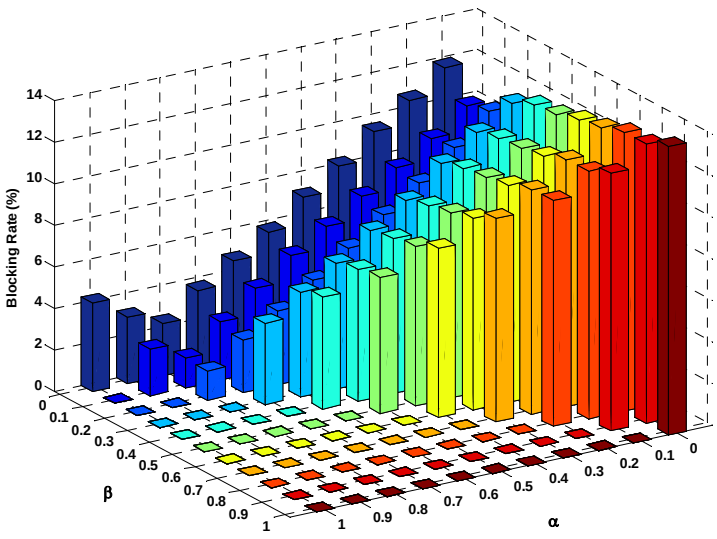


Figure 5.10: Blocking Rate for different α and β.

The distance between adjacent DRUs is 90 meters, the transmission power of each DRU is 36 dBm, and the bandwidth is 10 MHz (50 RBs) and path loss exponent of 3.76. In order to find the optimal control parameters α and β, we evaluate the performance of optimal solutions (DRU allocation) in terms of blocking rate (equation 5.10) with different α and β values (equatio0n 5.14). All the optimal solutions for different α and β values are obtained from an exhaustive search for 19-DRU. In other words, we set α and β to be $0, 0.1, ..., 1$ orderly with constraint $\alpha+\beta \leq 1$. For each pair of α and β, we perform an exhaustive search to find the optimal solution (DRU allocation). The blocking rate with different values of α and β is shown in Fig. 5.10. Placing a larger weight on the load balancing index (large α) until α = 0.8 improves the network balance and decreases the blocking rate. On the other hand, choosing α greater than 0.8 implies that we consider load balancing maximization too much more than average load minimization and average number of handover minimization. This results in allocating improper DRU containing users with inadequate channel conditions, which may consume too much resources. Therefore, the blocking rate increases. Based on the above explanations, we consider α = 0.8 and β = 0.1 which implies a 10$\%$ weighting on minimizing the number of handovers. Since handover performing requires a lot of signaling in target eNB, it is worth taking it into the consideration as an objective.

The Monte Carlo analysis is performed in such a way the DPSO algorithm is repeated 60 times over 50 different initial DRU allocations for each benchmarking problem and then takes an average of these results.

In each Monte Carlo repetition, one fifth of the DRUs are randomly chosen to have a high density of users. The high density DRUs has 4 times the number of users as compared to the non-dense DRUs. The users are uniformly distributed inside the DRU coverage area. For the simulation, we use 6 and 24 users in non-dense and high-dense DRU coverage areas, respectively. The performance of the DPSO algorithms will be investigated by comparing with ES.

Performance of DPSO algorithm:

Fig. 5.11, 5.12, 5.13 and .5.14 show the fitness function, the load balancing index, average load of the network and average number of handovers over the 2000 generations for both benchmark problems, respectively. All optimum values in figures are found after searching all possible solutions ($|N|^{|M|}$) by ES and do not depend on the number of generations. The optimum values in figures help to demonstrate the relative improvement after each generation of the DPSO algorithm. Note that super powerful computer is used in order to find the optimal solution for larg-scale scenario (61-DRU) and it took several days to obtained the final optimal solution.

Fig. 5.11 shows that the DPSO algorithm is converging to the optimal value. In 19-DRU, the convergence rate of the DPSO algorithm to the optimal solution is 0.7 ($CR = 0.7$) where the optimal solution is obtained 1400 times over 2000 generations (Fig. 5.11). In 61-DRU, although the optimal solution is not found during the 2000 generations of DPSO, the best fitness value found is 97% of

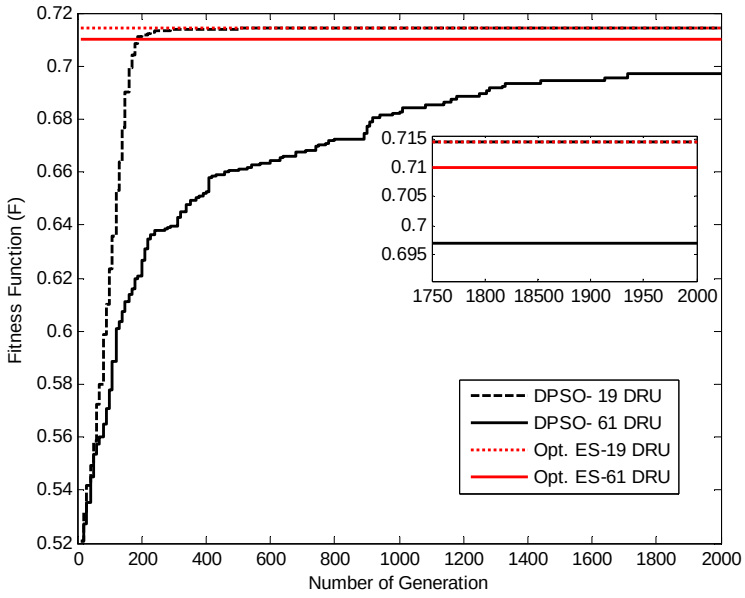


Figure 5.11: Fitness Function value for DPSO algorithm and ES.

the optimal value after only $1710 \times |\Delta|$ fitness evaluations, where the optimal value of ES is found after 12^{61} ($|N|^{|M|}$) fitness evaluations which is too enormous.

In 61-DRU, the convergence rate of the DPSO algorithm to the best solution found is 0.145 (CR=0.145) where the best solution found is obtained 290 times over 2000 generations (Fig. 5.11). Note that, in 61-DRU, DPSO algorithm is likely to yield a better solution if the algorithm runs more generations. Also, with the same number of generations, the algorithm is likely to produce a better solution if a larger population size is used in each generation.

Fig. 5.12 shows that the DPSO algorithm converges to the optimal load balancing index after the 260-th generation in 19-DRU. In 61-DRU, although the optimal solution is not found during the 2000 generations, the best load balancing index found is 98% of the optimal value after only $1730 \times |\Delta|$ fitness evaluations, where the optimal value is found after 12^{61} ($|N|^{|M|}$) fitness evaluations which is too enormous.

Fig. 5.13 and 5.14 shows that the DPSO algorithm converges to the optimal average load and average number of handover after 600-th and 550-th generation in 19-DRU. In 61-DRU, although the optimal solution is not found during the 2000 generations, the best value of average load and average number of handovere found are 98% and 99% of the optimal value after only $1750 \times |\Delta|$ and $1740 \times |\Delta|$ fitness evaluations, where the optimal value is found after 12^{61} ($|N|^{|M|}$) fitness evaluations which is too enormous.

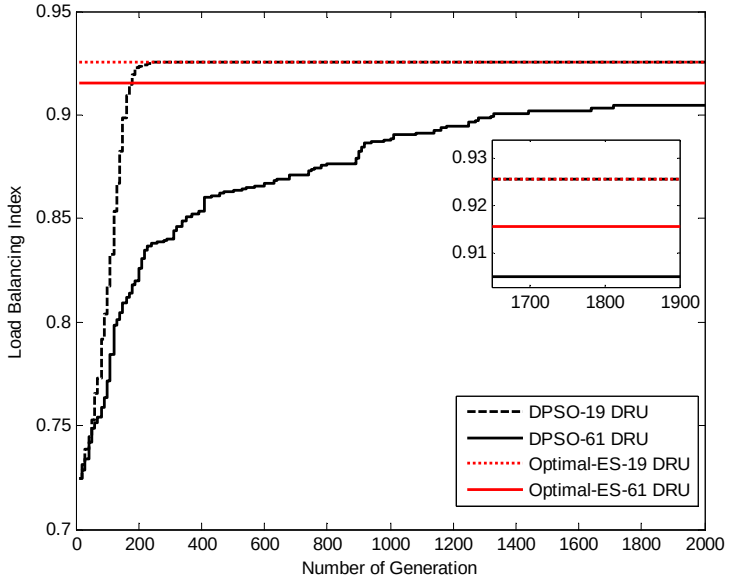


Figure 5.12: Load Balancing Index for DPSO algorithm and ES.

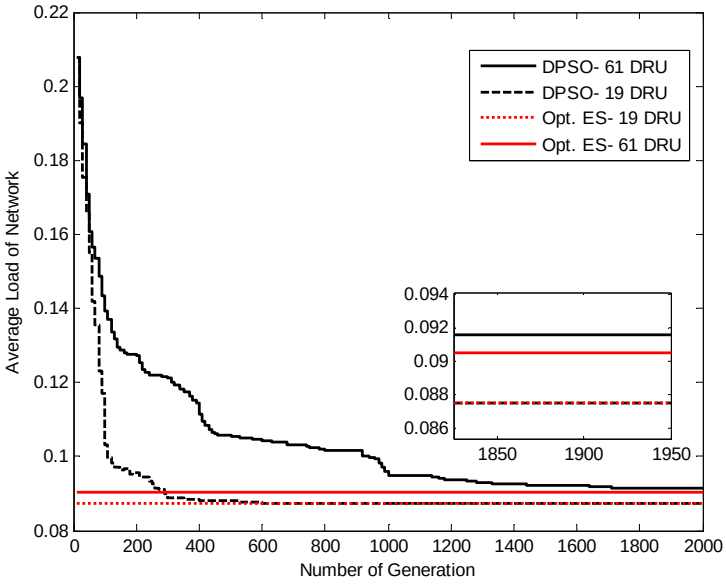


Figure 5.13: Average Load of Network for DPSO algorithm and ES.

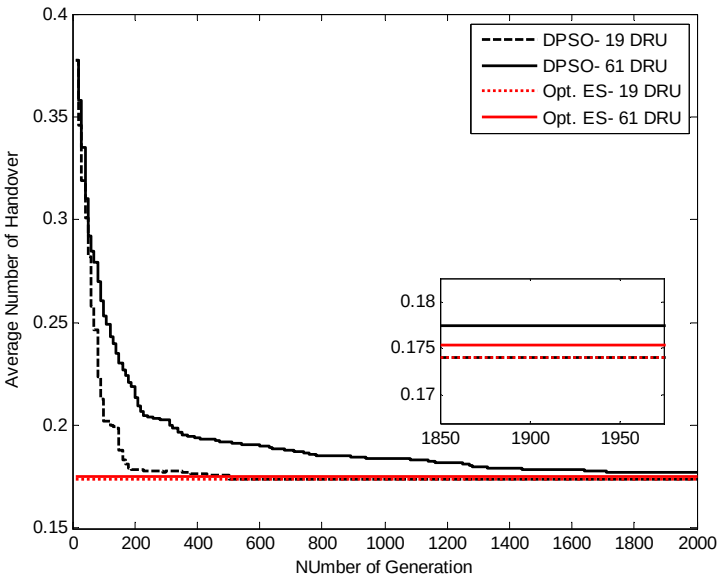


Figure 5.14: Average Number of Handover for DPSO algorithm and ES.

Complexity Comparison:

A good measure of the computational complexity of the DPSO algorithm as compared to the ES is comparing the number of fitness function evaluations. Computation complexity of ES is $O\left(|N|^{|M|}\right)$ where there are N eNBs and M DRUs, so the number of required fitness function evaluation for ES is $|N|^{|M|}$. On the other hand, the complexity of DPSO in the I-th generation is $O\left(|\Delta|\, I\right)$ where the number of required fitness function evaluations at each generation is $|\Delta|$, therefore the number of required fitness function evaluation for DPSO is $|\Delta|\, I$. Note that, the optimum value found by ES requires an enormous number of evaluated fitness functions which is $|N|^{|M|} - |\Delta|\, I$ more than number of evaluated fitness function at the I-th generation of DPSO algorithms.

In 19-DRU, the optimal solution is found after $600 \times 6 \times 10^3$ ($I \times |\Delta|$) fitness evaluations, where the optimal value of ES is found after 6^{19} ($|N|^{|M|}$) fitness evaluations. Note that, the optimum value found by ES requires 6.09×10^{14} ($|N|^{|M|} - |\Delta|\, I$) number of evaluated fitness functions more than number of evaluated fitness function at the 1400-th generation of DPSO algorithms.

In 61-DRU, the best solution is found after $1710 \times 6 \times 10^7$ ($I \times |\Delta|$) fitness evaluations, where the optimal value of ES is found after 12^{61} ($|N|^{|M|}$) fitness evaluations. Note that, the optimum value found by ES requires 6.76×10^{65} ($|N|^{|M|} - |\Delta|\, I$) number of evaluated fitness functions more than number of evaluated fitness function at the 1400-th generation of DPSO algorithms.

5.6 Summary.

DRU allocation of antenna modules is examined with the objective of balancing the traffic in an LTE DAS. Two different load balancing schemes (“Block Probability-triggered Load Balancing” and “Utility Based Load Balancing” schemes) are investigated. In “Block Probability-triggered Load Balancing”, proper DRU allocation is considered to be an effective use of resources in each sector that satisfies hard capacity limit and decreases the hand-off probability. EDA and GA algorithms are developed to solve the DRU allocation problem and their performances are compared with each other. The EDA algorithm produced a noticeably better and faster convergence than the GA algorithm and achieved the optimum performance.

In “Utility Based Load Balancing”, proper DRU allocation is considered to be an effective use of resources in each eNB that prevents high blocking rate. The DRU allocation was formulated as an integer linear programming problem which maximizes the load balancing index and minimizes the average load as well as minimizing the number of handovers. The DPSO algorithm is developed to solve the DRU allocation problem and the performance is compared with an exhaustive search. Using two test problems, the DPSO algorithm produced a noticeably faster convergence than the exhaustive search algorithm.

Chapter 6

Self-Organized Distributed Antenna System using Soft Frequency Reuse

The downlink performance of cellular networks is known to be strongly limited by inter-cell interference in multi-carrier based systems when full frequency resource is utilized. In order to mitigate this interference, a number of techniques have recently been proposed, e.g. the SFR scheme and DAS. In this chapter, we combine these two techniques in a unique cell architecture, which is called DAS-SFR. The central concept of this architecture is to distribute the antenna modules in a hexagonal cell such that the central antenna module transmits the signal using the entire frequency band while the remaining antenna modules utilize only a subset of the frequency bands based on a frequency reuse factor. We propose low-complexity bandwidth allocation scenarios, which do not require a complex processing for allocating bandwidths to users. The primary contribution of this work is the development of an analytical framework to evaluate the average spectral efficiency (ASE) in a DAS-SFR architecture to analyze the potential gains in multi-cell environment. This is an important metric to consider, specifically for users at the cell-edge since modern cellular networks are increasingly required to provide users with high data-rate and a guaranteed QoS, regardless of their geographic location, instead of simply a minimum user throughput which may be acceptable for applications like voice traffic.

Further, we investigate a throughput-balancing scheme that optimizes cellular performance according to the geographic traffic distribution in order to provide a high quality of service (QoS). To enable throughput balancing, we dynamically change the antenna module's carrier power to manage the inter-cell interference. A downlink power self-optimization algorithm is proposed for the DAS-SFR. The transmit powers are optimized in order to maximize the number of satisfied users under different users' distributions. Analytical derivations and simulations are performed to evaluate the system performance and the convergence of proposed algorithm to sub-optimal power level is

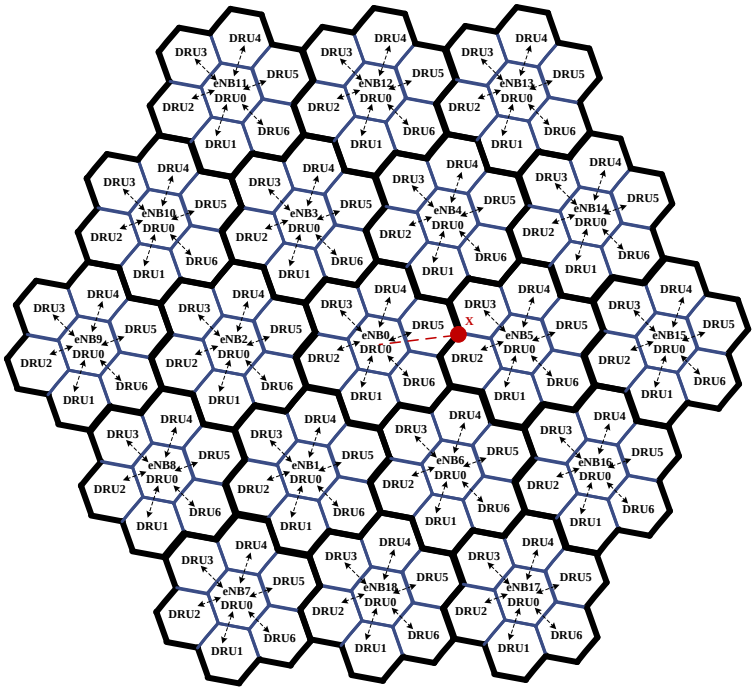


Figure 6.1: Structure of Distributed Antenna System.

presented.

6.1 System Model.

The general architecture of a DAS in an LTE multi-cell environment is shown in Fig. 6.1, where 7 antenna modules (DRUs) are connected to an eNB via an optical fiber and a DAU. We also consider the 2-tier cellular structure, where only two first tiers of eighteen cells surround a given cell. We assume each eNB has a separate DAU. The DAUs are not identified in Fig. 6.1, but are the interface between the eNodeBs and the DRUs. Each of the group of 7 DRUs comprises a geographic cell, which is surrounded by a thick line as seen in Fig. 6.1. The eNBs are linked to a public switched telephone network or a mobile switching center. For the simulcasting operation of DRUs allocated to a given eNB, the access network between each eNB and its DRUs should have a multi-drop bus topology.

We assume the entire spectral bandwidth F is divided into 3 parts (F_1, F_2, F_3). The combination of DAS and conventional cellular system with HFR1 and HFR3 are:

- **DAS-HFR3:** Eeach cell (all its 7 DRUs) transmits on one frequency part base on reuse factor 3 in such a way that neighbor cell DRUs do not use the same frequency (Fig. 6.2 (e)). Note

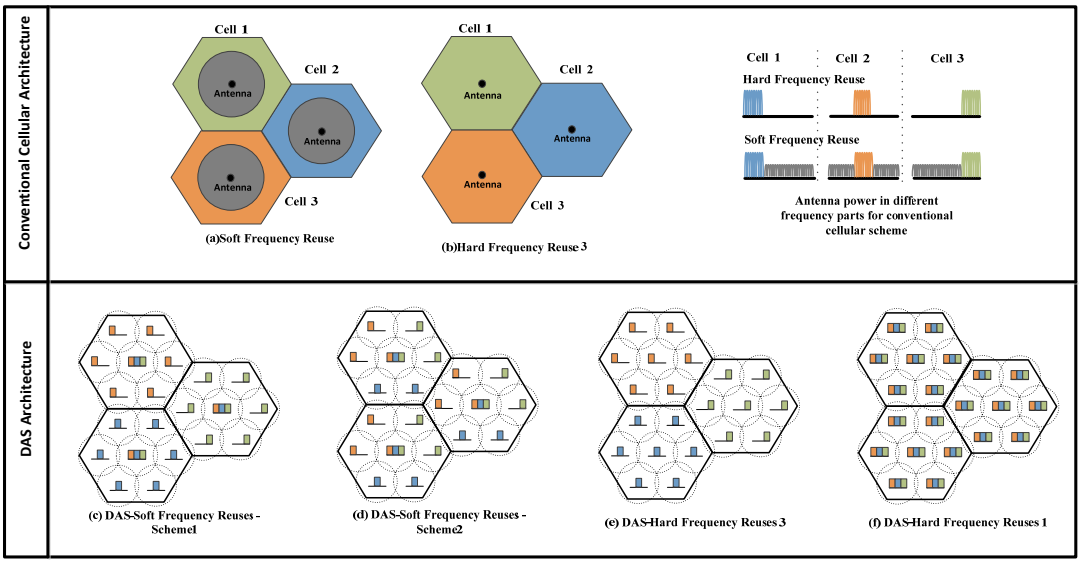


Figure 6.2: Different Frequency Reuse Techniques.

that in our analyses all 7 DRUs of the central cell transmit on frequency band 1 (F_1).

- **DAS-HFR1:** All DRUs in system transmit on all 3 frequency parts (Fig. 6.2 (f)).
- **Con-HFR3:** The central DRUs (DRU0s) of cells transmit the signal on one of (F_1, F_2 and F_3) based on reuse factor of 3 in such a way that DRU0s of neighbor cells do not use the same frequency (Fig. 6.2 (b)). For the analysis, all the edge cells' DRUs are turned off and the central cell DRU0 power is increased by a factor of 7. Note that in our analyses, central DRU of central cell use frequency part 1 (F_1) where central DRUs of neighbor cells use F_2 or F_3 .
- **Con-HFR1:** All the central DRUs transmit on all 3 frequency bands. For the analysis, all the edge cells' DRUs are turned off and the central cell DRU0 power is increased by a factor of 7.

Since combining HFR3 with either conventional cellular system or DAS (Con-HFR3 (Fig. 6.2 (b)) or DAS-HFR3 (Fig. 6.2 (e))) does not use the entire spectrum in each cell, it suffers from a reduced spectral efficiency. However, HFR3 is simple in implementation. HFR3 can be considered the opposite extreme to HFR1 in matters of frequency partitioning techniques. However, combining HFR1 with either conventional cellular system or DAS (Con-HFR1 or DASHFR1 (Fig. 6.2 (f))) uses the entire spectrum in each cell, but it strongly suffers from inter-cell interference. Therefore, the reason to utilize SFR is that, unlike HFR3, the SFR uses the entire spectrum in each cell and, unlike HFR1 the SFR reduces inter-cell interference. But utilizing SFR in a conventional cellular system (Con-SFR) can be impractical in realistic settings involving a large number of base stations and random traffic with high realistic path-loss gain. Con-SFR can be also impractical in realistic

terms since each base station antenna requires complex and expensive hardware tools to transmit on different restricted power bounds for different bands. However, an encouraging result is that by using SFR, signicant performance benefits can still be obtained.

The distance between transmit antenna modules and users is reduced by using a DAS architecture; therefore, low realistic path-loss gain can be attained. In this paper, we combine DAS and SFR to suppress inter-cell interference and make it practical in a realistic setting when path-loss gain is reduced by utilizing DAS. There is no need to acquire different restricted power bounds for different bands. In DAS-SFR, each remote antenna module transmits on same restricted power bound for different bands. Two proposed architectures of the combinations of DAS and SFR which are called DAS-SFR-scheme1 and DAS-SFR-scheme2 are shown in Figures 6.2 (c) and (d), respectively.

- **DAS-SFR-scheme1:** all central DRUs (DRU0s) of cells transmit the signal in all frequency bands (F_1, F_2 and F_3) and the other 6 edge DRUs (DRU1,2,...,6) of each cell transmit on one frequency part based on a reuse factor of 3 in such a way that neighbor cell edge DRUs do not use the same frequency (Fig. 6.2 (c)). Note that in our analyses the 6 edge DRUs of central cell transmit on frequency part 1 (F_1).

- **DAS-SFR-scheme2:** all central DRUs (DRU0s) of cells transmit the signal in all frequency bands (F_1, F_2 and F_3), all DRU4 and 5 of cells transmit the signal on F_1, all DRU1 and 6 of cells transmit the signal on F_2 and all DRU2 and 3 of cells transmit the signal on F_3 (Fig. 6.2 (d)).

The transmission schemes for different frequency parts and transmission power of DRUs are shown in Table 6.1. $P_n^{(i,f)}$ represents the total transmit power of n-th DRU of i-th cell at frequency band f where $P_n^{(i,F_1)} + P_n^{(i,F_2)} + P_n^{(i,F_3)} = 3P$ in all DAS architecture and $P_0^{(i,F_1)} + P_0^{(i,F_2)} + P_0^{(i,F_3)} = 21P$ in all conventional architecture.

Most of the recent publications have focused on investigating SINR advantages of DAS-HFR1 and analyzing its performance [31]. Some publications on DAS-HFR1 have focused on analyzing the uplink performance because of its analytical simplicity [32] and there are few studies on the downlink performance of DAS-HFR1 and DAS-HFR3, although the demand for high-speed data rate will be dominant in the downlink path. Some researchers considered the advantages of single users in DAS-HFR1 for a multi-cell context [33], while there are very few studies on multiple users in DAS-HFR1 and DAS-HFR3. Recent studies on Con-SFR have focused on optimal system design utilizing advanced techniques such as graph theory [34] and convex optimization [35] to maximize network throughput. On the other hand, few papers have studied different combinations of DAS-SFR ([36], [37]), but the comparison of different reference [36] applies SFR on DAS (DAS-SFR-scheme2) with two different frequency factors 3 and 6 where the central antenna modules utilize directional antenna, and the performance of DAS-SFR is compared with only DAS-HFR1.

	Frequency Band F_1					Frequency Band F_2					Frequency Band F_3				
	$P_0^{(0)}$	$P_1^{(0)}$ $P_6^{(0)}$	$P_2^{(0)}$ $P_3^{(0)}$	$P_4^{(0)}$ $P_5^{(0)}$	t	$P_0^{(0)}$	$P_1^{(0)}$ $P_6^{(0)}$	$P_2^{(0)}$ $P_3^{(0)}$	$P_4^{(0)}$ $P_5^{(0)}$	t	$P_0^{(0)}$	$P_1^{(0)}$ $P_6^{(0)}$	$P_2^{(0)}$ $P_3^{(0)}$	$P_4^{(0)}$ $P_5^{(0)}$	t
DAS-SFR-scheme1	P	$3P$	$3P$	$3P$	MISO(7×1)	P	0	0	0	SISO	P	0	0	0	SISO
DAS-SFR-scheme2	P	0	0	$3P$	MISO(3×1)	P	$3P$	0	0	MISO(3×1)	P	0	$3P$	0	MISO(3×1)
DAS-HFR3	$3P$	$3P$	$3P$	$3P$	MISO(7×1)	0	0	0	0	Nothing	0	0	0	0	Nothing
DAS-HFR1	P	P	P	P	MISO(7×1)	P	P	P	P	MISO(7×1)	P	P	P	P	MISO(7×1)
Con-HFR3	$21P$	0	0	0	SISO	0	0	0	0	Nothing	0	0	0	0	Nothing
Con-HFR1	$7P$	0	0	0	SISO	$7P$	0	0	0	SISO	$7P$	0	0	0	SISO

Table 6.1: Transmission classification of central cell for different transmission schemes.

Unlike [36], our proposed DAS-SFR-scheme2 utilize omnidirectional antennas at the central antenna modules. Reference [37] applies SFR on DAS (DAS-SFR-scheme1) and compares it with only DAS-HFR3 and DAS-HFR1. The comparison of different combinations of frequency reuse techniques and DAS (DAS-SFR-scheme1 and 2, DAS-HFR1, DAS-HFR3) with each other and conventional cellular (Con-HFR1 and Con-HFR3) systems have never been studied.

6.2 Low-Complexity Bandwidth Allocation of DAS and SFR Combinations.

In contrast to single user condition where all bandwidths are allocated to only one user without requiring bandwidth allocation scenario, in multi-user condition, we need to consider different scenarios of allocating bandwidths to users. In the following section, we consider different bandwidth allocation scenarios for different DAS-SFR schemes. Most studies on SFR system design have focused on how to determine the size of the frequency partitions [39], for example, in a typical LTE system with a bandwidth of 5 MHz, 25 RBs may be available to serve users for each frequency part ($F_i, i = 1, 2, 3$) [40].

Distinguishing the exterior region from interior region is required to propose different bandwidth allocation scenarios. Assuming that each cell has radius *R*, we define parameter α ($0 < \alpha < 1$) such that the interior users are located in the (0, αR) region and the exterior users are located in the ($\alpha R, R$) region.

In LTE systems, eNB distinguishes between the interior and exterior users based on their corresponding uplink power received at the central DRU (DRU0). To implement this technique, we propose a threshold T_p as a parameter in eNB such that users with uplink power higher than T_p are assigned as interior users (located in interior region), and vice versa. By using this technique,

the bandwidths are allocated to users based on received power which is in contrast with performing complex scheduling (like proportional fair). In complex scheduler, eNB needs to schedule resources at each Transmission Time Interval (TTI) which utilizes a significant amount of hardware resources. Note that the resource allocation scheme employed in the existing literature involves complex scheduling procedures, i.e. proportional fair scheduler, max-min scheduler. More specifically, the complex schedulers are responsible for allocating resources to users at each TTI, and utilize large amount of complex hardware resources processed in the MAC layer. In the proposed bandwidth allocation schemes, however, the three available frequency bands, i.e. F_1, F_2 and F_3 are allocated to different users based on the received signal power. This process does not require complex hardware utilization in any of the upper layers, and can be done at antenna modules front end.

By knowing α,

$$\begin{aligned} N_{interior} &= \alpha^2 N_{users} \\ N_{exterior} &= \left(1-\alpha^2\right) N_{users} \end{aligned} \tag{6.1}$$

where $N_{interior}$ represents the number of interior users, $N_{exterior}$ represents the number of exterior users, and N_{users} represents the total number of users.

6.2.1 Bandwidth Allocation Scenarios for DAS-SFR-scheme1.

For typical central cell, in DAS-SFR-scheme1, we assumed the central DRU (DRU0) transmits the signal in all frequency bands (F_1, F_2 and F_3) and the other 6 edge DRUs (DRU 1, 2,..., 6) transmit on frequency part 1 (F_1). Note that, in DAS-SFR-scheme1 central DRU of each cell transmits on all three frequency bands. Then, three bandwidth allocation scenarios are considered for the central cell of DAS-SFR-scheme1:

- **Scenario 1 (Round Robin):** In this scenario, all F_1, F_2 and F_3 frequency bandwidths are shared and allocated equally to all users in the cell (entire region). Note that in this scenario, F_2 and F_3 frequency bandwidths are inefficiently allocated to the very low SINR exterior region (Fig. 6.4 (DAS-SFR-scheme1-F2 and F3))(the region is more close to edge cell DRUs than central DRU). This is because, the edge DRUs (DRU 1, 2,..., 6) do not transmit on F_2 and F_3 frequency bands at central cell.

- **Scenario 2:** In this scenario, F_1 frequency bandwidth is shared and allocated equally to all users but F_2 and F_3 frequency bandwidths are solely shared and allocated equally to interior users (interior region). Note that in this scenario, all available F_2 and F_3 frequency bandwidths are fully allocated to the interior users (those users are more close to central DRU than edge cell DRUs), which is in contrast with scenario 1 where F_2 and F_3 bandwidths allocated to all users (entire region). This leads to a big gap between the numbers of allocated bandwidths to the interior users as compared to the exterior users.

- **Scenario 3:** In this scenario, same as scenario 2, F_2 and F_3 frequency bandwidths are shared and allocated equally to interior users, but in contrast with scenario 2 all F_1 frequency bandwidth is solely allocated to the exterior users (exterior region). In this scenario, considering SINR map, all three frequency bands are more fairly allocated between all users, as compared to the previously mentioned scenarios. Moreover, in this scenario the frequency bands are allocated to the users following an SINR-based approach, in which the F_1 frequency band is allocated to the exterior users and F_2 and F_3 frequency bands are allocated to interior users.

6.2.2 Bandwidth Allocation Scenarios for DAS-SFR-scheme2.

Since at DAS-SFR-scheme2, we assumed central DRU transmits the signal in all frequency bands (F_1, F_2 and F_3), DRU4 and 5 transmit the signal on F_1, DRU 1 and 6 transmit the signal on F_2 and DRU 2 and 3 transmit the signal on F_3, two bandwidth allocation scenarios are considered for DAS-SFR-scheme2:

- **Scenario 1 (Round Robin):** In this scenario, all F_1, F_2 and F_3 frequency bandwidths are shared and allocated equally to all users in the cell (entire region). Note that in this scenario, F_1, F_2 and F_3 frequency bands are inefficiently allocated to the very low SINR region (Fig. 6.4 (DAS-SFR-scheme2-F1, F2 and F3)). This is because, each edge DRU (DRU 1, 2,..., 6) transmits on one of F_1, F_2 and F_3 frequency bands. Therefore, the region close to edge DRU that transmits on particular frequency band is in low SINR region for the other two frequency bands, e.g. the region close to DRU4 and 5 is in low SINR region for F_2 and F_3 frequency bands because DRU 4 and 5 transmit only on F_1.

- **Scenario 2:** The interior region is supported by F_1, F_2 and F_3 frequency bands. The exterior region is supported following SINR-based approach, whereas each part of the exterior region is supported by only one frequency which has high SINR compared to the two other frequency bands, i.e. the region close to DRU 4 and 5 is supported by only F_1, the region close to DRU 1 and 6 is supported by only F_2 and the region close to DRU 2 and 3 is supported by F_3.

The first scenario of both DAS-SFR-scheme 1 and 2 is non-SINR-based bandwidth allocation (similar to Round Robin scheduling) and the other scenarios are SINR-based bandwidth allocations (similar to SINR-based scheduling such as proportional fair scheduling). Note that, these scenarios do not utilize any complex processing for allocating bandwidths to users where most of the known bandwidth allocation methods (such as proportional fair scheduler) require complex processing.

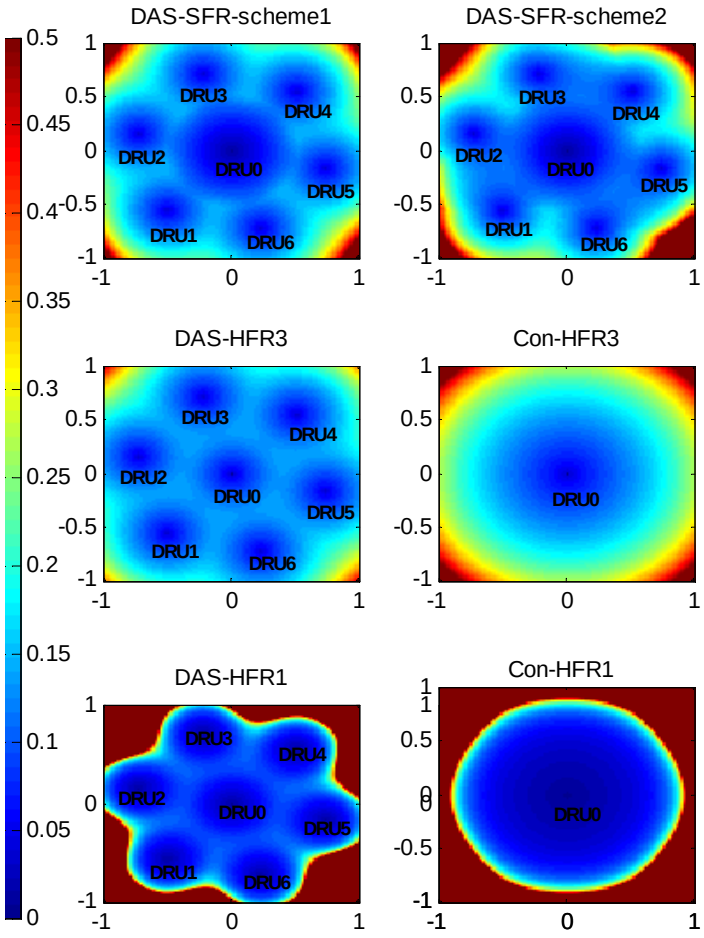


Figure 6.3: Outage map for all different transmission techniques.

6.3 Analysis of DAS and Frequency Reuse Techniques Combinations.

6.3.1 Outage Probability Analysis of Combinations.

The outage occurs when the achieved capacity of each technique $C^{tech} = \sum_{i=1}^{3} C^{tech}_{(F_i)}$ is less than the threshold value R_{th}.

$$\begin{aligned} P_{out}\left[C^{tech} < R_{th}\right] &= P_{out}\left[\sum_{i=1}^{3} \log_2(1+\gamma^{tech}_{F_i,ave}) < R_{th}\right] \\ &= \Pr\left[\gamma^{tech}_{ave} < 2^{R_{th}} - 1\right] \end{aligned} \tag{6.2}$$

where γ_{ave}^{tech} is $\sum_{i=1}^{3}\gamma_{F_i,ave}^{tech} + \gamma_{F_1,ave}^{tech}\gamma_{F_2,ave}^{tech} + \gamma_{F_1,ave}^{tech}\gamma_{F_3,ave}^{tech} + \gamma_{F_2,ave}^{tech}\gamma_{F_3,ave}^{tech} + \prod_{i=1}^{3}\gamma_{F_i,ave}^{tech}$ and, the $C_{(F_i)}^{tech}$ and $\gamma_{F_i,ave}^{tech}$ are the achieved capacity and average SINR of a given technique for the frequency part F_i. And $tech \in \{$ DAS-SFR-scheme1-SC1, DAS-SFR-scheme1-SC2, DAS-SFR-scheme1-SC3, DAS-SFR-scheme2-SC1, DAS-SFR-scheme2-SC2, DAS-HFR3, Con-HFR3, DAS-HFR1, Con-HFR1$\}$. As it is discussed in Appendix A.1, $\gamma_{F_i,ave}^{tech}$ is a lognormal random variable with mean $\mu_{\gamma_{ave}}^{tech,F_i}$ and variance $\sigma_{\gamma_{ave}}^{tech,F_i}$ which can be obtained by equations (A.8) and (A.9) considering what combination technique and which frequency band is being used.

Since the summation and multiplication of lognormal random variables is a lognormal random variable, the γ_{ave}^{tech} becomes a lognormal variable with mean $\mu_{\gamma_{ave}}^{tech}$ and variance $\left(\sigma_{\gamma_{ave}}^{tech}\right)^2$ in dB which can be obtained by Appendix A.2.

Fig. 6.3 represents the outage probability map for all different combination techniques at central cell when the pathloss exponent is 4.0. We assume power control compensates for the lognormal shadowing in the central cell, but the standard deviation of the lognormal shadowing for the signals from neighbor cells is 4 dB. The outage threshold (R_{th}) is assumed as 3.5.

6.3.2 Regional Capacity Analysis of Combinations for different frequency bands.

The extension of the analytical framework of the ASE (Appendix A.3) in DAS-SFR architecture is also presented to analyze the regional capacity in multi-users multi-cell environment. In the multi-users case scenario where users are uniformly distributed, knowing each position's ASE and user numbers, as well as the pre-defined bandwidth allocation scheme in different regions, it is possible to derive the regional capacity in a network with multiple participating users. Due to the uniform distribution of users' positions in different zones, instead of deriving individual throughput for each user, we study the regional capacity.

The amount of allocated F_i bandwidth to each user in different regions (interior or exterior) for different scenarios is $\frac{W}{N_{users}}\beta_{(F_i)}^{(region,tech)}$ where $\beta_{(F_i)}^{(region,tech)}$ is expressed in Table 6.2. We also assume equal bandwidth for each of the frequency bands ($W_{(F_1)}$=$W_{(F_2)}$=...=$W_{(F_M)}$=W_F).

If we consider the bandwidths which are allocated to the different regional users under the different scenarios, the capacity metric of the total system can be obtained by,

$$C^{(\text{region,tech})} = C_{(F_1)}^{(\text{region,tech})} + C_{(F_2)}^{(\text{region,tech})} + C_{(F_3)}^{(\text{region,tech})} \tag{6.3}$$

where

$$C_{(F_i)}^{(\text{region,tech})} = \frac{W}{N_{users}}\beta_{(F_i)}^{(\text{region,tech})}.\vartheta^{(t(F_i,\text{tech},Cell_{num}))}$$

tech $\in \{$ DAS-SFR-scheme1-SC1, DAS-SFR-scheme1-SC2, DAS-SFR-scheme1-SC3, DAS-SFR-scheme2-SC1, DAS-SFR-scheme2-SC2, DAS-HFR3, Con-HFR3, DAS-HFR1, Con-HFR1$\}$, and

	Interior region DRU 0			Exterior region								
				DRU 4 & 5			DRU 1 & 6			DRU 2 & 3		
	F_1	F_2	F_3	F_1	F_2	F_3	F_1	F_2	F_3	F_1	F_2	F_3
DAS-SFR-scheme1-SC1	1	1	1	1	1	1	1	1	1	1	1	1
DAS-SFR-scheme1-SC2	1	A*	A	1	0	0	1	0	0	1	0	0
DAS-SFR-scheme1-SC3	0	A	A	B*	0	0	B	0	0	B	0	0
DAS-SFR-scheme2-SC1	1	1	1	1	1	1	1	1	1	1	1	1
DAS-SFR-scheme2-SC2	C***	C	C	C	0	0	0	C	0	0	0	C
DAS-HFR3	1	0	0	1	0	0	1	0	0	1	0	0
Con-HFR3	1	0	0	1	0	0	1	0	0	1	0	0
DAS-HFR1	1	1	1	1	1	1	1	1	1	1	1	1
Con-HFR1	1	1	1	1	1	1	1	1	1	1	1	1

* $A=\frac{1}{\alpha^2}$, ** $B=\frac{1}{\left(1-\alpha^2\right)}$, *** $C=\frac{1}{\alpha^2+\frac{1}{3}\left(1-\alpha^2\right)}$

Table 6.2: $\beta_{(F_i)}^{(region,tech)}$

.

region $\in$ {interior region, exterior region}, and SC means scenario and $t(F_i, tech, Cell_{num}) \in \{SISO, MISO\}$, e. g. t is MISO for F_1 at central cell in DAS-SFR-scheme1 (Table 3.1).

6.3.3 Analytical and Simulation Results.

In simulation results, zero noise variance ($\sigma^2_{n^{(0)}} = 0$) is considered. Fig. 6.4 represents the average SINR map for individual frequency bands for all different combination techniques at central cell, which is explained in Appendix A. It also justifies the applied bandwidth allocation schemes at different instances, and helps with better understanding of different combination techniques.

In DAS-HFR3 and Con-HFR3, the frequency bands F_2 and F_3 are not utilized at central cell, therefore, only F_1 SINR is presented. On the other hand, the F_1, F_2 and F_3 SINR maps of DAS-HFR1 and Con-HFR1 are presented and they are the same as each other.

Fig. 6.5 and Fig. 6.6 represent the ASE of cellular DAS for different frequency reuse techniques versus the normalized distance from the central cell DRU0 in the direction of the worst case position *X* when the path loss exponent is 3.76 (*X* is presented in Fig. 6.1 and Fig. 6.4).

The total ASE of the system can be obtained by adding the ASE of the individual frequency bands. Therefore $\vartheta_{\text{total}} = \sum_{i=1}^{3} \vartheta^{(t(F_i,tech,Cell_{num}))}$ where $t(F_i, tech, Cell_{num}) \in$ {SISO,MISO}, e.g., t is MISO for F_1 at central cell in DAS-SFR-scheme1 (Table 6.1).

Each scenario is plotted for the individual ASEs with different frequency bands, and for the total ASE, ϑ_{total}.

Fig. 6.5 compares the ASE performance of the 6 different combinations, for the individual frequency bands at a particular central cell. This figure shows an interesting non-monotonic relationship between ASE and the normalized distance from the central DRU (eNB0s DRU 0) for certain

frequency bands. This is because, as the distance increases from the central DRU0 to position *X*, corresponding signal of edge DRUs to that certain frequency band dominates at approximately 0.6*R* which is shown in SINR-based figure (Fig. 6.4), e.g, SINR map shows, DRU 4 and 5 signals at F_1 dominate at approximately $0.6R$. On the other hand, the rest of the frequency bands ASEs decrease when the normalized distance increases from the central DRU0. This is because, as the distance increases from the central DRU0 to position *X*, the received powers of those frequency bands and their SINR level decrease when the normalized distance increases (based on SINR-based figure in Fig. 6.4).

Fig. 6.5 shows that F_1 ASE of DAS-HFR3 is greater than the other DAS architecture due to using more power at DRUs for this frequency (Table 6.1) and absence of interfering signal from close neighboring cells. However, the same power level is utilized for transmitting F_1 in DAS-HFR3 and Con-HFR3 at central cell, but F_1 ASE of DAS-HFR3 and Con-HFR3 are different during travelling user, this is because, all power is aggregated at central DRU for Con-HFR3 where the power is distributed equally between all 7 DRUs for DAS-HFR3. F_1 ASE of DAS-SFR-scheme1 is greater than DAS-SFR-scheme2, this is because, all 7 DRUs transmit on F_1 for DAS-SFR-scheme1 where only 3 DRUs transmit on F_1 for DAS-SFR-scheme2 and also only one DRU (central DRU) introduces interfering signal from each neighboring cell in DAS-SFR-scheme1. Note, 3 DRUs introduce interfering signals from each neighboring cell in DAS-SFR-scheme2. F_2 and F_3 ASE of DAS-SFR-scheme1 are less than DAS-SFR-scheme2 while the distance is increased. This is because as Fig. 3 shows SINR level of F_2 and F_3 bands in DAS-SFR-scheme1 are less than that of DAS-SFR-scheme2. Note that only one DRU (Central DRU) transmits on F_2 and F_3 for DAS-SFR-scheme1 where 3 DRUs transmit on F_2 and F_3 for DAS-SFR-scheme2. It is worthwhile mentioning that the difference between F_2 and F_3 capacities for DAS-SFR-scheme2 is because of the different SINR levels as is shown in Fig. 6.4.

Fig. 6.5 shows that the DAS-HFR1 and Con-HFR1 methods achieve a higher total ASE than the other methods inside a normalized distance 0.5. However the significant reduction happens outside a normalized distance 0.8, due to the high level interference introduced from the edge DRUs of neighboring cell.

Note that, in the DAS-SFR techniques, all 3 frequency bands are assigned, whereas, in the HFR3 methods, only the F_1 band is used. Therefore, the results show that the DAS-SFR methods achieve higher ASE than Con-HFR3 and DAS-HFR3. More specifically, the total ASE achieved using the DAS-SFR methods are higher inside the normalized distance 0.5 than Con-HFR3 and DAS-HFR3, where the central DRU0 transmits signal on 3 frequency parts. It is worth to mention that the ASE of DAS-SFR methods are slightly lower than that DAS-HFR3 and Con-HFR3 outside the normalized distance 0.8. This is because there is weak interference introduced from the DRUs of neighboring cells when the DAS-SFR-scheme1 and DAS-SFR-scheme2 methods are applied.

The difference between the ASE of DAS-SFR-scheme1 and scheme2 is because of the different

SINR experienced (Fig. 6.4).

About 50 percent of the users are located inside the $(0.8R,\ R)$ region, assuming a uniform distribution, while the other 50 percent are farther from the interfering cells. Therefore, it is more important to improve the system capacity inside the $(0.8R,\ R)$ region. In Fig. 6.7, the regional capacity results ($C^{(\text{region},\ tech)}\frac{N_{\text{users}}}{W}$) are provided considering a uniform distribution of users. Note that, in this example, we assume α is 0.4 and path loss exponent is 3.76.

As Fig. 6.7 shows, by comparing different schemes and bandwidth allocation scenarios of DAS-SFR, the achieved regional capacity of DAS-SFR-scheme2-SC2 is more fairly distributed inside the entire system and there is not a significant gap between different geographic areas. This is because, using DAS-SFR-scheme2-SC2, the frequency bandwidths are more fairly allocated to the regions in comparison to the other architectures. Note that, using the other DAS-SFR-scheme1-SC2 and SC3, there is a significant gap between interior and exterior regions.

Due to the no-fair allocation of bandwidths to users at DAS-SFR-scheme1-SC1 and DAS-SFR-scheme2-SC1, the achieved regional capacity of these scenarios is lower than the other techniques. This is because, lots of portion of bandwidths are inefficient allocated to regions with low SINR level. The achieved capacity of the DAS-SFR-scheme1-SC2 and SC3 is higher than that of DAS-SFR-scheme1-SC1, especially near the central DRU. This is due to the fact that the F_2 and F_3 bandwidths are not allocated to the exterior users which have very low SINR level (Fig. 6.4) in DAS-SFR-scheme1-SC2 and SC3. In DAS-SFR-scheme1-SC2 and SC3, all F_2 and F_3 bandwidths are allocated to interior users.

Fig. 6.7 shows, by comparing DAS-SFR-scheme1-SC2 and SC3, the achievable regional capacity for the interior region in DAS-SFR-scheme1-SC2 is slightly higher than that of SC3 due to the partial allocation of F_1 bandwidths to interior region in SC2 as well as F_2 and F_3 bandwidths, where only the F_2 and F_3 bandwidths are allocated to the interior region in SC3. Conversely, the achieved capacity for the exterior region in a DAS-SFR-scheme1-SC3 is slightly higher than that of SC2, due to the fact that the F_1 bandwidth is only allocated to the exterior region, whereas in SC2 the F_1 is partially allocated to the exterior users.

Fig. 6.7 illustrates that the DAS-SFR schemes significantly improve the system performance by making up for the dead spots, which appear at the cell edges. There is a significant amount of interference happening in the $(0.8R, R)$ region, causing dead spots in DAS-HFR1 and Con-HFR1 techniques.

Let's assume that $TBps/Hz$ is the guaranteed capacity of the system. A more desirable transmission technique is the one whose capacity is greater than $TBps/Hz$ in a wider geographic area in comparison to other techniques. Therefore, the most appropriate combination and bandwidth allocation scenario not only depend on achieved capacity inside the cell, but also on T level.

Fig. 6.8 shows the regional capacity distribution of each techniques in terms of CDF. It also compares different combination techniques with different guaranteed threshold throughput. The

CDF index shows the percentage of the coverage area that enjoys a certain guaranteed throughput. We compare different techniques with each other for $T = 10$, $T = 25$ and $T = 35$. As an example, when $\alpha = 0.4$, at $T = 10$, Fig. 6.8 shows 36, 64, 70, 78, 80, 82, 85, 86 and 94 percent of coverage area has regional capacity less than 10 (bps/Hz) in DAS-SFR-scheme2-SC2, DAS-HFR1, DAS-SFR-scheme1-SC3, DAS-SFR-scheme1-SC2, Con-HFR1, DAS-SFR-scheme1-SC1, DAS-SFR-scheme2-SC1, DAS-HFR3 and Con-HFR3, respectively. Therefore, DAS-SFR-scheme2-SC2 covers more geographic area than the other transmission techniques at $T = 10$. At $T = 25$, Fig. 6.8 shows 88, 90, 92, 94, 96, 99, 99, 100 and 100 percent of coverage area has regional capacity less than 25 (bps/Hz) in DAS-SFR-scheme2-SC2, DAS-SFR-scheme1-SC2, DAS-SFR-scheme1-SC3, DAS-HFR1, Con-HFR1, DAS-SFR-scheme1-SC1, DAS-SFR-scheme2-SC1, DAS-HFR3 and Con-HFR3, respectively. Therefore, DAS-SFR-scheme2-SC2 covers more geographic area than the other transmission techniques at $T = 25$. At $T = 35$, Fig. 6.8 shows 93, 94, 96, 98, 99, 100, 100, 100 and 100 percent of coverage area has regional capacity less than 35 (bps/Hz) in DAS-SFR-scheme1-SC2, DAS-SFR-scheme1-SC3, DAS-SFR-scheme2-SC2, DAS-HFR1, Con-HFR1, DAS-SFR-scheme1-SC1, DAS-SFR-scheme2-SC1, DAS-HFR3 and Con-HFR3, respectively. Therefore, DAS-SFR-scheme1-SC2 covers more geographic area than the other transmission techniques at $T = 35$. By comparing results of different T, it is concluded that the more desirable transmission technique depends on the T level.

Figs. 6.9 and 6.10 present the impact of variations in α from 0.2 to 0.8 on the regional capacity for different geographical zones (edge and non-edge region). More importantly, changes in α directly result in changes in the amount of bandwidths allocated to different zones in DAS-SFR-scheme1-SC2 and SC3, and DAS-SFR-scheme2-SC2.

Fig. 6.9 and 6.10 illustrate the analytical results for the non-edge region (inside (0, 0.8R)) and edge region (inside (0.8R, R)), where α varies between 0.2 and 0.8. Note that as α changes, the amount of bandwidths allocated to the different regions in DAS-SFR-scheme1-SC2 and SC3, and DAS-SFR-scheme2-SC2 changes accordingly (TABLE. II). However, in other techniques, as α changes, the amount of allocated bandwidths to the different regions remains constant as α changes.

As α increases in Fig. 6.9, average regional capacity of non-edge region decreases in DAS-SFR-scheme1-SC2 and SC3 and DAS-SFR-scheme2-SC2. This is because by increasing α, the constant amount of bandwidths is allocated to more geographic area. The average regional capacity of DAS-SFR-scheme1-SC2 and SC3 is higher than DAS-SFR-scheme2-SC2 when α is small. This is because, all F_2 and F_3 bandwidths are only allocated to interior region with high SINR region shown in Fig. 6.4 (when α is small) for DAS-SFR-scheme1-SC2 and SC3, but in DAS-SFR-scheme2-SC2 each frequency band is allocated not only to interior region but also to a portion of exterior region.

By increasing α, the rate of decrease in non-edge region capacity for DAS-SFR-scheme1-SC2

and SC3 is more than that of DAS-SFR-scheme2-SC2. By increasing α, F_2 and F_3 bandwidths in DAS-SFR-scheme1-SC2 and SC3 are allocating to lower SINR region than DAS-SFR-scheme2-SC2. Note that, in DAS-SFR-scheme1, the rate of decrease in capacity for non-edge region for SC3 is more than that of SC2, as α increases. The reason behind this is that the rate of decrease in the portion of the bandwidth allocated to the non-edge region in SC3 is more than that of SC2. Specifically, as was previously mentioned, since no F_1 bandwidth is allocated to the non-edge region in DAS-SFR-scheme1-SC3, non-edge capacity in SC3 is less than that of SC2 for DAS-SFR-scheme1.

As α increases in Fig. 6.10, same as the non-edge region, the average capacity of the edge region decreases in DAS-SFR-scheme2-SC2, due to the fact that the constant portion of bandwidth is allocated to more geographic area. But, using DAS-SFR-scheme1-SC3, as α increases, the average capacity of the edge region increases, due to the fact that the constant portion of bandwidth is allocated to less geographic area. Specifically, as α increases, fewer regions use F_1 bandwidth in DAS-SFR-scheme1-SC3. However, in other techniques, even DAS-SFR-scheme1-SC2, edge region experiences a constant capacity, this is because, the portions of bandwidth allocated to the edge users remain constant as α changes.

Fig. 6.11 presents the impact of variations in α on the regional capacity for different combination architectures in terms of CDF where CDF index shows the percentage of the coverage area that enjoys a certain guaranteed throughput. Note that, only the capacity of these three DAS-SFR-scheme1-SC2, DAS-SFR-scheme1-SC3 and DAS-SFR-scheme2-SC2 techniques changes by altering α. Different transmission techniques can be compared with each other by choosing α and T.

By comparing DAS-SFR schemes, DAS-HFR3 and DAS-HFR1 architectures result in Fig. 6.5 and 6.6, it is worth to mention that, each architecture can outperform the others inside one of the three different $(0, 0.45R)$, $(0.45R, 0.8R)$ and $(0.8R, R)$ regions. So the average capacity of three $(0, 0.45R)$, $(0.45R, 0.8R)$ and $(0.8R, R)$ regions for different architecture are shown in Figures. 6.12, 6.13 ,and 6.14, which are individually plotted versus pathloss exponent where the α is 0.5.

As an example, at pathloss exponent 4, Fig. 6.12 shows the DAS-SFR-scheme1-SC2, SC3 and DAS-SFR-scheme2-SC2 outperform the other techniques inside $(0, 0.45R)$ where more bandwidths are allocated to the interior users ($\alpha^2 N_{users}$) than the other techniques. Note that the DAS-SFR-scheme1-SC2, SC3 and DAS-SFR-scheme2-SC2 are the first, second and third techniques, which have higher average capacities, respectively. Fig. 6.13 shows the DAS-SFR-scheme2-SC2, DAS-HFR1 and DAS-SFR-scheme1-SC3 outperform the other techniques inside $(0.45R, 0.8R)$. Note that the DAS-SFR-scheme2-SC2, DAS-HFR1 and DAS-SFR-scheme1-SC3 are the first, second and third techniques which have higher average capacity, respectively. Fig. 6.14 shows the DAS-SFR-scheme2-SC2, DAS-SFR-scheme1-SC3 and DAS-HFR3 outperform the other techniques inside $(0.8R, R)$. Note that the DAS-SFR-scheme2-SC2, DAS-SFR-scheme1-SC3 and DAS-HFR3 are the

PARAMETERS	VALUE
Channel Bandwidth for each Frequency Part	5 MHz
Carrier Frequency	2.14 GHz
FFT size	1024
Number of Resource Blocks for each Frequency Part	25
Subcarrier Spacing	15 kHz
Cellular Layout	Hexagonal grid, 19 sites
Inter-eNB Distance	500 meters
Log-normal Shadowing	8 dB
Propagation loss	128.1+37.6 $\log_{10}(R(km))$
White Noise Power Density	-174 dBm/Hz
Scheduling	Round Robin
TTI	1 ms

Table 6.3: Simulation Parameters.

first, second and third techniques which have higher average capacity, respectively.

The performance of the above mentioned architectures is also investigated through LTE system level simulations. We consider the two-ring hexagonal cellular system with nineteen eNBs, such that each cell has 7 DRUs, as depicted in Fig. 6.1, where the eNBs distance is 500 meters. We use the simulation parameters listed in Table 6.3. The presented numerical results corroborate the analytical results.

6.4 User Throughput Analysis.

The real user throughput can be obtained by multiplying the quantity of effective bandwidth per user by its ASE. Therefore, having the quantity of effective bandwidth of *f* assigned to user *k* ($W_k^{(f)}$), the real throughput at user *k* can be written in terms of *bps* (bit per second) as follows,

$$C_k^{real}(\mathbf{\Phi}) = \sum_{i=1}^{3} C_k^{(F_i,tech)}(\mathbf{\Phi}) \tag{6.4}$$

where

$$C_k^{(F_i,tech)}(\mathbf{\Phi}) = W_k^{(F_i)} \vartheta^{(t(F_i,tech,Cell_{num}))}(\mathbf{\Phi})$$

where $t(F_i, tech, Cell_{num}) \in \{SISO, MISO\}$ and $tech$ denotes the transmission scheme, for example t is MISO for F_1 at central cell in DAS-SFR (Table 6.1). $C_k^{(F_i,tech)}(\mathbf{\Phi})$ is the user throughput *k* where $\mathbf{\Phi} = \left\{ P_n^{(i,f)} \, | n = 0, 1, ..., 6 \ , \ i = 0, 1, ..., 18 \right\}$. $P_n^{(i,f)}$ represents the total power of n-th DRU of i-th cell at frequency band f.

6.5 The Power Self-Organized Distributed Antenna System using Soft Frequency Reuse.

So far, only uniform user distribution is considered in order to compare techniques with each other where in real case users are not distributed uniformly. Therefor power self organized algorithm is proposed in order to proved better quality of service for non-unifomly user distribution by changing the antenna modules power. This algorithm is only applied on DAS-SFR-scheme1 in further studies.

6.5.1 Formulation of Power Allocation.

In a multi-user system, we are interested in a proper power allocation which maximizes the number of satisfied users and their throughput. Without proper power allocation, there may be cases of unbalanced throughput between users where a few users can have ultra-high throughput and most of the users have ultra-low throughput. Proper power allocation can control the signal to interference ratio in order to balance throughput between users and decrease the number of unsatisfied users.

In this section, we formulate the power allocation problem to maximize the number of satisfied users and also maximize the total satisfied users' throughput.

We consider the following KPIs in the power allocation problem:

1. $\mathbf{KPI_{SU}}$ ***(Number of Satisfied Users):*** We can derive a metric defining a percent of satisfied users (i.e., users that can achieve the targeted service bit rate (C_{th})). The fraction of satisfied users would be,

$$KPI_{SU}(\mathbf{\Phi}) = \frac{\sum_{k=1}^{N_{users}} G_k(\mathbf{\Phi})}{N_{users}} \tag{6.5}$$

where N_{user}^{total} is total number of users and,

$G_k(\mathbf{\Phi}) = \begin{cases} 1 & when \quad C_k^{real}(\mathbf{\Phi}) \geq C_{th} \\ 0 & otherwise \end{cases}$.

$G_k(\mathbf{\Phi})$ is equal to 1 when user k achieves the target bit rate C_{th} ,

2. $\mathbf{KPI_{CSU}}$ ***(Capacity of Satisfied Users):*** The total achieved capacity of satisfied users would be,

$$KPI_{CSU}(\mathbf{\Phi}) = \frac{\sum_{k \in SUS} C_k^{real}(\mathbf{\Phi})}{\left(\sum_{i=1}^{3} W_{(F_i)} \right) / 3} = \frac{3 \sum_{k \in SUS} C_k^{real}(\mathbf{\Phi})}{\sum_{i=1}^{3} W_{(F_i)}} \tag{6.6}$$

where $W_{(f)}$ is the bandwidth of frequency band *f* and $SUS = \{k\,|G_k = 1, k = 1, 2, ..., N_{users}\}$ is the satisfied users set.

Now, our QoS function is the weighted combination of the two KPIs (cost factors) which we have already introduced. Our objective function is to maximize the QoS function.

$$\underset{\mathbf{\Phi}}{Maximize} \quad QoS(\mathbf{\Phi}) = \alpha KPI_{SU}(\mathbf{\Phi}) + \beta KPI_{CSU}(\mathbf{\Phi}) \tag{6.7}$$

Note that, the parameters α and β are chosen in such a way that which KPI metric has more priority than the other KPI.

We can further simplify the KPIs in (6.5) and (6.6) based on the following arguments $(\mathbf{A})$ and $(\mathbf{B})$:

(A): Since it is not practical to calculate the ASE for the individual users, the simplification below is valid for the theoretical analysis. However, in practice, the number of the satisfied users and therefore the KPIs are found based on the real users' throughput ($C_k^{real}(\mathbf{\Phi})$) measurement after the power allocation procedure.

We assume bandwidth of f are equally shared to those users which are located in f supporting area according to different bandwidth allocation scenarios at section 6.2. e.g., F_1 are equally shared to exterior users for bandwidth allocation scenario 3 for DAS-SFR-scheme1. We also assume equal bandwidth for each of the frequency bands ($W_{(F_1)}$=$W_{(F_2)}$=$W_{(F_3)}$=W_F). Therefore, we can rewrite the real capacity in (16) as,

$$\begin{aligned} C_k^{real}(\mathbf{\Phi}) &= \sum_{i=1}^{3} W_k^{(F_i)} \vartheta^{(t(F_i,\mathsf{tech},Cell_{num}))}(\mathbf{\Phi}) \\ &= \sum_{i=1}^{3} \frac{W_{(F_i)}}{N_{user}^{(F_i)}} \vartheta^{(t(F_i,\mathsf{tech},Cell_{num}))}(\mathbf{\Phi}) \qquad ,(By\ \ Equal\ \ Sharing) \\ &= W_F \sum_{i=1}^{3} \frac{\vartheta^{(t(F_i,\mathsf{tech},Cell_{num}))}(\mathbf{\Phi})}{N_{user}^{(F_i)}} \quad ,(By\ , W_{(F_1)} = W_{(F_2)} = W_{(F_3)} = W_F) \end{aligned} \tag{6.8}$$

where $N_{user}^{(f)}$ is the number of users which can be supported by frequency band *f*, and $W_{(F_i)}$ is F_i bandwidth.

(B): This simplification is considered to avoid complexity and decrease the number of optimization variables. The primary contribution of this work lies in the proposed algorithm's minimal complexity. More specifically, we assume all center DRUs frequency band power, which is not assigned to the edge DRUs, are perturbed equally, either increasingly, or decreasingly.

Note that, before simplification, the number of optimization problem variable ($|\mathbf{\Phi}|$) is $19 \times (6+3)$, where the first term in the product is due to the fact that we have 19 cells, and the second term is because each cell of DAS-SFR-scheme1 has $(6+3)$ changeable user frequency band powers. These $(6+3)$ changeable user frequency band powers are comprised of 6 frequency band powers for the edge DRUs and 3 frequency band powers for central DRUs.

We decrease size of the optimization variable from $19 \times (6+3)$ to 1 in such a way that only the central DRUs frequency bands power, which are not assigned to the edge DRUs, are perturbed in

a similar fashion. The central DRUs F_2, F_3 power, which are not assigned to the edge DRUs, are perturbed equally for the central cell (eNB0) in a DAS-SFR-scheme1 configuration.

So the optimization problem (6.7) is simplified to,

$$\underset{\Delta P}{Maximize} \quad QoS(\Delta P) = \alpha KPI_{SU}(\Delta P) + \beta KPI_{CSU}(\Delta P) \tag{6.9}$$

where

$$KPI_{SU}(\Delta P) = \frac{\sum_{k=1}^{N_{users}} G_k(\Delta P)}{N_{users}}$$

$$KPI_{CSU}(\Delta P) = \sum_{k \in SUS} \sum_{i=1}^{M} \frac{\vartheta^{(t(F_i,\text{tech},Cell_{num}))}(\Delta P)}{N_{user}^{(F_i)}}$$

where,

$$G_k(\Delta P) = \begin{cases} 1 & when \quad \sum_{i=1}^{M} \frac{\vartheta^{(t(F_i,\text{tech},Cell_{num}))}(\Delta P)}{N_{user}^{(F_i)}} \geq \frac{C_{th}}{W_F} \\ 0 & otherwise \end{cases}$$

$$\Delta P(dB) = P'(dBm) - P(dBm)$$

- All cell's DRUs transmit the signal with power P (dBm) on one frequency part based on a reuse factor 3 in such a way that neighbor cell edge DRUs do not use the same frequency. And also each cell's central DRU transmits the signal with power $P'(dBm)$ on the other frequency parts which are not assigned to its 6 edge DRUs.

In our analysis, we assume that P is fixed (43 dBm) and only P' changes in magnitude in such a way that ΔP is altered between -10 to 20 dB. We also assume the α is much larger than the β , i.e., $\alpha = 10^6$ and $\beta = 10^{-6}$ in order to assign KPI_{SU} and KPI_{CSU} as the first and second priority to maximize QoS.

Following that, the optimization process is performed in an orderly basis by PSO algorithm. PSO algorithm initially maximize the number of satisfied users (KPI_{SU}), and at the event of reaching the same KPI_{SU} for different power level scenarios, the algorithm performs to maximize the achieved capacity of satisfied users (KPI_{CSU}).

6.5.2 The Power Self-Optimization (PSO) Algorithm.

According to the above analysis, we propose a PSO technique which is based on a simple and decentralized algorithm that runs on the application layer.

In the PSO algorithm, the expected network gain, which is comprised by one or both system KPIs, is used in order to determine whether to increase or decrease the transmission power of the central DRUs. Finally, the central processor which can be located beyond the eNBs is in charge of adjusting all central DRUs ΔP based on the calculated KPIs by performing the PSO algorithm in the central processor. Fig.6.15 depicts the block diagram of the self-optimization algorithm.

Observing the block diagram shown in Fig. 6.15, it is possible to note that the transmission power is adjusted by comparing the current KPI, calculated at the end of the current phase, and the last KPI, calculated at the end of last phase. Moreover, it is important to highlight that the central DRUs have a predefined minimum and maximum transmission power (P^{min} and P^{max}), which cannot be exceeded by the algorithm. Thus, the self-optimization algorithm increases or decreases the ΔP step-by-step by $p(dB)$ for each central DRUs. Parameter t can take two values, 1 and -1, where 1 shows that the algorithm starts by increasing the power level. Conversely, -1 indicates that the algorithm starts by decreasing the power level. The ΔP at the first iteration can be arbitrarily chosen from an interval of ΔP from -10 dB to 20 dB. Note that, we assume, ΔP is -10 dB and t is 1 at the first iteration.

Whenever the algorithm starts off by increasing the power level, the central DRUs increase ΔP by the fixed parameter p. The central DRU keeps increasing the power by the fixed parameters p as long as the current calculated KPI_{SU} is greater than the last calculated KPI_{SU}. If the current calculated KPI_{SU} is equal to the last calculated KPI_{SU}, the central DRUs keep increasing the power as long as the current calculated KPI_{CSU} is not smaller than the last calculated KPI_{CSU}; otherwise, it decreases its power level. Note that whenever the algorithm starts off by decreasing the power level, all the above mentioned statements should be reversed i.e., the increasing behavior should be changed to a decreasing behavior and vice versa.

We defined the parameter c to find and stop at the local optimal power. Since we do not want the power (ΔP) to oscillate around the local optimal power forever, choosing the parameter c equal to 2, helps the algorithm stop. $c = 2$ stops the algorithm if the convergence attempts to change the power monotonically (increasing or decreasing ΔP) for the second time. This aids in convergence on a local maximum.

The PSO algorithm seeks to maximize the number of satisfied users as the first priority. Meanwhile it seeks to maximize the capacity of the satisfied users as the second priority in order to have a better QoS. Even though it does not achieve a global optimal solution, it provides stable power updates toward the local optimal solution.

6.5.3 Analytical and Simulation Results.

Fig. 6.16 represents the ASE (Appendix A.3) of a cellular DAS's central cell for different frequency reuse techniques versus the normalized distance from the eNB0 DRU0 in the direction of the worst case position *X*, for a pathloss exponent of 3.76. Each scenario is plotted for the individual frequency in Fig. 6.16 (a) and also for the total ASE (ϑ_{total}) in Fig. 6.16 (b).

As it can be observed in Fig. 6.16 (a), when applying the SFR methods, by increasing from -10 dB to 20 dB, the ASE of central cell at frequency bands F_2 and F_3, increase. This, however, increases the interference associated with the edge DRUs of the neighboring cells which are using

F_2 and F_3 as their main frequency band. It is necessary to note that, increasing ΔP from -10 dB to 20 dB, significantly increases the associated interference with the F_1 frequency band in the central cell, imposed from the neighboring cells, and thus decreases ASE of central cell at frequency F_1.

It is important to notice that, considering SFR methods in Fig. 6.16 (b), as power increases, the total ASE does not change monotonically, which means the ASE associated with the cell's interior regions increases, and that of the cell's exterior regions decreases. Therefore, the users' distribution within the cell's area plays a significant role when deciding the optimal ΔP.

A secondary consideration when deciding the optimal ΔP is the minimum required capacity (C_{th}) to distinguish the satisfied users from unsatisfied users. As an example, with a high C_{th} ($\frac{C_{th}}{W_F} = \sum_{i=1}^{3} \vartheta^{t(F_i,\text{tech},Cell_{num})}$ = 14 bit/Hz, (Fig. 6.16 (b))), as ΔP increases, a wider radiance in a cell will be covered by ASE higher than 14 bit/Hz. With a low C_{th} ($\frac{C_{th}}{W_F} = \sum_{i=1}^{3} \vartheta^{t(F_i,\text{tech},Cell_{num})}$ = 3 bit/Hz, (Fig. 6.16 (b))), as ΔP increases, a shorter radiance in the cells will be covered by ergodic ASE than 3 bit/Hz.

The FFR (HFR1) method fully uses the frequency bands, therefore, the cell's interior region achieved an ASE higher than when applying the HFR3 method (HFR with frequency reuse factor 3) (Fig. 6.16 (b)). For example, at ϑ_{total}=14, the FFR method outperforms the HFR3 method, considering the users' satisfaction probability. However, when applying the FFR method, due to the interference caused by the neighboring cells, the edge cells frequently experience dead spots. As an example, considering the users' satisfaction probability, for ϑ_{total}l=3, the HFR3 method outperforms the FFR method.

In multi-users case, the user distributions are defined as depicted in Table 6.4 ,where,

$$N_{users} = N_A^d + N_B^d + N_C^d + N_D^d \tag{6.10}$$

$$\frac{1}{X_A^d}\frac{N_A^d}{S_A} = \frac{1}{X_B^d}\frac{N_B^d}{S_B} = \frac{1}{X_C^d}\frac{N_C^d}{S_C} = \frac{1}{X_D^d}\frac{N_D^d}{S_D}$$

where X_i^d is presented in Table 6.4, and N_{users} is total number of users, and N_i^d is the number of users in region i for distribution d, and S_i is the area of region *i* ($S_i = \pi((R_2^i)^2 - (R_1^i)^2)$), and $i \in \{Region-A, Region-B, Region-C, Region-D\}$ and $d \in \{UD, DCD, DCED, DED\}$. We perform Monte Carlo simulations to corroborate the analytical results. It is assumed that the total number of users (N_{users}) is 200.

Figures 6.17 and 6.18 demonstrate the two KPI_{SU} and KPI_{CSU} for different ΔP, considering four different user distributions. In multi-users case, since the 200 users (N_{users}) is assumed, the high C_{th}= 0.07W_F ($\frac{14}{200} = 0.07$) and the low C_{th}= 0.01W_F ($\frac{3}{200} = 0.01$) is considered for each user in (6.9). In theoretical analysis, the interior region is distinguished from the exterior region, based on T_P . In other words, the region with ASE higher than T_P is considered as interior region, and

	R_1	R_2	UD[(1)]	DCD[(2)]	DCED[(3)]	DED[(4)]
Region A	0	0.25	1	7	1	1
Region B	0.25	0.50	1	1	7	1
Region C	0.50	0.75	1	1	1	1
Region D	0.75	1	1	1	1	7

[(1)]UD : Uniform Distribution

[(2)]DCD: Dense at the Center Distribution

[(3)]DCED: Dense at the mid. of Cent. And Edge cell Distribution

[(4)]DED: Dense at the Edge cell Distribution

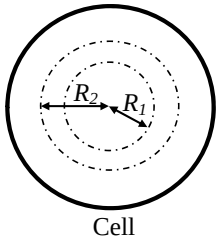


Table 6.4: $X_i^d, i \in \{A, B, C, D\}, d \in \{UD, DCD, DCED, DED\}$.

the region with ASE lower than T_P is considered as exterior region. This T_P is associated to the region's ASE of the frequency bands that are only allocated to the central DRUs, i.e., ASE of F_2 and F_3 are considered for central cell. We assume T_P=2 (*bit/Hz*) in our theoretical analyses.

Since both KPI functions are dependent on G_k , it is reasonable to consider each of these functions as a criterion to measure the QoS. We analyze two different cases, i.e., QoS = KPI_{SU} and QoS = KPI_{CSU}. In the first case, we presume KPI_{SU} as our QoS function whereas in the second case we consider KPI_{CSU} as our QoS function.

The evaluated KPI_{CSU} in figures 6.17 and 6.18 are presented as $\frac{KPI_{CSU}(\Delta P)}{\max\limits_{\Delta P} KPI_{CSU}}$ where $KPI_{CSU}(\Delta P)$ is the evaluated KPI_{CSU} at ΔP and $\max\limits_{\Delta P} KPI_{CSU}$ is the maximum value of evaluated KPI_{CSU} for all ΔP. The $\frac{KPI_{CSU}(\Delta P)}{\max\limits_{\Delta P} KPI_{CSU}}$ helps to demonstrate the relative degradation of each $KPI_{CSU}(\Delta P)$ comparing to $\max\limits_{\Delta P} KPI_{CSU}$.

As it is seen in Fig. 6.17, when C_{th} takes a low value, i.e., C_{th}=0.01W_F , except for the FFR method, applying the rest of the frequency reuse methods (HFR3, SFR) results in the highest number of the satisfied users (KPI_{SU}). Note that, as ΔP increases, when applying DAS-SFR-scheme1, the number of the satisfied users asymptotically decreases. The above mentioned results hold for all 4 different user distributions.

As was shown in Fig. 6.17, there exists an optimal ΔP at which the KPI_{CSU} is maximum, for all four different distributions. For instance, when applying the DAS-SFR-scheme1-Scenario3 method, for the UD (Uniform Distribution), DCD (Dense at the Center Distribution), DCED (Dense at the mid. of Cent. And Edge cell Distribution) and DED (Dense at the Edge cell Distribution), the maximum KPI_{CSU} happens at ΔP = -5 dB, 2 dB, 8 dB, and -4 dB, respectively.

One has to consider the optimal ΔP is different for dissimilar distribution scenarios. Moreover,

the DAS-SFR-scheme1-Scenario3 method outperforms the other two DAS-SFR-scheme1 methods, for all the distributions under consideration.

Fig. 6.18 reveals that, when C_{th} takes a large value, i.e., C_{th} =0.07 W_F , the FFR method outperforms the HFR3 method, considering the number of the satisfied users (KPI_{SU}). This corroborates our analytical results from Fig. 6.16, as it was explained previously. However, the DAS-SFR-Scenario2 method outperforms all the other methods, at different optimal ΔP values for dissimilar distribution scenarios.

As it can be perceived from Fig. 6.18, there exists an optimal ΔP at which the $KPI_{SU}$and KPI_{CSU} are maximum, for all four different distributions. As an example, when applying the DAS-SFR-scheme1-Scenario2 method, for the UD, DCD, DCED and DED, the maximum $KPI_{SU}$and KPI_{CSU} happen at ΔP = 6 dB, 4 dB, -2 dB, and 13 dB, respectively.

Note that, the optimal ΔP is different for different distribution scenarios. Moreover, the DAS-SFR-scheme1-Scenario2 method outperforms the other two SFR methods, for all distributions under consideration. Since the DAS-SFR-scheme1-Scenario2 uses all the frequency bands in the interior cell region, along with the fact that the users with throughput above the C_{th} are mainly located in the interior cell region, leads to the final conclusion that DAS-SFR-scheme1-Scenario2 outperforms the other methods.

Fig. 6.19 demonstrates two evaluated KPI_{SU} and KPI_{CSU} at each iteration using PSO algorithm for DAS-SFR-scheme1-Scenario3 when C_{th}=0.01W_F (Fig. 6.19(a)) and DAS-SFR-scheme1-Scenario2 when C_{th}=0.07W_F (Fig. 6.19(b)). The KPI_{SU} is presented as $KPI_{SU}\%$ which is $KPI_{SU} \times 100$. Note that, we assume, the ΔP is -10 dB at the first iteration. Fig. 6.19 shows the convergence of the algorithm to local optimal ΔP for different users distributions (UD, DCD, DCED and DED). In DAS-SFR-scheme1-Scenario3 when C_{th}=0.01W_F, the PSO algorithm is stopped at iteration 3, 7, 7 and 4 which are corresponding to = -6, 2, 2 and -4 dB for 4 different user distributions UD, DCD, DCED and DED, respectively. In DAS-SFR-scheme1-Scenario2 when C_{th}=0.07W_F, the PSO algorithm is stopped at iteration 9, 10, 5 and 16 which are corresponding to = 6, 8, -2 and 20 dB for 4 different user distributions UD, DCD, DCED and DED, respectively.

Fig. 6.20 represents the outage probability of different combination techniques in terms of CDF which is explained at section 6.3 when ΔP increases. The power control compensates are assumed for the lognormal shadowing in the central cell but the standard deviation of the lognormal shadowing for the signals from neighbor cells is 4 dB. The outage threshold (R_{th}) is assumed as 14.

6.6 Summary.

Two cell architecture DAS-SFR-scheme1-scheme1 and scheme2 that combine two known inter-cell interference mitigation techniques, DAS and SFR, are introduced in order to improve cell edge user's

throughput when the system has full spectral efficiency. Moreover, the simple bandwidth allocation techniques are investigated, which are specific to DAS-SFR architecture schemes. We have also provided comprehensive studies to compare the performance of our proposed methods with the other combinations of DAS and other frequency reuse techniques. The results from performance comparisons between different combinations lead to the conclusion that the most appropriate combination and bandwidth allocation scenario not only depend on achieved capacity inside the cell, but also on minimum guaranteed throughput.

The primary results show that DAS-SFR architectures effectively address inter-cell interference in a multi-cell environment, especially at the cell edges when compared to a DAS-HFR1 and Con-HFR1 architectures and achieves a non-trivial capacity increase over a DAS-HFR3 and Con-HFR3 architecture. More specifically, the results show that by controlling the amount of bandwidths allocated to users located in different areas, we can increase the frequency reuse and also improve the data rate for exterior users (users near to cell edge). If the throughput requirement for the interior users (users near to the cell center) is small, more bandwidths are allocated to the exterior users.

Then, a power self-optimization algorithm that aims at maximizing the number of satisfied users while trying to increase their capacity was also proposed. In more detail, the self-optimization algorithm uses the KPIs computed in the last phase and current phase to adjust the power level for the next phase.

Analytical and simulation results demonstrated the advantage of using the self-optimization algorithm instead of setting a fixed power level. When a DAS-SFR without the PSO algorithm is considered, the transmission power is set at the beginning of the communication and remains the same during its entire network lifetime. This characteristic can be negative considering a DAS-SFR in a real environment where the inherent user distribution is not constant.

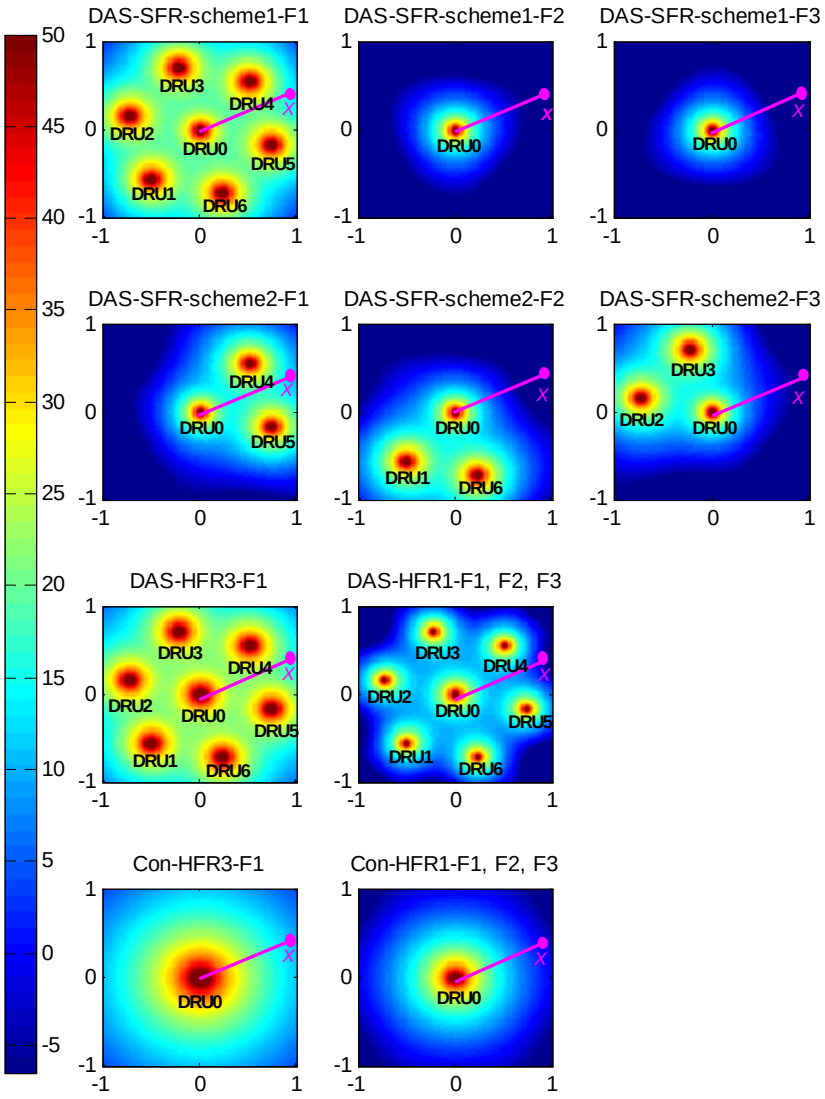


Figure 6.4: SINR map for individual frequency bands for all different transmission techniques.

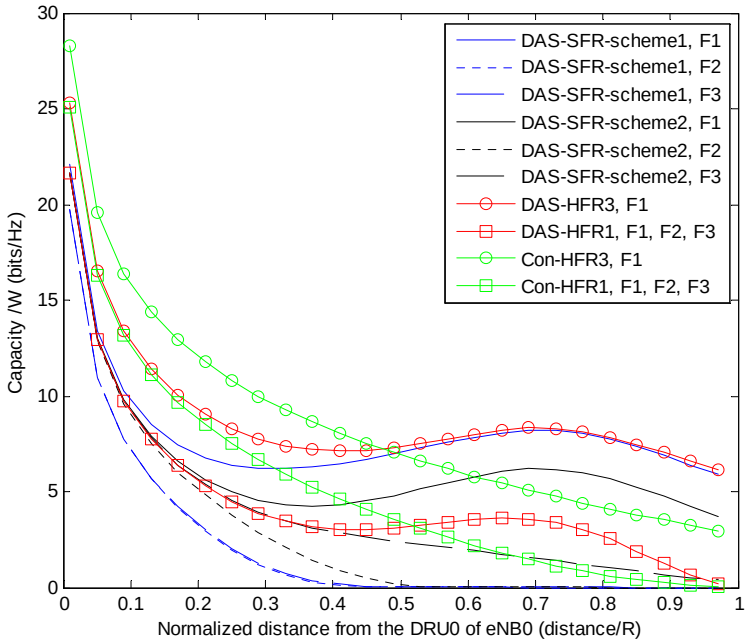


Figure 6.5: ASE versus the normalized distance the DRU0.

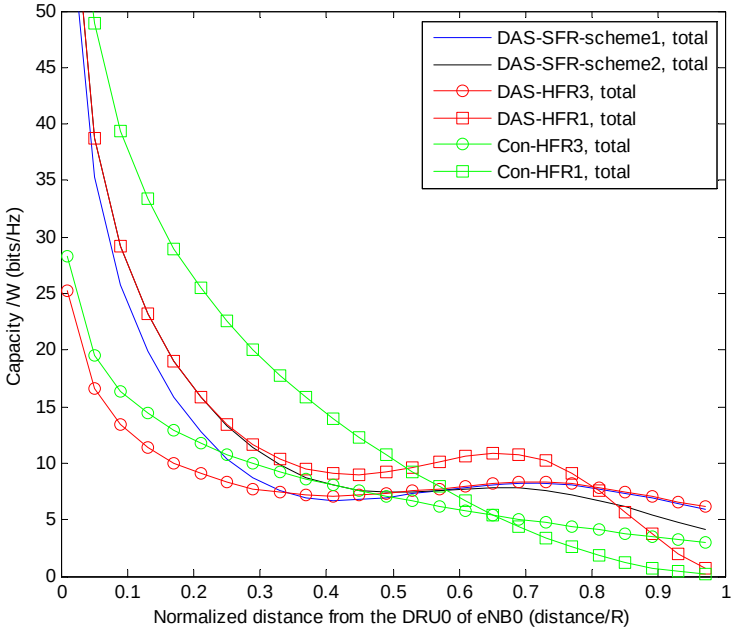


Figure 6.6: ASE versus the normalized distance the DRU0.

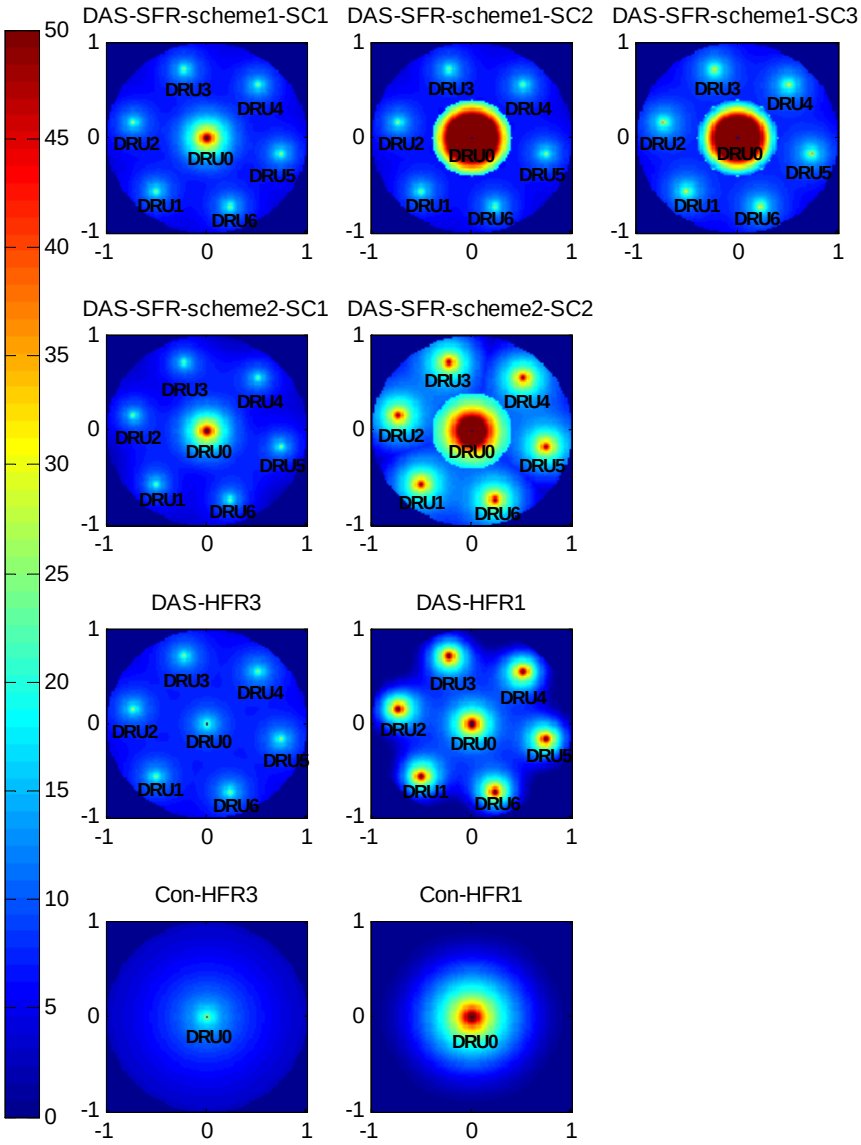


Figure 6.7: Regional capacity ($C^{(\text{region,tech})}\frac{N_{users}}{W}$) for multiuser case versus the normalized distance from the DRU0 in eNB0(central cell) area.

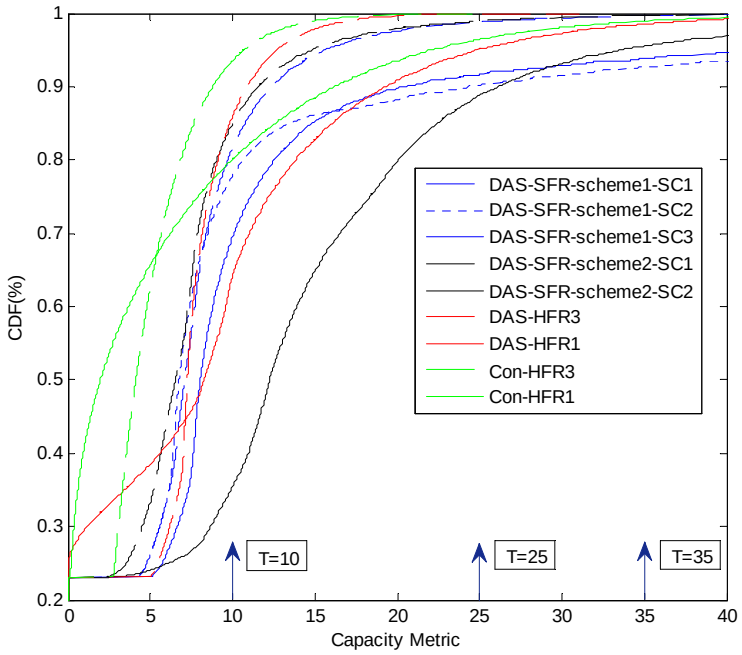


Figure 6.8: CDF of regional capacity($C^{(\text{region,tech})}\frac{N_{users}}{W}$) for multiuser case in eNB0(central cell) area when pathloss exponent is 3.76 and α=0.4.

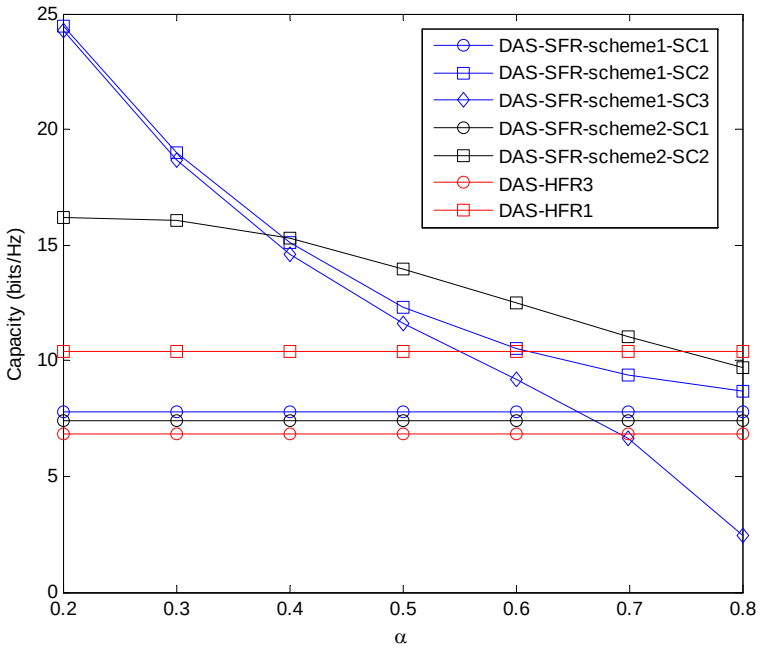


Figure 6.9: Average regional capacity ($C^{(\text{region,tech})}\frac{N_{users}}{W}$) for non-edge cell $(0, 0.8R)$ users versus α when pathloss exponent is 3.76.

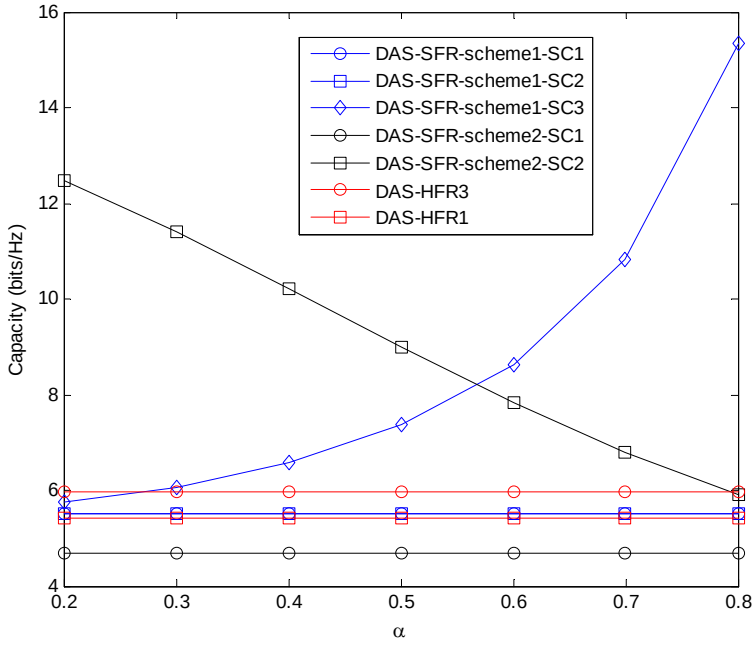


Figure 6.10: Average regional capacity ($C^{(\text{region,tech})}\frac{N_{users}}{W}$) for edge cell $(0.8R, R)$ users versus α when pathloss exponent is 3.76.

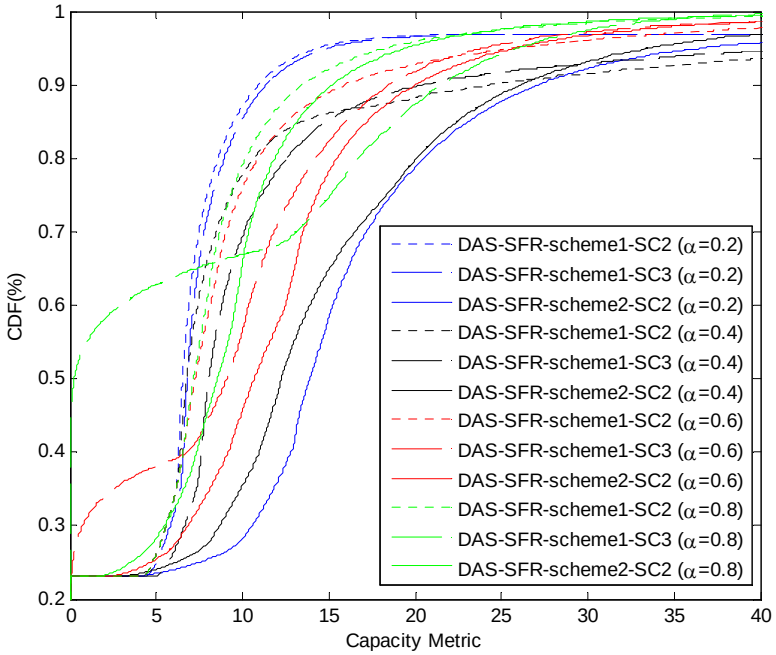


Figure 6.11: CDF of regional capacity($C^{(\text{region,tech})}\frac{N_{users}}{W}$) for multiuser case in eNB0(central cell) area for different α when pathloss exponent is 3.76.

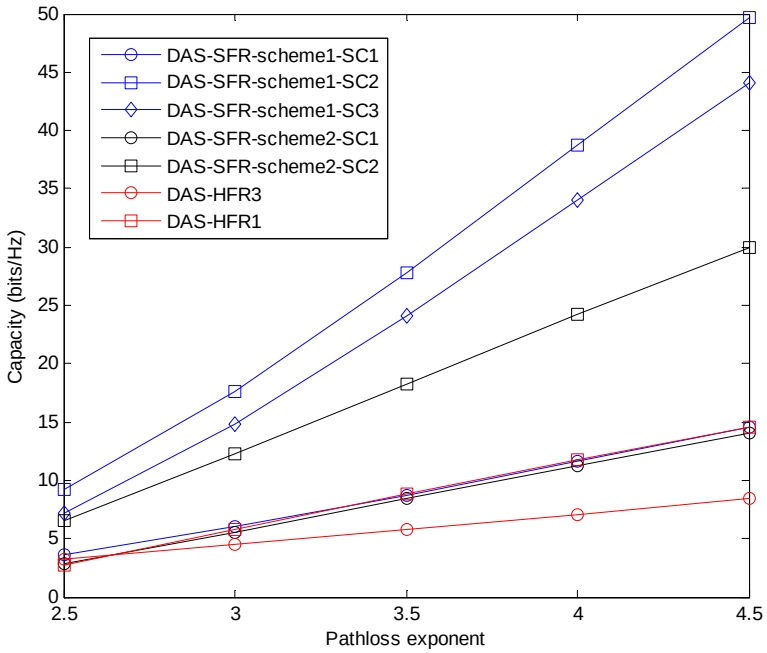


Figure 6.12: Average regional capacity ($C_{metric}^{(\text{region,tech})} \frac{N_{users}}{W}$) for edge cell $(0, 0.45R)$ users versus pathloss exponent when $alpha$=0.5.

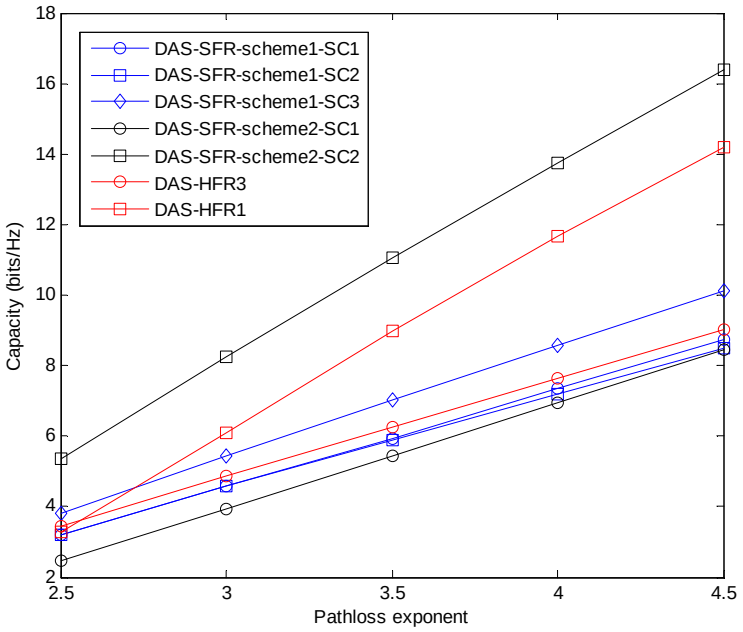


Figure 6.13: Average regional capacity ($C^{(\text{region,tech})}\frac{N_{users}}{W}$) for edge cell $(0.45R, 0.8R)$ users versus pathloss exponent when $alpha$=0.5.

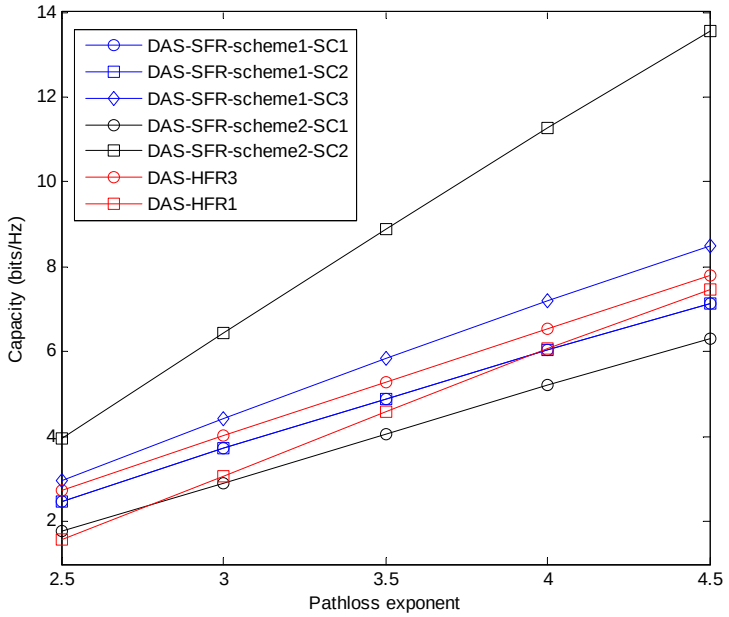


Figure 6.14: Average regional capacity ($C^{(\text{region,tech})}\frac{N_{users}}{W}$) for edge cell $(0.8R, R)$ users versus pathloss exponent when $alpha$=0.5.

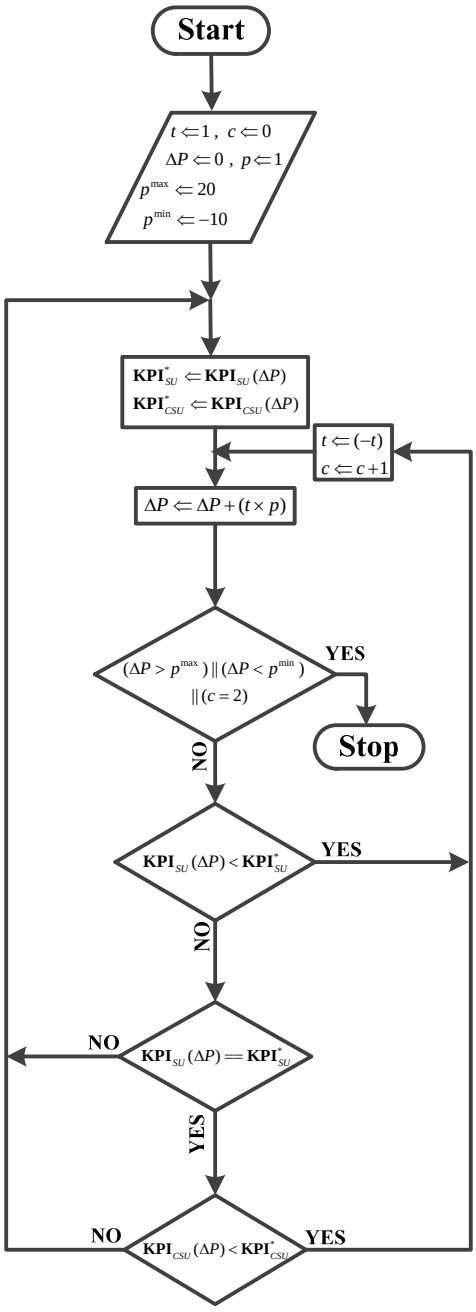
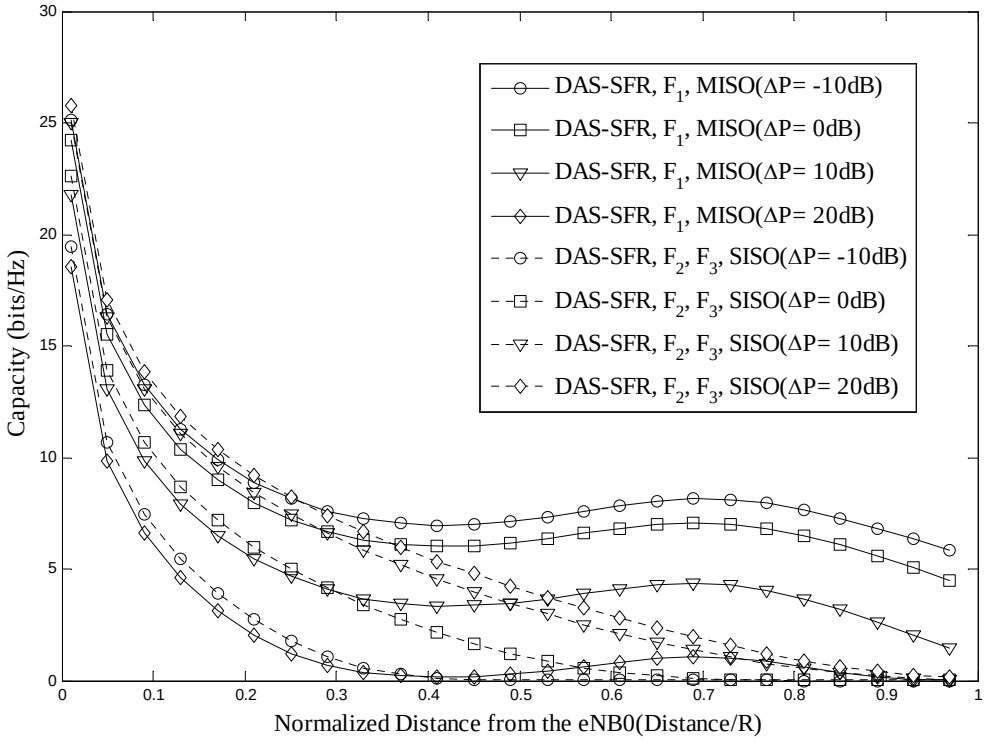
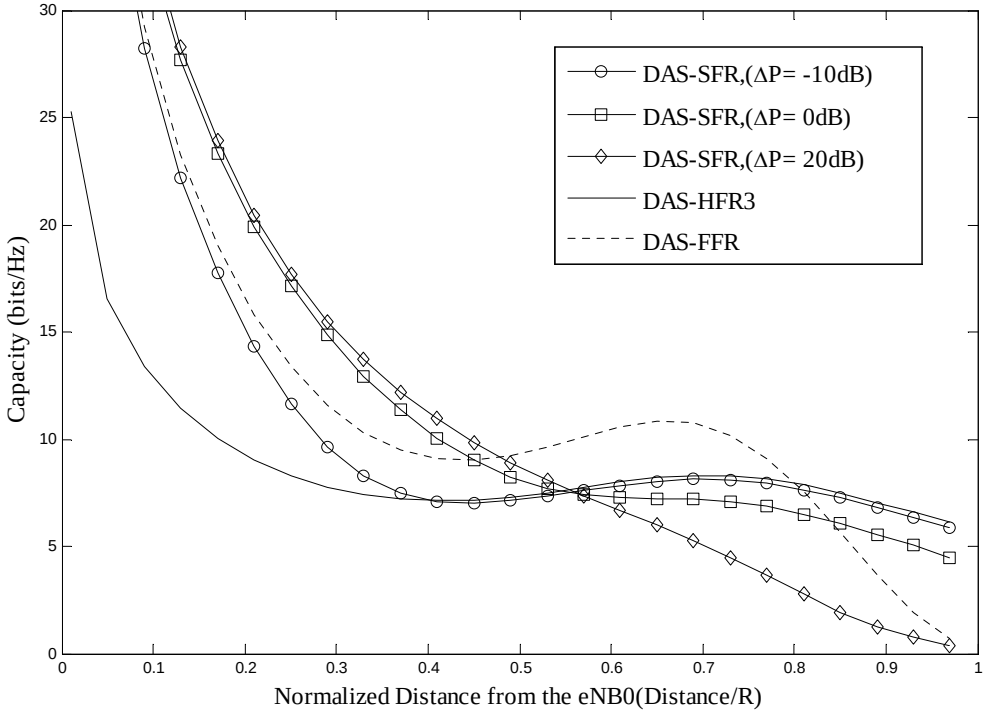


Figure 6.15: PSO Algorithm.



(a)



(b)

Figure 6.16: ASE versus the normalized distance the DRU0.

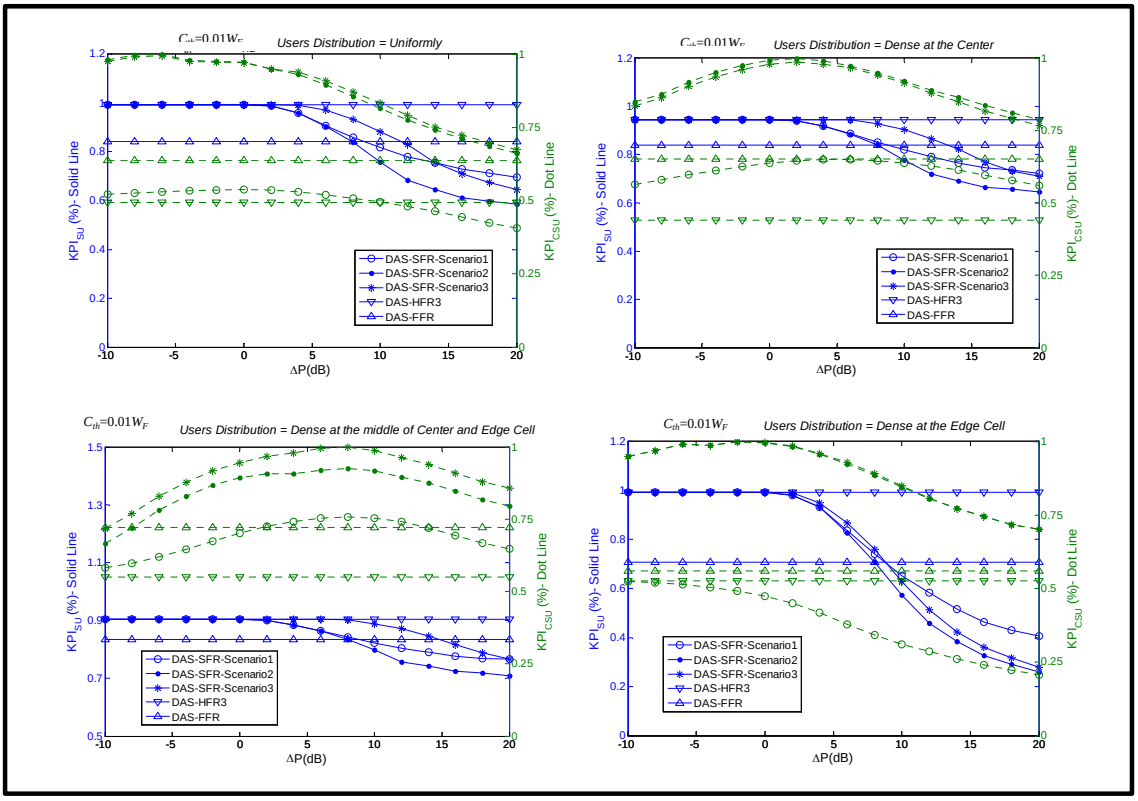


Figure 6.17: KPIs versus the ΔP for different distribution users scheme where $C_{th} = 0.01W_F$.

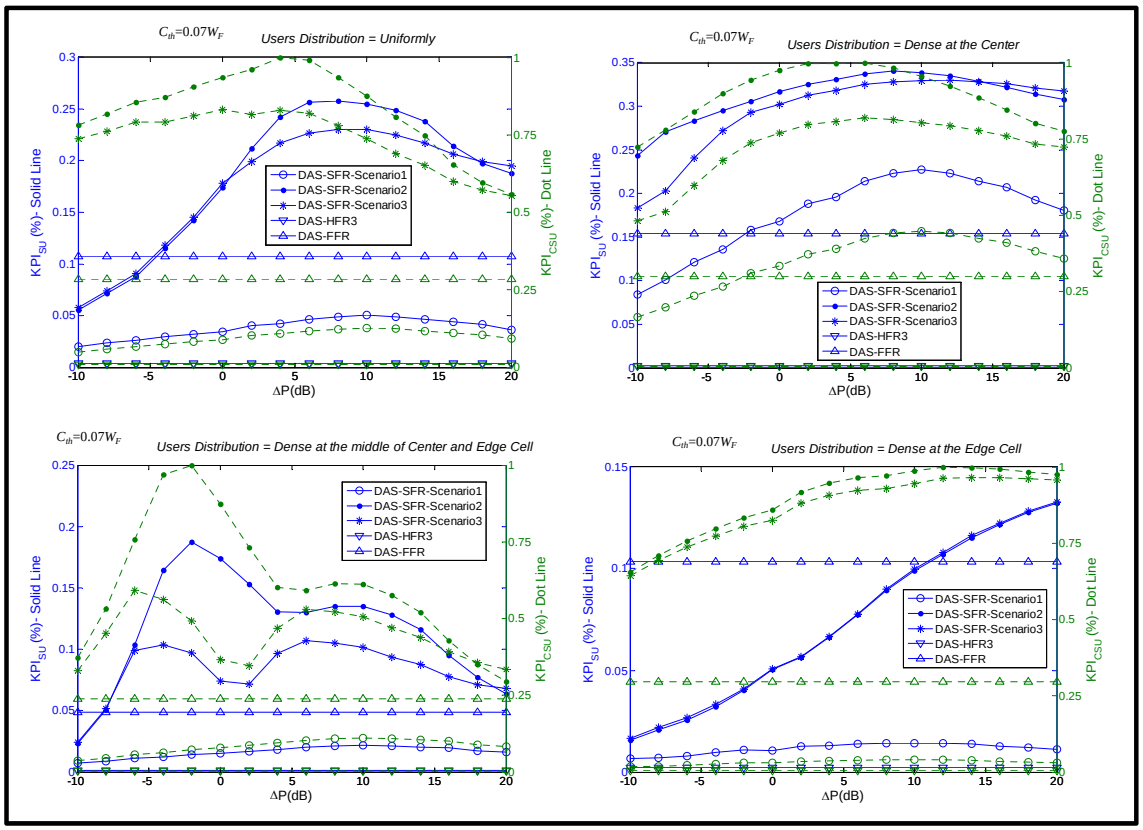


Figure 6.18: KPIs versus the ΔP for different distribution users scheme where $C_{th} = 0.07W_F$.

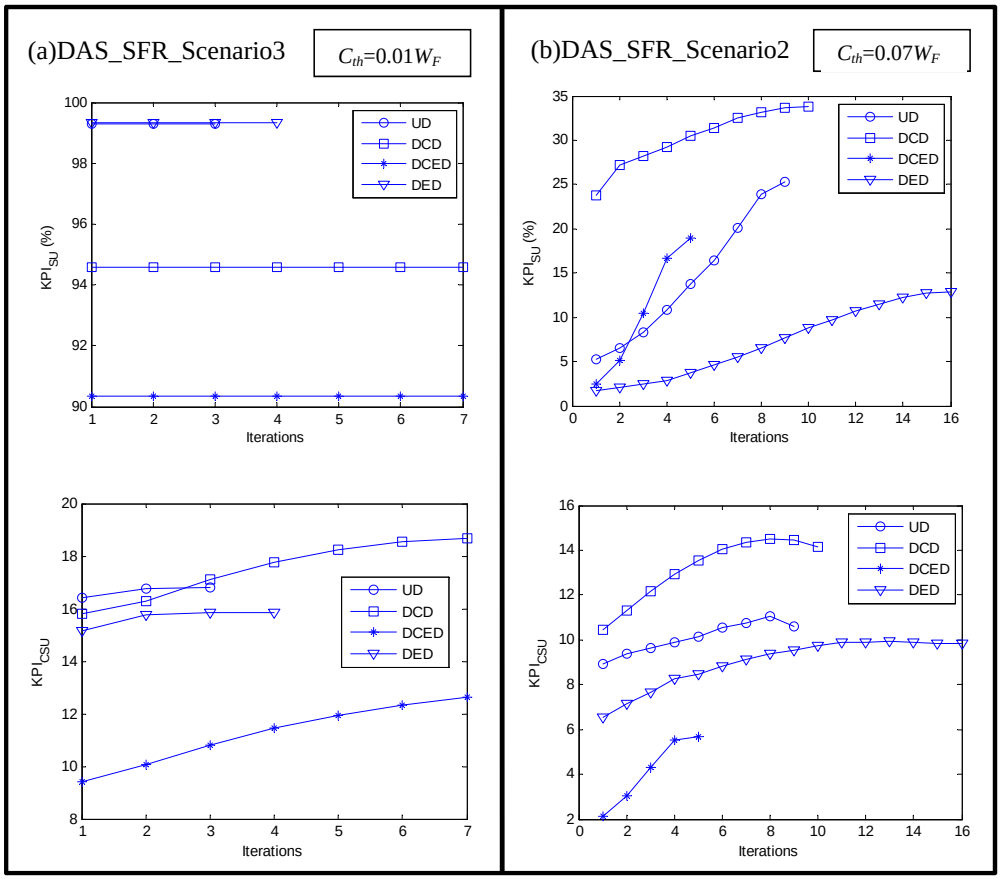


Figure 6.19: The convergence behavior of proposed PSO algorithm for two scenario.

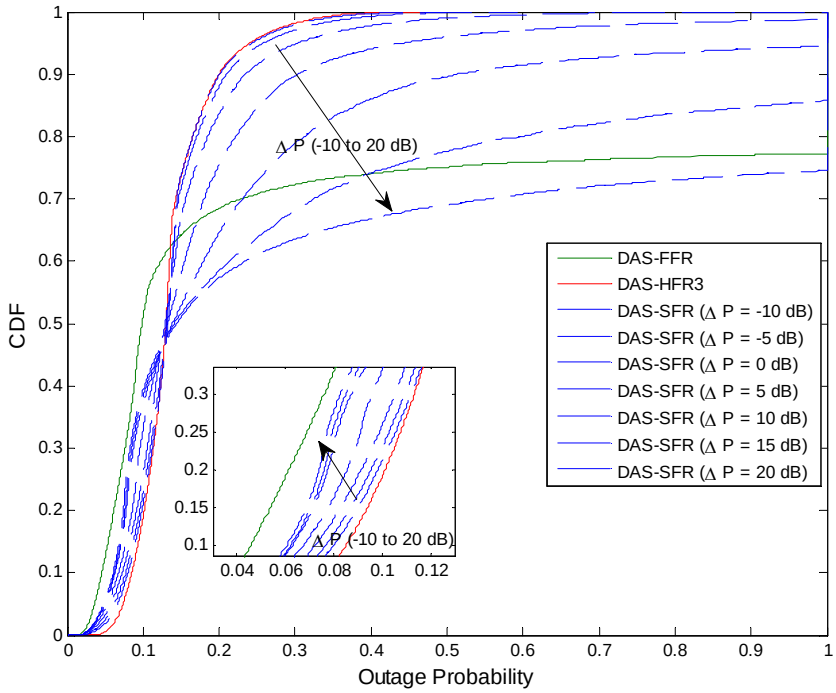


Figure 6.20: CDF of outage probability for different transmission techniques.

Chapter 7

Conclusion

In this thesis we studied how to employ distributed antenna system (DAS) in order to balance load in SON, and combine DAS with different frequency reuse techniques in order to mitigate inter-cell interference and contribute to the SON development for providing high QoS in a time-varying environment.

In particular, in chapter four, the performance analysis for a Virtual Cell distributed architecture and a Small Cell architecture was undertaken. The results indicate that the inter-cell interference from Small Cells have a significant impact on the user performance. Without any coordination between Small Cells, the SINR over the geographic area is significantly poorer. This results in the inability for the users to operate at the highest spectral efficiency available from LTE. The increased hand-offs from a poor SINR profile will increase the burden on the network and negatively impact the user experience.

Small Cells provide increased capacity to a given geographic area, whereas a traditional DAS architecture has been viewed as a coverage solution. However, a Virtual Cell distributed network that has the ability to provide capacity on demand by altering the simulcast ratio will, ultimately, provide the most optimal solution. A Virtual Cell distributed network provides the throughput demands at peak loads and maximizes the user experience throughout the entire coverage area. This investigation demonstrates that the throughput data rate of a Virtual Cell distributed network is the same as that of a Small Cell at a hotspot yet at a fraction of the required number of eNodeBs. The Virtual Cell distributed network architecture provides the best spectral efficiency and subsequently the highest useable capacity.

DRU allocation of eNB resources is examined in order to balance the traffic for a Virtual Cell network. DRU allocation of eNB resources is considered an effective technique for load balancing traffic on a network. Compact DRU allocations are proposed to reduce handoffs and interferences. DRU allocation of eNB resources is formulated as an integer linear programming problem, which minimizes a performance based cost function.

SON algorithms are developed to solve the DRU allocation of eNB resources. A significant throughput improvement is demonstrated using the SON algorithm.

In Chapter five, DRU allocation of antenna modules is examined with the objective of balancing the traffic in an LTE DAS. Two different load balancing schemes ("Block Probability-triggered Load Balancing" and "Utility Based Load Balancing" schemes) are investigated. In "Block Probability-triggered Load Balancing", proper DRU allocation is considered to be an effective use of resources in each sector that satisfies hard capacity limit and decreases the hand-off probability. EDA and GA algorithms are developed to solve the DRU allocation problem and their performances are compared with each other. The EDA algorithm produced a noticeably better and faster convergence than the GA algorithm and achieved the optimum performance.

In "Utility Based Load Balancing", proper DRU allocation is considered to be an effective use of resources in each eNB that prevents high blocking rate. The DRU allocation was formulated as an integer linear programming problem which maximizes the load balancing index and minimizes the average load as well as minimizing the number of handovers. The DPSO algorithm is developed to solve the DRU allocation problem and the performance is compared with an exhaustive search. Using two test problems, the DPSO algorithm produced a noticeably faster convergence than the exhaustive search algorithm.

Finally, in chapter six, two cell architecture DAS-SFR-scheme1 and scheme2 that combine two known inter-cell interference mitigation techniques, DAS and SFR, are introduced in order to improve cell edge user's throughput when the system has full spectral efficiency. Moreover, the simple bandwidth allocation techniques are investigated, which are specific to DAS-SFR architecture schemes. We have also provided comprehensive studies to compare the performance of our proposed methods with the other combinations of DAS and other frequency reuse techniques. The results from performance comparisons between different combinations lead to the conclusion that the most appropriate combination and bandwidth allocation scenario not only depend on achieved capacity inside the cell, but also on minimum guaranteed throughput.

The primary results show that DAS-SFR architectures effectively address inter-cell interference in a multi-cell environment, especially at the cell edges when compared to a DAS-HFR1 and Con-HFR1 architectures and achieves a non-trivial capacity increase over a DAS-HFR3 and Con-HFR3 architecture. More specifically, the results show that by controlling the amount of bandwidths allocated to users located in different areas, we can increase the frequency reuse and also improve the data rate for exterior users (users near to cell edge). If the throughput requirement for the interior users (users near to the cell center) is small, more bandwidths are allocated to the exterior users.

Then, a power self-optimization algorithm that aims at maximizing the number of satisfied users while trying to increase their capacity was also proposed. In more detail, the self-optimization algorithm uses the KPIs computed in the last phase and current phase to adjust the power level for the next phase.

Analytical and simulation results demonstrated the advantage of using the self-optimization algorithm instead of setting a fixed power level. When a DAS-SFR without the PSO algorithm is considered, the transmission power is set at the beginning of the communication and remains the same during its entire network lifetime. This characteristic can be negative considering a DAS-SFR in a real environment where the inherent user distribution is not constant.

Bibliography

[1] LTE-A, “Requirements for Further Advancements for EUTRA,” 3GPP TR 36.913, 2008.

[2] LTE-A, “Physical layer aspects for evolved Universal Terrestrial Radio Access (UTRA) (Release 7),” 3GPP TR 25.814 ver 7.1.0, 2006.

[3] R. Giuliano, C. Monti, and P. Loreti, “Wireless technologies advances for emergency and rural communications - WiMAX fractional frequency reuse for rural environments,” in *IEEE Wireless Communications*, 15(3), 60-65, 2008.

[4] S. Hamalainen,“Self-Organizing Networks in 3GPP LTE” , *VTC* 2009-Fall ,pp: 1 2, 2009

[5] The SOCRATES Project, EB/OL. ht.tp:/./w. w.w.fp7-socrates.eu 2010

[6] 3GPP TS 32.500. “Self-Organizing Networks (SON) Concepts and requirements” 2009

[7] 3GPP TR 36.902.“Self-configuring and self-optimizing network use cases and solutions” 2009

[8] 3GPP TR 32.816. “Study on Management of Evolved Universal Terrestrial Radio Access Network and Evolved Packet Core” .2008.

[9] I. Balan, “Self-optimization in 3GPP LTE networks,” *FirW PhD Symposium*, Ghent, Belgium, December 2009.

[10] 3GPP, RI-050764, “ Inter-cell interference handling for E-UTRA,” *Ericsson Inc.*.

[11] R. Bosisio and U. Spagnolini, “Interference coordination vs. interference randomization in multicell 3GPP LTE system,” *IEEE WCNC’08*, pp. 824-829, April 2008.

[12] W. Choi, J. Andrews, “Downlink performance and capacity of distributed antenna system in multi-cell environment ,” *IEEE Transaction on Wireless Communication*,vol. 6, no. 1, Jan. 2007, pp. 1676-1680.

[13] Z. Wang, R.A. Stirling-Gallacher, “Frequency reuse scheme for cellular OFDM systems,” *Electronics Letters*, 38(8), 387-388 (2002)

[14] David Gonzalez G, Mario Garcia-Lozano, Silvia Ruiz Boque, Dong Seop Lee, "Optimization of Soft Frequency Reuse for Irregular LTE Macrocellular Networks" *IEEE Transaction on Wireless Communication*, vol. 12, no. 5, pp. 2410-2423, 2013.

[15] Huawei, "Soft frequency reuse scheme for UTRAN LTE," 3GPP, Huawei R1-050507, May 2005.

[16] NEC Corporation, "Self Organizing Networks - NEC's proposals for next generation radio network management," White Paper, Feb. 2009.

[17] O. K. Tonguz, and E. Yanmaz, "The mathematical theory of dynamic load balancing in cellular networks," *IEEE Transaction on Mobile Computing*, vol. 7, no. 12, pp.1504-1518, Dec 2008.

[18] B. Eklundh, "Channel utilization and blocking probability in a cellular mobile telephone system with directed retry," *IEEE Transaction on Communication*, vol. 34, no. 3, pp. 329-337, Apr. 1986.

[19] H. Jiang, and S. S. Rappaport, "CBWL: A new channel assignment and sharing method for cellular communication systems," *IEEE Transaction on Vehicular Technology*, vol. 43, no. 4, pp. 313-322, May 1994.

[20] S. Das, H. Viswanathan, and G. Rittenhouse, "Dynamic load balancing through coordinated scheduling in packet data systems," *IEEE, INFOCOM*, 2003.

[21] T. Bu, L. Li, and R. Ramjee, "Generalized proportional fair scheduling in third generation wireless data networks," *IEEE, INFOCOM*, Apr. 2006.

[22] K. Son, S. Chong, and G. Veciana, "Dynamic association for load balancing and interference avoidance in multi-cell networks," *IEEE Transaction on Wireless Communication*, vol. 8, no. 7, pp. 3566-3576, Jul. 2009.

[23] T. S. Rappaport, *Wireless communications : principles and practice*. Upper Saddle River, N.J., London: Prentice Hall PTR, Prentice-Hall International, 1996.

[24] A. Lindsay-Stewart, W. C. Y. Lee, M. A. Schulz, and C. Xu, "Incremental capacity gains for high blocking sites using dynamic channel sharing," *IEEE Transactions on Vehicular Technology*, vol. 50, no. 1, pp. 1-11, January 2001.

[25] Y. Argyropoulos, S. Jordan, and S. P. R. Kumar, "Dynamic channel allocation in interference-limited cellular systems with uneven traffic distribution," *IEEE Transactions on Vehicular Technology*, vol. 48, no. 1, pp. 224-232, January 1999.

[26] F. Delli Priscoli, N. P. Magnani, V. Palestini, and F. Sestini, "Application of dynamic channel allocation strategies to the GSM cellular network," *IEEE Journal on Selected Areas in Communications*, vol. 15, pp. 1558- 1567, October 1997.

[27] Andreas Lobinger, et al. “Load balancing in downlink LTE self-optimizing networks.” *Vehicular Technology Conference (VTC 2010-Spring)*, 2010 IEEE 71st. IEEE, 2010.

[28] Hao He, et al. “Game theory based load balancing in self-optimizing wireless networks.” *Computer and Automation Engineering (ICCAE)*, 2010 The 2nd International Conference on. Vol. 4. IEEE, 2010.

[29] J. S. Wu, J. K. Chung, and C. C. Wen, “Hot-spot traffic relief with a tilted antenna in CDMA cellular networks,” *IEEE Transactions on Vehicular Technology*, vol. 47, no. 1, pp. 1-9, Febuary 1998.

[30] A. Jalali, “On cell breathing in CDMA networks,” in *The Proceedings of IEEE ICC 98*, vol. 2, June 1998, pp. 985-988.

[31] W. Roh and A. Paulraj, “Outage performance of the distributed antenna systems in a composite fading channel,”in Proc., *IEEE Vehicular Technology Conference*, Sept. 2002, pp. 1520-1524.

[32] H. Fujii and H. Yoshino, “Theoretical capacity and outage rate of OFDMA cellular system with fractional frequency reuse,” in Proc.*IEEE Vehicular Technology Conference*, May 2008, pp. 1676-1680.

[33] W. Choi, J. Andrews, “Downlink performance and capacity of distributed antenna system in multi-cell environment ,” *IEEE Transaction on Wireless Communication*,vol. 6, no. 1, Jan. 2007, pp. 1676-1680.

[34] R. Chang, Z. Tao, J. Zhang, and C. Kuo, “A graph approach to dynamic fractional frequency reuse (FFR) in multi-cell OFDMA networks,” in Proc., *IEEE International Conference on Communications*, June 2009, pp. 1-6.

[35] M. Assad, “Optimal fractional frequency reuse (FFR) in multicellular OFDMA system,” in Proc. *IEEE Vechicular Technology Conference*,Sep. 2008, pp. 1-5.

[36] Huiling Zhu, “On frequency reuse in cooperative distributed antenna systems.” in *Communications Magazine, IEEE* , no. 4, pp. 85-89, 2012.

[37] Zhang Rong, Li Guangjun; L. Hanzo, “Distributed Antenna Systems in Fractional-Frequency-Reuse-Aided Cellular Networks,” in Proc., *IEEE Transaction Vehicular Technology*, vol.62, pp. 1340-1349, March 2013.

[38] Castaneda-Trujillo, Eduardo Daniel, Ramiro Samano-Robles, and Atilio Gameiro.“Frequency-reuse planning of the down-link of distributed antenna systems with maximum-ratio-combining (MRC) receivers.” *Latin America Transactions, IEEE (Revista IEEE America Latina)*, no. 3, pp. 1703-1709, Sept. 2012.

[39] L. Chen and D. Yuan, “Generalizing FFR by flexible sub-band allocation in OFDMA networks with irregular cell layout,” in Proc. *IEEE Wireless Communications and Networking Conference*, Apr. 2010, pp. 1-5.

[40] A. Ghosh, J. Zhang, J. G. Andrews, and R. Muhamed, *Fundamentals of LTE.*, Prentice Hall, 2010.

[41] M. Clark, T. Willis III, L. Greenstein, A. Rustako Jr, V. Erceg, and R. Roman, “Distributed versus centralized antenna arrays in broadband wireless networks,” *Vehicular Technology Conference, IEEE VTS 53rd*, vol. 1. IEEE, 2001, pp. 33-37.

[42] Z. Cao, U. Tureli, Y. D. Yao,“low-complexity orthogonal spectral signal construction for generalized OFDMA uplink frequency synchronization,” in *IEEE Transaction on Vehicular Technology*, on 56, no. 3 (2007): pp. 1143-1154.

[43] Koyluoglu, O. Ozan, and Hesham El Gamal, “Cooperative Encoding with Partially Known Non-Causal Interference,” in *IEEE Transaction on Information Theory*, on 57, no. 9 (2011): pp. 5682-5694.

[44] Seyed Amin Hejazi, and Shawn P. Stapleton. “Traffic Monitoring in a LTE Distributed Antenna System.” *Journal of Selected Areas in Telecommunications*, p. 19-25, May 2013.

[45] I. Cotains, “ Aspects of Cost Effective Network Performance Management,” in *The Proceedings of the 8th World Wireless Congress(WWC’2007)*, San Francisco, CA, May 2007, pp. 91-94.

[46] G. L. Nemhauser and L. A. Wolsey, *Integer and Combinatiorial Optimization*, John Wiley and Sons, Inc., 1988.

[47] M. Mitchell, *An Introduction to Genetic Algorithms*. MIT Press, 1996.

[48] A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, NY, USA, 1965.

[49] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products.* UK: Academic Press, 2003.

[50] P. Larranaga and J. A Lozano, *Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation*. Kluwer Academic Publishers, 2001.

[51] R. Kreher and K. Gaenger, *LTE Signaling, Troubleshooting and Optimization*, John Wiley and Sons, Inc., 2011.

[52] H. Jihai and W. Bingyang, “Handover in the 3GPP long term evolution (LTE) systems” *Mobile Congress (GMC)*, pp. 1-6, Febuary 2010.

[53] Sesia Stefania, Issam Toufik, and Matthew Baker, *LTE: the UMTS long term evolution*, New York: John Wiley Sons, 2009.

[54] Erik Dahlman, Stefan Parkvall, and Johan Skold, *4G: LTE/LTE-Advanced for Mobile Broadband*, Academic Press, 2011.

[55] S. Schwartz and Y. Yeh, "On the distribution function and moments of power sums with lognormal components," *Bell Syst. Tech. Jour.*, vol. 61, no. 7, pp. 14411462, Sept. 1982.

[56] H. Jihai and W. Bingyang, "Handover in the 3GPP long term evolution (LTE) systems" *Mobile Congress (GMC)*, pp. 1-6, Febuary 2010.

[57] Dah-Ming Chiu, and Raj Jain, "Analysis of the increase and decrease algorithms for congestion avoidance in computer networks", *Computer Networks and ISDN systems* 17.1 pp: 1-14, (1989)

[58] A. Coello Carlos, Gregorio Toscano Pulido, and Maximino Salazar Lechuga, "Handling multiple objectives with particle swarm optimization", *IEEE Transactions on Evolutionary Computation*, Vol 8, Issue 3, June , pp:256 279, (2004).

[59] J. Kennedy and R.C. Eberhart , "Particle Swarm Optimization", *IEEE Conference on Neural Networks*, Piscataway, NJ, USA, pp. 1942-1948, (1995).

[60] Quan-Ke Pan, M. Fatih Tasgetiren, and Yun-Chia Liang, "A discrete particle swarm optimization algorithm for the no-wait flowshop scheduling problem", *Computers and Operations Research* 35.9, pp: 2807-2839, (2008)

[61] I. Rivas, L.B. Lopes,"Transmitter macrodiversity in radio fibre microcellular networks", *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, PIMRC, vol. 3, pp. 1074 - 1078, 1997.

[62] Zhu Huiling, "Performance Comparison between Distributed Antenna and Microcellular Systems", *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 6, pp. 1151-1163, June 2011.

[63] Tania Villa, Ruben Merz, Pablo Vidales, "Performance evaluation of OFDMA femtocells link-layers in uncontrolled deployments", *Wireless Conference (EW)*, 2010 European, pp. 825 -832, 2010.

[64] D. Castanheira and A. Gameiro, "Distributed antenna system capacity scaling [coordinated and distributed mimo]", *IEEE Wireless Communication Mag.*, vol. 17, no. 3, pp. 6875, 2010.

[65] Yi Wu, D. Zhang, H. Jiang and Ye Wu,"A Novel Spectrum Arrangement Schemes for Femto Cell Deployment in LTE Macro Cells", *Personal, Indoor and Mobile Radio Communications*, 20th International Symposium, 2009.

[66] liu Zhen, et al., "A site-specific study of in-building wireless solutions", *IEEE Vehicular Technology Conference*,VTC 2010 spring.

[67] T. Ahmad, S. Al-Ahmadi, H. Yanikomeroglu, and G. Boudreau, "Downlink linear transmission schemes in a single-cell distributed antenna system with port selection," *IEEE Vehicular Technology Conference*. VTC Spring, Budapest, Hungary, 15-18 May 2011, pp. 15.

[68] S. Sesia, Issam Toufik, M. Baker *LTE - The UMTS Long Term Evolution: From Theory to Practice*. A John Wiley and Sons, Ltd 6 (2009): 136-144.

[69] *LTE/SAE Techonology Product Roadmap*, White Paper, Huawey Techonologies, Sep. 2007.

[70] 3GPP, Rl-050738, "Interference mitigation considerations and results on frequency reuse," Siemens Inc.

[71] 3GPP, Rl-050833, "Interference mitigation in evolved UTRA/UTRAN," LG Electronics Inc.

[72] J. Andrews, W. Choi, and R. Heath, "Overcoming interference in spatial multiplexing mimo cellular networks," *IEEE Wireless Comm.*, vol. 14, pp. 95-104, 2007.

[73] G. Boudreau, J. Panicker, N. Guo, R. Chang, N. Wang, and S. Vrzic, "Interference coordination and cancellation for 4g networks," *IEEE Comm. Magazine*, vol. 47, pp. 74-81, April 2009.

[74] 3GPP, TR 25.814, "Physical layer aspects for evolved universal terrestrial radio access (UTRA)," www.3gpp.org.

[75] R. W. Heath, Jr., S. Peters, Y. Wang, and J. Zhang, "A current perspective on distributed antenna systems for the downlink of cellular systems," *IEEE Communications Magazine*, vol. 51, no. 4, pp. 161167, April 2013.

[76] L. Dai, S. Zhou, and Y. Yao, "Capacity analysis in CDMA distributed antenna systems," *IEEE Transactions on Wireless Communications*, vol. 4, no. 6, pp. 26132620, 2005.

[77] P. Chow, A. Karim, V. Fung, and C. Dietrich, "Performance advantages of distributed antennas in indoor wireless communication systems," *IEEE Vehicular Technology Conference*, pp. 15221526, 1994.

[78] H. Yanikomeroglu and E. Sousa, "CDMA distributed antenna system for indoor wireless communications," *IEEE Intl. Conf. on Universal Personal Comm.*, vol. 2, pp. 990994, 1993.

[79] G. Chen, C. Yu, and C. Huang, "A simulation study of a distributed antenna-based CDMA system," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 2, pp. 517521, 1996.

[80] A. Obaid and H. Yanikomeroglu, “Reverse-link power control in CDMA distributed antenna systems,” *IEEE Wireless Communications and Networking Conf.*, vol. 2, pp. 608612, 2000.

[81] K. Kerpez, “A radio access system with distributed antennas,” *IEEE Trans. on Info. Theory*, vol. 45, no. 2, pp. 265275, 1996.

[82] M. Clark, T. Willis III, L. Greenstein, A. Rustako Jr, V. Erceg, and R. Roman, “Distributed versus centralized antenna arrays in broadband wireless networks,” *IEEE Vehicular Technology Conference*, pp. 3337, 2001.

[83] W. Choi and J. Andrews,“Downlink performance and capacity of distributed antenna systems in a multicell environment,” *IEEE Trans. on Communications*, vol. 6, no. 1, pp. 6973, 2007.

[84] H. Zhuang, L. Dai, L. Xiao, and Y. Yao, “Spectral efficiency of distributed antenna system with random antenna layout,” *Electronics Letters*, vol. 39, no. 6, pp. 495496, 2003.

[85] I. Toufik and R. Knopp, “Wideband channel allocation in distributed antenna systems,” *IEEE 64th Vehicular Technology Conference* (VTC Fall), pp. 15, 2006.

[86] H. Zhu, S. Karachontzitis, and D. Toumpakaris, “Low-complexity resource allocation and its application to distributed antenna systems,” *IEEE Wireless Communications*, vol. 17, no. 3, pp. 4450, 2010.

[87] X. Wang, P. Zhu, and M. Chen, “Antenna location design for generalized distributed antenna systems,” *IEEE Communications Letters*, vol. 13, no. 5, pp. 315317, 2009.

[88] E. Park, S.-R. Lee, and I. Lee, “Antenna placement optimization for distributed antenna systems,” *IEEE Transactions on Wireless Communications*, vol. 11, no. 7, pp. 24682477, 2012.

[89] E. Park and I. Lee, “Antenna placement for downlink distributed antenna systems with selection transmission,” *IEEE 73rd Vehicular Technology Conference (VTC Spring)*, pp. 15, 2011.

[90] L. Xiao, L. Dai, H. Zhuang, S. Zhou, and Y. Yao, “Information-theoretic capacity analysis in MIMO distributed antenna systems,”*57th IEEE Vehicular Technology Conference (VTC Spring)*, vol. 1, pp. 779782, 2003.

[91] H. Dai, “Distributed versus co-located MIMO systems with correlated fading and shadowing,” *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 4, pp. 561 564, 2006.

[92] H. Dai, H. Zhang, and Q. Zhou, “Some analysis in distributed MIMO systems,” *IEEE Trans. Commun.*, vol. 2, no. 3, pp. 4350, 2007.

[93] W. Roh and A. Paulraj, "Performance of the distributed antenna systems in a multi-cell environment," *IEEE Vehicular Technology Conference*, vol. 1, pp. 587591, 2003.

[94] W. Feng, Y. Li, J. Gan, S. Zhou, J. Wang, and M. Xia, "On the size of antenna cluster in multi-user distributed antenna systems," *IEEE CHINACOM*, pp. 11011105, 2008.

[95] J. Gan, Y. Li, S. Zhou, and J. Wang, "On sum rate of multi-user distributed antenna system with circular antenna layout," *66th IEEE Vehicular Technology Conference (VTC Fall)*, pp. 596600, 2007.

[96] L. Dai, "Distributed antenna system: Performance analysis in multi-user scenario," *42nd Annual Conference on Information Sciences and Systems*, pp. 8589, 2008.

[97] W. Feng, Y. Li, J. Gan, S. Zhou, J. Wang, and M. Xia, "On the deployment of antenna elements in generalized multi-user distributed antenna systems,"*Mobile Networks and Applications*, vol. 16, no. 1, pp. 3545, 2011.

[98] R. W. Heath, Jr., T. Wu, Y. H. Kwon, and A. Soong, "Multiuser MIMO in Distributed Antenna Systems With Out-of-Cell Interference," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 48854899, 2011.

[99] B. Song, R. L. Cruz, and B. D. Rao, "Downlink optimization of indoor wireless networks using multiple antenna systems," *Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, vol. 4, pp. 27782789, 2004.

[100] T. Wu and P. Hosein, "Radio resource management strategies for distributed antenna systems," *IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 16, 2010.

[101] Sara Mendes Duarte, "Analysis of Technologies for Long Term Evolution in UMTS", *Master of Science thesis*, Technical University of Lisbon, 2008.

Appendix A

Received Signal, Outage Probability and Average Spectral Efficiency of DAS

The downlink path of a DAS can be considered as an equivalent MISO system with additive interference and noise. The received signal at the location in the central cell of DAS shown at Fig. 6.1, can be expressed as,

$$\begin{aligned} y^{(0)} &= \text{signal+ interference + noise} \\ &= \sum_{n \in SoS} h_n^{(0)} x_n^{(0)} + \sum_{i=1}^{18} \sum_{n \in SoS} h_n^{(i)} x_n^{(i)} + n^{(0)} \end{aligned} \tag{A.1}$$

where $h_n^{(i)} \in C$, $i = 0, 1, ..., 18$, denotes the channel between the n-th DRU in the *i*-th cell and the user in the central cell, $x_n^{(i)} \in C$, $i = 0, 1, ..., 18$ is the transmitted signal of the n-th DRU in the *i*-th cell, $n^{(0)} \in C$ denotes the additive Gaussian noise with variance $\sigma^2_{n^{(0)}}$ and SoS represents the set of DRUs with transmit power greater than zero (Table 6.1). The transmit signal and its power of the *n*-th DRU of *i*-th cell are denoted by $x_n^{(i)}$ and $P_n^{(i)}$ respectively, where the central DRU of each cell is index by *n*= 0. Where the distributed antenna power constraint is considered, we have,

$$\mathbf{E}\left[\left|x_n^{(i)}\right|^2\right] \leq P_n^{(i)}\,, \tag{A.2}$$

The composite fading channel $h_n^{(i)}$, $i = 0, 1, ..., 18$, encompasses not only small-scale fading but also large-scale fading, which is modeled as $h_n^{(i)} = h_n^{*(i)} l_n^{(i)}$, where $h_n^{*(i)}$ and $l_n^{(i)}$ reflect the small-scale channel fading and the large-scale channel fading between the n-th DRU in the *i*-th cell and the user in the central cell, respectively.

$\left\{h_n^{*(i)} \,|n=0,1,...,6; i=0,1,...,18\right\}$ are independent and identically distribute (*i.i.d*) circularly symmetric complex Gaussian variables with zero mean and unit variance [33], and $l_n^{(i)}$ can be modeled as,

$$l_n^{(i)} = \sqrt{\left[D_n^{(i)}\right]^{-\gamma} \chi_n^{(i)}}\,,\ n=0,1,...,6\,,\ i=0,1,...,18 \tag{A.3}$$

where $D_n^{(i)}$ and $\chi_n^{(i)}$ are independent random variables representing the distance and the shadowing between the user in the central cell and the *n*-th DRU in the *i*-th cell, respectively, γ denotes the path loss exponent. $\left\{\chi_n^{(i)} \,|n=0,1,...,6; i=0,1,...,18\right\}$ are *i.i.d* random variables with probability density function (PDF)

$$f_\chi(\chi) = \frac{1}{\sqrt{2\pi}\lambda\sigma_\chi\chi} \exp\left(-\frac{\left(\ln\chi - \mu_\chi\right)^2}{2\lambda^2\sigma_\chi^2}\right) \quad , \quad \chi > 0, \tag{A.4}$$

where σ_χ and μ_χ are the shadowing standard deviation and mean respectively, and $\lambda = \frac{\ln 10}{10}$ Therefore, SINR at a given location for the DAS system can be modeled by,

$$\gamma = \frac{\sum\limits_{n\in SoS} \left[D_n^{(0)}\right]^{-\delta} \chi_n^{(0)} \left[h_n^{*(0)}\right]^2 P_n^{(0)}}{\sum\limits_{i=1}^{18}\sum\limits_{n\in SoS} \left[D_n^{(i)}\right]^{-\delta} \chi_n^{(i)} \left[h_n^{*(i)}\right]^2 P_n^{(i)} + \sigma^2_{n^{(0)}}} \tag{A.5}$$

A.1 Outage Probability.

In this section the closed-form derivation of the outage probability is expressed for a DAS system. In general, outage occurs when the achieved capacity $C = \log_2(1+\gamma_{ave})$ is less than threshold value R.

$$P_{out}[C<R] = \mathsf{Pr}\left[\gamma_{ave} < 2^R - 1\right] \tag{A.6}$$

By assuming the insignificant noise power compared to the interference power and the fast flat fading in the SINR equation (A.5), the SINR averaged over the fast fading can be modeled by,

$$\gamma_{ave} = \frac{\sum\limits_{n\in SoS} \left[D_n^{(0)}\right]^{-\delta} \chi_n^{(0)} P_n^{(0)}}{\sum\limits_{i=1}^{18}\sum\limits_{n\in SoS} \left[D_n^{(i)}\right]^{-\delta} \chi_n^{(i)} P_n^{(i)}} \tag{A.7}$$

Since the summation and ratio of lognormal random variables can be approximated as a lognormal variable [55] (Appendix A.2), the SINR averaged γ_{ave} becomes a lognormal random variable with mean $\mu_{\gamma_{ave}}$ and variance $\sigma^2_{\gamma_{ave}}$ in dB by,

$$\mu_{\gamma_{ave}} = \lambda^{-1}\ln\left(\tfrac{\mu_N}{\mu_D}\right) - \tfrac{\lambda^{-1}}{2}\ln\left(\tfrac{\mu_D^2}{\mu_N^2}\tfrac{\mu_N^2+\sigma_N^2}{\mu_D^2+\sigma_D^2}\right) \tag{A.8}$$

$$\sigma^2_{\gamma_{ave}} = \lambda^{-2} \ln\left(\frac{(\sigma_N^2+\mu_N^2)(\sigma_D^2+\mu_D^2)}{\mu_N^2\mu_D^2}\right) \tag{A.9}$$

where μ_N [1] and σ_N^2 [2] are the mean and the variance of equation (A.7) nominator, respectively. And μ_D [3] and σ_D^2 [4] are the mean and the variance of equation (A.7) denominator, respectively (Appendix A.2).

Then, the outage probability at a given location is given by,

$$\mathsf{Pr}\big[\gamma_{ave} < 2^R - 1\big] = \mathsf{1} - Q\left(\frac{\mathsf{10log}_{10}(2^R-1)-\mu_{\gamma_{ave}}}{\sigma^2_{\gamma_{ave}}}\right) \tag{A.10}$$

Since the outage derivation of a DAS is a generalization of the conventional cellular system, the outage derivation of the conventional cellular system can be obtained in the same way.

A.2 Lognormal Random Variable Property.

The summation, multiplication and ratio of lognormal random variables can be approximated as a lognormal variable [55].

The ψ_i is the a lognormal random variable with probability density function, ($f_\psi(\psi)$), where σ_ψ and μ_ψ are the shadowing standard deviation and mean respectively, and $\lambda = \frac{\ln 10}{10}$.

$$f_\psi(\psi) = \frac{1}{\sqrt{2\pi}\lambda\sigma_\psi\psi}\exp\left(-\frac{(\ln\psi-\mu_\psi)^2}{2\lambda^2\sigma_\psi^2}\right) \quad , \quad \psi > 0, \tag{A.11}$$

The summation of lognormal random variables ($\sum_{i=1}^{m} d(i)\psi_i$) can be approximated as a lognormal variable with mean μ_{sum} and variance σ_{sum} as,

$$\mu_{sum} = \sum_{i=1}^{m} d(i)\mathsf{exp}\left(\lambda\mu_{\psi_i} + \lambda^2\frac{\left[\sigma_{\psi_i}\right]^2}{2}\right) \tag{A.12}$$

---

[1] $\mu_N = \sum_{n\in SoS}\left[D_n^{(0)}\right]^{-\delta} P_n^{(0)}\mathsf{exp}\left(\lambda\mu_n^{(0)} + \lambda^2\frac{\left[\sigma_n^{(0)}\right]^2}{2}\right)$

[2] $\sigma_N^2 = \sum_{n\in SoS}\left[D_n^{(0)}\right]^{-2\delta}\left[P_n^{(0)}\right]^2\mathsf{exp}\left(2\lambda\mu_n^{(0)} + \lambda^2\left[\sigma_n^{(0)}\right]^2\right)\left(\mathsf{exp}\left(\lambda^2\left[\sigma_n^{(0)}\right]^2\right) - 1\right)$

[3] $\mu_D = \sum_{i=1}^{18}\sum_{n\in SoS}\left[D_n^{(i)}\right]^{-\delta} P_n^{(i)}\mathsf{exp}\left(\lambda\mu_n^{(i)} + \lambda^2\frac{\left[\sigma_n^{(i)}\right]^2}{2}\right)$

[4] $\sigma_D^2 = \sum_{i=1}^{18}\sum_{n\in SoS}\left[D_n^{(i)}\right]^{-2\delta}\left[P_n^{(i)}\right]^2\mathsf{exp}\left(2\lambda\mu_n^{(i)} + \lambda^2\left[\sigma_n^{(i)}\right]^2\right)\left(\mathsf{exp}\left(\lambda^2\left[\sigma_n^{(i)}\right]^2\right) - 1\right)$

$$\sigma_{sum}^2 = \sum_{i=1}^{m} d(i)\exp\Big(2\lambda\mu_{\psi_i} + \lambda^2\left[\sigma_{\psi_i}\right]^2\Big)\Big(\exp\Big(\lambda^2\left[\sigma_{\psi_i}\right]^2\Big) - 1\Big) \tag{A.13}$$

where $d(i) \in \{-1, 1\}$. The multiplication of lognormal random variables ($\prod_{i=1}^{m} \psi_i^{d(i)}$) can be approximated as a lognormal variable with mean μ_{mul} and variance σ_{mul} as,

$$\mu_{mul} = \sum_{i=1}^{3} d(i)\left(\lambda^{-1}\ln\left(\mu_{\psi_i}\right) - \tfrac{\lambda^{-1}}{2}\ln\left(\tfrac{\sigma_{\psi_i}^2}{\mu_{\psi_i}^2} + 1\right)\right) \tag{A.14}$$

$$\sigma_{mul}^2 = \sum_{i=1}^{3} d(i)\lambda^{-2}\ln\left(\tfrac{\sigma_{\psi_i}^2}{\mu_{\psi_i}^2} + 1\right) \tag{A.15}$$

where $d(i) \in \{-1, 1\}$.

A.3 Average Spectral Efficiency.

In this section, the closed-form derivation of the ASE is expressed for a DAS system. Since the number of interfering sources are sufficiently large and interfering sources are independent of each other, the interference plus noise is assumed to be a zero-mean complex Gaussian random variable as follows:

$$N = \sum_{i=1}^{18}\sum_{n\in SoS} h_n^{(i)} x_n^{(i)} + n^{(0)} \tag{A.16}$$

The variance of *N* is derived by Central Limit Theorem as

$$\begin{aligned} Var(N) &= \sum_{i=1}^{18}\sum_{n\in SoS}\left[l_n^{(i)}\right]^2 P_n^{(i)} + \sigma_{n^{(0)}}^2 \\ &= \sigma^2 \end{aligned} \tag{A.17}$$

Therefore, the received signal at the mobile station at a given symbol duration is given by

$$y^{(0)} = \sum_{n\in SoS} h_n^{*(0)} l_n^{(0)} x_n^{(0)} + N \tag{A.18}$$

By assuming that the channel is ergodic and the channel state information is known only at the receiver (CSIR), the average spectral efficiency (ASE) at a given location of the target UE for the central cell can be calculated by

$$\begin{aligned} \vartheta &= \mathbf{E}_{h^{*(0)}}\left[\log_2(1 + \tfrac{1}{\sigma^2}\sum_{n\in SoS}\left|h_n^{*(0)}\right|^2\left[l_n^{(0)}\right]^2 P_n^{(0)})\right] \\ &= \int_{\gamma_f=0}^{\infty} \log_2(1+\gamma_f) f_{\gamma_f}(\gamma_f) d\gamma_f \end{aligned} \tag{A.19}$$

where $\gamma_f = \frac{1}{\sigma^2} \sum\limits_{n \in SoS} \left| h_n^{*(0)} \right|^2 \left[l_n^{(0)} \right]^2 P_n^{(0)}$ is a weighted chi-squared distributed random variable with p.d.f given by [48]

$$f_{\gamma_f}(\gamma_f) = \sum_{n \in SoS} \frac{\sigma^2 \pi_n}{\left[l_n^{(0)} \right]^2 P_n^{(0)}} \exp\left(- \frac{\sigma^2 \gamma_f}{\left[l_n^{(0)} \right]^2 P_n^{(0)}} \right) \tag{A.20}$$

where $\pi_n = \prod\limits_{k \in SoS, k \neq n} \frac{\left[l_n^{(0)} \right]^2 P_n^{(0)}}{\left[l_n^{(0)} \right]^2 P_n^{(0)} - \left[l_k^{(0)} \right]^2 P_k^{(0)}}$. The simple form of ASE can be obtained using integral table [49] for MISO,

$$\mathbf{MISO}: \quad \vartheta^{(MISO)} = -\frac{1}{\ln 2} \sum_{n \in SoS} \pi_n \exp\left(- \frac{\sigma^2}{\left[l_n^{(0)} \right]^2 P_n^{(0)}} \right) Ei\left(- \frac{\sigma^2}{\left[l_n^{(0)} \right]^2 P_n^{(0)}} \right) \tag{A.21}$$

where, $(Ei(t) = -\int\limits_{-x}^{\infty} e^{-t}/t \, dt)$ can be easily calculated with popular numerical tools such as MATLAB. Note that the derivations for ASE and the received signal are independent of the number of participating users. More specifically, the underlying derivations for ASE and the received signal mainly rely on the users' positions.

Since the derivation for this MISO vector channel is a generalization of a SISO channel, the ASE for SISO channel is given, respectively, by

$$\mathbf{SISO}: \quad \vartheta^{(SISO)} = -\frac{1}{\ln 2} \exp\left(- \frac{\sigma^2}{\left[l_0^{(0)} \right]^2 P_0^{(0)}} \right) Ei\left(- \frac{\sigma^2}{\left[l_0^{(0)} \right]^2 P_0^{(0)}} \right) \tag{A.22}$$

Appendix B

Evolutionary Algorithms

B.1 Genetic Algorithm (GA) and Estimation Distribution Algorithm (EDA).

GAs are stochastic search methods that mimic genetic phenomena such as gene recombination, mutation and survival of the fittest. GAs have been applied to a large number of scientic and engineering problems, including many combinatorial optimization problems in networks. GAs [47], operate on a set of candidate solutions, called a population. Each solution is represented by a string, which is called a chromosome. Each chromosome is assigned a fitness value that measures how well the chromosome solves the problem at hand, compared with other chromosomes in the population. From the current population, a new population is generated typically using three genetic operators: selection, crossover and mutation. Chromosomes for the new population are selected randomly (with replacement) in such a way that fitter chromosomes are selected with higher probability. For crossover, survived chromosomes are randomly paired, and then two chromosomes in each pair exchange a subset of their strings to create two offspring. Chromosomes are then subject to mutation, which refers to random ips of the strings element applied individually to each of the new chromosomes. The process of evaluation, selection, crossover and mutation forms one generation in the execution of a GA. The above process is iterated with the newly generated population successively replacing the current one. The GA terminates when a certain stopping criterion is reached, e.g., after a predefined number of generations. There are several aspects of our problem suggesting that a GA-based method may be a promising candidate: GA has proven to work well if the space to be searched is large, but known not to be perfectly smooth, or if the space is not well understood, and if finding a global optimum is not critical [47]. In general, they use a penalty function to encode problem constraints and allow a search for illegal solutions. Allowing a search for illegal solutions may prevent falling down into a local minimum and generate a better solution.

During each generation of the GAs, individuals in the current population are rated for their fitness as domain solutions.

Unlike other evolutionary algorithms, in EDA a new population of individuals in each generation is generated without crossover and mutation operators. Instead, in EDA a new population is generated based on a probability distribution, which is estimated from the best selection of individuals in our previous generations [50].

In general, conventional GAs and EDAs can be characterized by parameters and notations:

1) I_s is the space of all potential solutions (entire search space of GAs chromosomes or EDAs individual).

2) F denotes a fitness function.

3) Δ_l is the set of GA's chromosomes or EDA's individuals (population) at the l_{th} generation. $|\Delta_l|$ is population size at the l_{th} generation. Population Δ_l can be specified by the matrix $\mathbf{Q}_l = \left\{Q^j | j = 1, 2, ..., \Delta_l\right\}$ where Q^j=$\left\{q_1^j, q_2^j, ..., q_N^j\right\}$ is an individual (or a chromosome). Each member (q_i^j) in an individual (or a chromosome) represents the sector to which the corresponding DRUs belong where $q_i^j = k$ means DRU_i belongs to $Sector_k$ in j-th individual (or chromosome) of population.

4) η_l is the set of best candidate solutions selected from set Δ_l at the l_{th} generation.

5) We denote $\beta_l \equiv \Delta_l - \eta_l \equiv \Delta_l \cap \eta_l^c$. where η_l^c is the complement of η_l.

6) p_s is the selection probability. The GA and EDA algorithms select $p_s\,|\Delta_l|$ individuals from set Δ_l to make up set η_l.

7) In GA: P_c is the crossover probability. In the uniform crossover process two chromosomes are randomly chosen with a probability P_c and genes are exchanged with a rate C_R. P_m is the mutation probability. For the mutation a gene is randomly chosen with a probability P_m and the value of the gene is changed. That is, the DRU changes its sector. In EDA: We denote by Γ the distribution estimated from η_l (the set of selected candidate solutions) at each generation.

8) I_{Ter} are the maximum number of generation.

A typical GA and EDA are described in the following steps, and Fig. B.1 presents the algorithms:

Step 0: Generate initial population Δ_0. The initial population ($|\Delta_l|$ chromosomes or EDA's individuals) is typically obtained by sampling according to the uniform (equally likely) distribution, $p\left(\theta_1, \theta_2, \cdots, \theta_N\right) = \prod_{i=1}^{N} p_i(\theta_i)$,$p_i(\theta_i = 1) = p_i(\theta_i = 2) = ... = p_i(\theta_i = K) = \frac{1}{K}$.

Step 1: Evaluate the chromosomes or EDA's individuals in the current population Δ_{l-1} according to the fitness function F. Sort the candidate solutions (GA's chromosomes or EDA's individual in the current population) according to their fitness orders.

Step 2: If the best candidate solution satisfies the convergence criterion or the number of iterations exceeds its limit l_{Ter} , then terminate; else go to step 3.

Step 3: Select the best $p_s\Delta_{l-1}$ candidate solutions (chromosomes/Individuals) from current population Δ_{l-1}. This selection is accomplished according to the sorted candidate solutions.

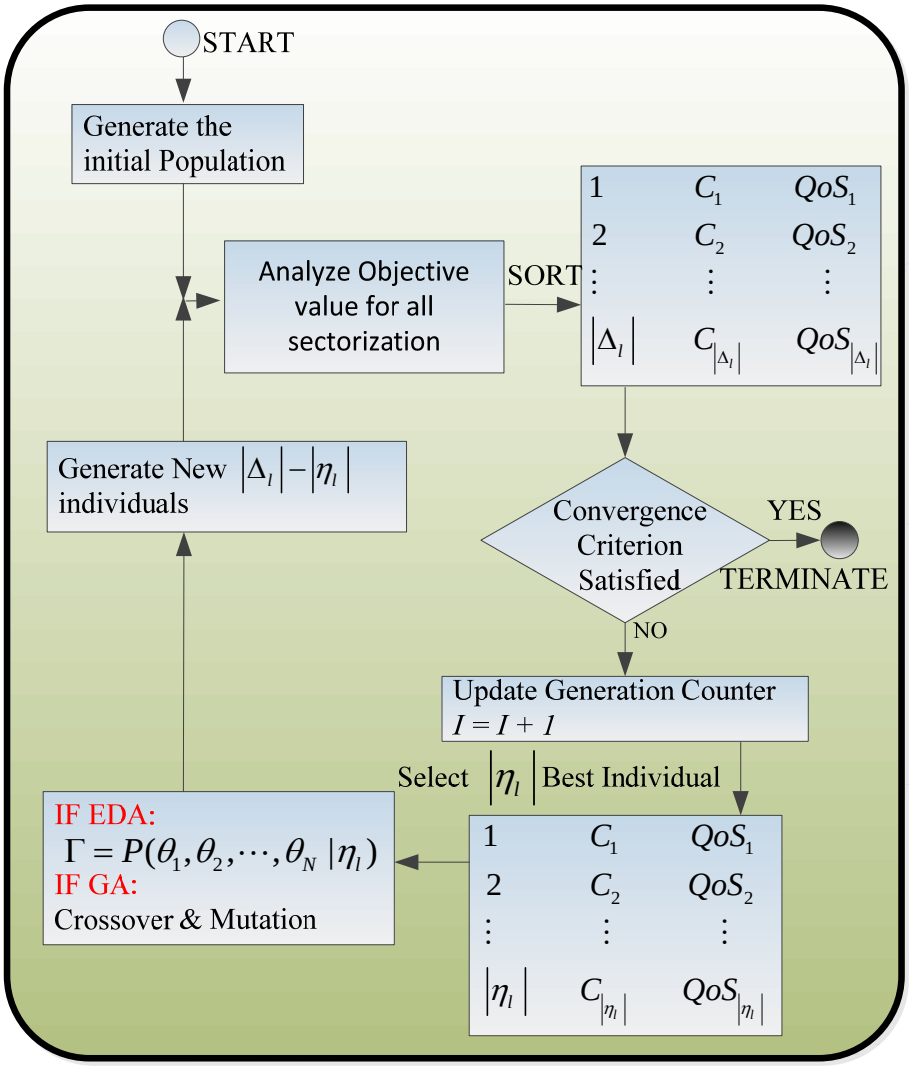


Figure B.1: Block diagram of the EDA and GA algorithm.

Step 4: In GA: generating new $|\beta_l|$ chromosomes according to uniform crossover and flipping mutation. Newly generated chromosomes are replaced by the previous $|\beta_{l-1}|$ chromosomes. Uniform crossover process is performed on the two chromosomes which are randomly chosen from η_l and they are called parents. The uniform crossover is performed with probability P_c. In uniform crossover with exchanging rate C_R, 100$\times C_R$ percent of the genes from the first parent is chosen and the other genes are chosen from the second parent. For the mutation, gene is randomly chosen from each produced children chromosome (results of crossover) with probability P_m and then the value of gene is changed to the other value (choosing randomly from the other possible value)

In EDA: Estimate the probability distribution $p\left(\theta_1, \theta_2, \cdots, \theta_N\right)$ according to the $|\eta_{l-1}|$ best candidate solutions. $\Gamma = P(\theta_1, \theta_2, \cdots, \theta_N | \, \eta_{l-1})$ denotes this probability distribution estimation. Generate new $|\Delta_{l-1}| - |\eta_{l-1}|$ individuals according to this new estimated probability distribution Γ . Replace the bad $|\beta_{l-1}|$ individuals with newly generated $|\Delta_{l-1}| - |\eta_{l-1}|$ individuals ($|\beta_{l-1}| = |\Delta_{l-1}| - |\eta_{l-1}|$). Γ is obtained by:

$$\Gamma = p(\theta_1, \theta_2, \cdots, \theta_N | \, \eta_{l-1}) = \prod_{i=1}^{n} p_i(\theta_i | \, \eta_{l-1}) = \prod_{i=1}^{n} \left(\frac{\sum_{j=1}^{|\eta_{l-1}|} \delta\left(q_i^j = \theta_i | \, \eta_{l-1}\right)}{|\eta_{l-1}|} \right) \tag{B.1}$$

where δ is expressed as an indicator function by,

$$\delta\left(q_i^j = \theta | \, \eta_{l-1}\right) = \begin{cases} 1 & \text{if } q_i^j = \theta \\ 0 & \text{otherwise} \end{cases} \tag{B.2}$$

Step 5: Go to step 1 and repeat the steps.

B.2 Discrete Particle Swarm Optimization (DPSO).

Particle Swarm Optimization (PSO) is derived from the social behavior of bird flocking and fish schooling. The PSO is proposed to solve global optimization problems [58]. PSO is a population-based search algorithm where each individual is called a particle, and a collection of particles is called population (or swarm). Considering the optimization of the fitness function, PSO tries to find the value of a vector variable $\{u_1, u_2, ..., u_M\}$ that best fits in accordance with the fitness measure $F(\{u_1, u_2, ..., u_M\})$. The vector $\{u_1, u_2, ..., u_M\}$ can be analogically viewed as a position of a particle in the M-dimensional space and finding the best solution can be analogically viewed as particles flying in the space searching for the optimum position. In a PSO system, particles search around (fly and change their position) in a multidimensional search space (set of all potential solutions). The combination of local search methods with global search methods is utilized in a PSO system, where each particle adjusts its position on the basis of its own best position experience and the neighboring particles best position experience during the flight. Therefore, each particle has a tendency to fly (move) towards better solutions [59]. The standard PSO algorithm cannot be used

for a discrete problem where positions are real-valued. Pan et al. Reference [60] introduced the new position update method for particles based on discrete job permutations.

In general, conventional DPSO can be characterized by parameters and notations:

1)F denotes a fitness function.

2)$|\Delta|$ is the size of population, i.e., the number of particles.

3)I_{gen} are the maximum number of generation.

4)M is the dimension of the particle position (the length of the discrete string ($\{u_1, u_2, ..., u_M\}$) to represent the particle position).

5)The position of the particles at the I-th generation is represented by $\mathbf{U}^I = [\mathbf{u}_1^I; \mathbf{u}_2^I; ...; \mathbf{u}_{|\Delta|}^I]$, where $\mathbf{u}_i^I, i = 1, 2, ..., |\Delta|$ represent the position of the particle indexed by i. Components of vector $\mathbf{u}_i^I$ are denoted as $\mathbf{u}_i^I = (u_{i,1}^I, u_{i,2}^I, ..., u_{i,M}^I)$, where $u_{i,j}^I \in \{1, 2, ..., N\}, j = 1, 2, ..., M$.

6)P_{bi}^I represents the best position of particle i up to the I-th generation (best among the history of particle is positions); that is, $P_{bi}^I \equiv \arg\max_{0 \le s \le I} F(\mathbf{u}_i^s)$.

7)G_B^I represents the globally best position ever visited by any particle up to the I-th generation in terms of fitness function F. That is, $G_B^I \equiv \arg\max_{1 \le i \le |\Delta|} F(P_{bi}^I) \equiv \arg\max_{1 \le i \le |\Delta|, 1 \le s \le I} F(\mathbf{u}_i^s)$.

A typical DPSO are described in the following steps and Fig. B.2 presents the algorithms:

Step 0: Generate initial population ($\mathbf{U}^0$). At the initiation I=0, initialize each particles best position $P_{bi}^0 = \mathbf{u}_i^0, 1 \le i \le |\Delta|$, initialized the global best as, $G_B^0 = (g_{b1}^0, g_{b2}^0, ..., g_{bM}^0) = \arg\max_{1 \le i \le |\Delta|} F(P_{bi}^0)$.

Step 1: Update the generation counter ($I = I + 1$).

Step 2: Update the position of particle i as,

$$\mathbf{u}_i^I = c_2 \otimes A_3 \left(c_1 \otimes A_2 \left(w \otimes A_1 \left(\mathbf{u}_i^{I-1} \right), P_{bi}^{I-1} \right), G_B^{I-1} \right), 1 \le i \le |\Delta| \tag{B.3}$$

- $\lambda_i^I = w \otimes A_1 \left(\mathbf{u}_i^{I-1} \right)$ represents the velocity of the particle where A_1 represents the mutation operator with the probability of w. In other words,

$$\lambda_i^I = \begin{cases} A_1 \left(\mathbf{u}_i^{I-1} \right) & if \ \ U(0,1) < w \\ \mathbf{u}_i^{I-1} & otherwise \end{cases} \tag{B.4}$$

  where $U(0,1)$ is the random variable uniformly distributed in interval (0,1).

- $\delta_i^I = c_1 \otimes A_2 \left(\lambda_i^I, P_{bi}^{I-1} \right)$ is the "cognition" part of the particle representing the private thinking of the particle itself, where A_2 represents the crossover operator with the probability of c_1. The first and second parents for the crossover operator are λ_i^I and P_{bi}^{I-1}, respectively. In other words,

$$\delta_i^I = \begin{cases} A_2 \left(\lambda_i^I, P_{bi}^{I-1} \right) & if \ \ U(0,1) < c_1 \\ \lambda_i^I & otherwise \end{cases} \tag{B.5}$$

- $\mathbf{u}_i^I = c_2 \otimes A_3 \left(\delta_i^I, G_B^{I-1}\right)$is the "social" part of the particle representing the collaboration among particles where A_3 represents the crossover operator with the probability c_2. The first and second parents for the crossover operator are δ_i^I and G_B^{I-1}, respectively.

$$\mathbf{u}_i^I = \begin{cases} A_3 \left(\delta_i^I, G_B^{I-1}\right) & if \ \ U(0,1) < c_2 \\ \delta_i^I & otherwise \end{cases} \tag{B.6}$$

Step 3: If the convergence criteria are satisfied, then terminate. Otherwise go to step 4.
Step 4: For $i = 1, 2, ..., |\Delta|$, update the particle i's best as,

$$P_{bi}^I = \begin{cases} P_{bi}^{I-1} \ \ if \ F(\mathbf{u}_i^I) \leq F(P_{bi}^{I-1}) \\ \mathbf{u}_i^I \ \ if \ F(\mathbf{u}_i^I) \geq F(P_{bi}^{I-1}) \end{cases} \tag{B.7}$$

Step 5: Update the global best as,

$$G_B^I = \begin{cases} \underset{1 \leq i \leq |\Delta|}{\arg\max} \ F(P_{bi}^{I-1}) & if \quad \underset{1 \leq i \leq |\Delta|}{\max} \ F(P_{bi}^I) > F(G_B^{I-1}) \\ G_B^{I-1} & otherwise \end{cases} \tag{B.8}$$

Step 6: Go to step 1 and repeat the steps.

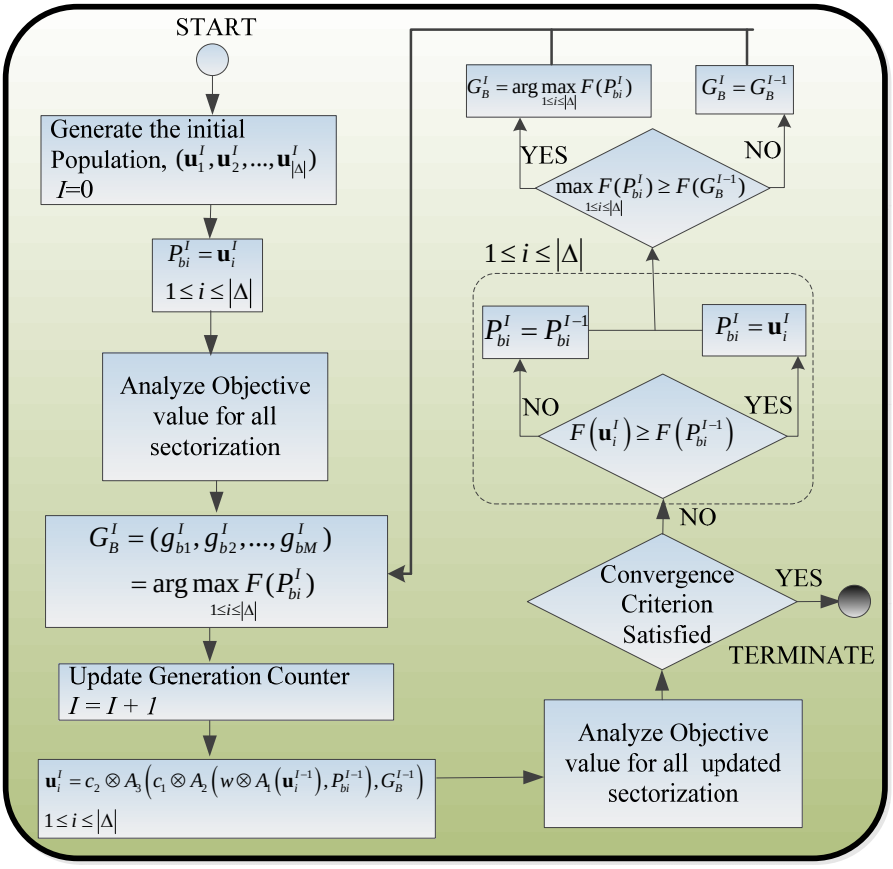


Figure B.2: Block diagram of the DPSO algorithm.

Appendix C

Traffic Monitoring in a LTE Distributed Antenna System

A method is required for monitoring the information to drive a SON algorithm such as the channel quality between each individual pair of DRUs and UEs (User Equipment), the number of UEs associated with a given DRU, and which UE is associated with which DRU.

C.1 Traffic Monitoring Solution.

The uplink received power at each DRU is primarily influenced by the distance between the UE and the DRU. Therefore, UE_i is associated with DRU_j when the received power of UE_i at DRU_j is greater than the UE_i received power at the other DRUs. One solution to determine that UE_i is associated with a given DRU is by comparing the received uplink powers of UE_i at the different DRUs. Although demodulating the Physical Uplink Control Channel (PUCCH) will extract the uplink control information such as the Channel Quality Indicator (CQI) [53, Ch. 16.3], this technique will require complex hardware resources. On the other hand, the CQIs transmitted by UEs represent the quality of the channel between the eNB and UE, which is not sufficient for the SON algorithm to design a new DRU allocation configuration. The SON algorithm needs the channel quality between the DRU and UE, not between eNB and UE, to properly optimize the DRU allocation configuration. Therefore, extracting the physical reference signals from the received uplink signal at the DRU, for each individual UE, is one solution to estimate the channel quality between DRU_j and UE_i. The Demodulation Reference Signals (DM-RSs) associated with the physical uplink channel are primarily provided for channel estimation, and therefore present in every transmitted uplink slot [54, Ch. 10.2.2]. A DM-RS is intended for a specific UE and is only transmitted in the RBs (Resource

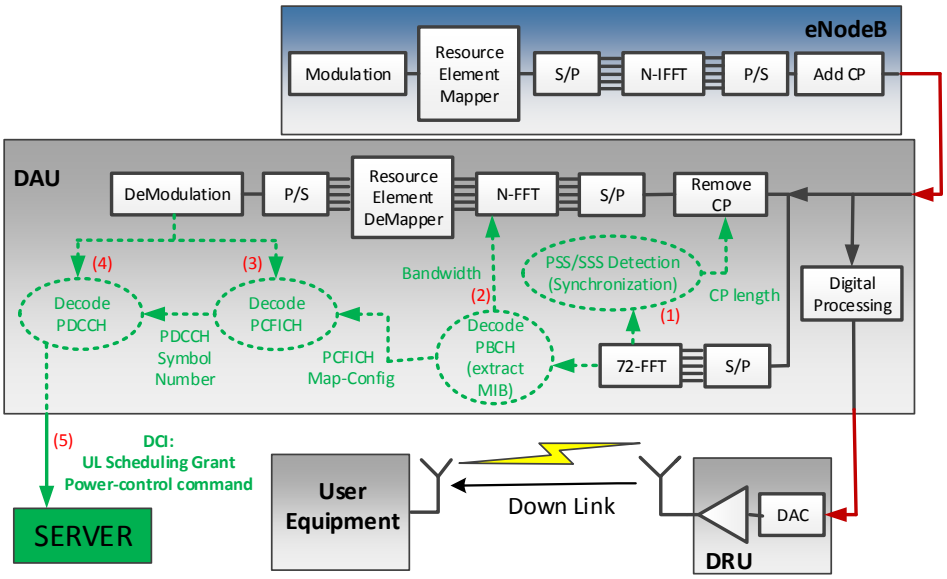


Figure C.1: EDCI: UL control information extracting procedure from DL signal.

Block (group of subcarriers)) assigned for transmission to that UE. The DM-RS are based on Zadoff-chu sequences with constant amplitude. “Extracting Downlink Control Information” and “Extracting Uplink Radio Frame” are two required procedures which are explained in the following subsections,

C.1.1 Extracting Downlink Control Information (EDCI).

The Downlink Control Information (DCI) which includes downlink scheduling assignments, uplink scheduling grants, power-control commands and other control information for UEs [54, CH. 10.4.3], will be obtained from the received downlink signal at each DAU via the eNB by using the following steps (Fig. C.1).

Step 1- Synchronization and cell search: The detection of two Primary Synchronization signal (PSS) and Secondary Synchronization Signal (SSS) not only enables time and frequency synchronization, but also provides the identity of cell and cyclic prefix (CP) length, as well as whether the cell uses Frequency Division Duplex (FDD) or Time Division Duplex (TDD) [53, Ch. 7.2]. The PSS and SSS are each comprised of a 62 length Zadoff-chu sequence [53, Ch 7.2.1] symbols, mapped to the central 62 subcarriers around the D.C. subcarrier, which is left unused [54, Ch. 14.1.2]. This structure enables the detection of the PSS and SSS using a size-72 Fast Fourier Transform (72-FFT) (Fig. C.1, EDCI (1)).

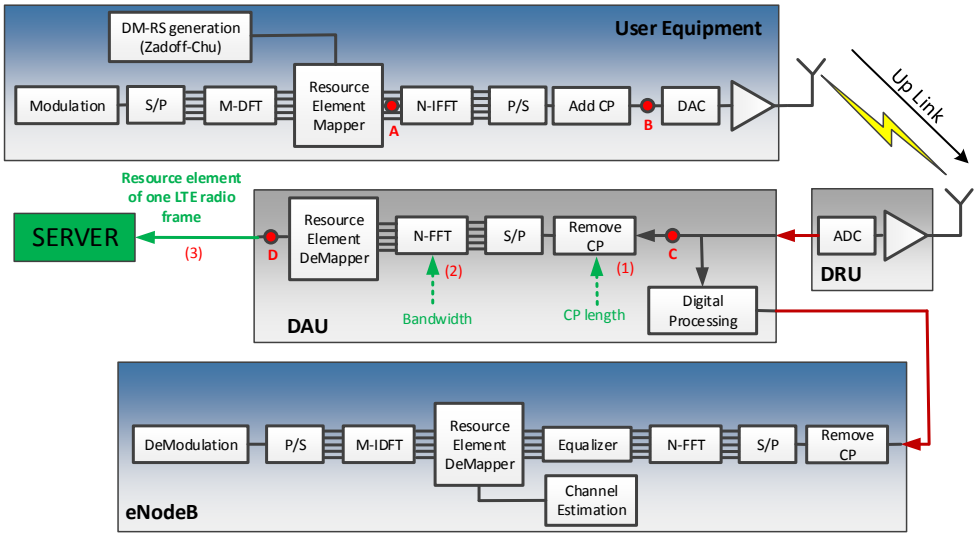


Figure C.2: EURF: De-mapping the resource element of one radio frame from UL signal.

Step 2- Physical Broadcast Channel (PBCH) decoding: The Master Information Block (MIB) transmitted using PBCH, consists of a limited amount of system information such as cell bandwidth and Physical Control Format Indicator Channel (PCFICH) configuration of the cell. Detectability without knowing the system bandwidth is achieved by mapping the PBCH only to the central 72 subcarrier (minimum possible LTE bandwidth of 6 RBs), regardless of the actual system bandwidth. This structure enables decoding of the PBCH using 72-point FFT [53, Ch. 9.2.1] (Fig. C.1, EDCI (2)).

Step 3- Physical Control Format Indicator Channel (PCFICH) decoding: The PCFICH carries a Control Format Indicator (CFI) which indicates the size of a control region in terms of the number of OFDM symbols (i.e. 1, 2 or 3) [53, Ch. 9.3.3]. Figuring out the value of the CFI is possible by decoding the PCFICH (PCFICH Map configuration is obtained at step 2). Note that CP length obtained at step 1 helps to remove the CP from the received digital symbol and the actual bandwidth obtained at step 2 helps to determine the FFT size (Fig. C.1, EDCI (3)).

Step 4- Physical Downlink Control Channel (PDCCH) decoding: The PDCCH carries the DCIs. DAU need to blindly detect all UEs' PDCCH by searching the PDCCH region. PDCCH region is detected by decoding the CFI in step 3. Note that, the blind decoding is performed for all possible UEs to collect the control information for all users, separately [54, CH. 10.4.3] (Fig. C.1, EDCI (4)).

Step 5- Transmitting to Server: transmit all control information for all users to the server via an optical fiber (Fig. C.1, EDCI (5)).

C.1.2 Extracting Uplink Radio Frame (EURF).

The LTE UL radio frame can be extracted with the following steps (Fig. C.2).

Step 1- Removing CP and Performing the N-FFT: CP length and actual bandwidth were obtained in step 1 and 2 of "Extracting Downlink Control Information", respectively (Fig. C.2, EURF (1)).

Step 2- Build LTE radio Frame: LTE Frame Builder saves all the OFDM symbols after the N-FFT and builds a LTE radio frame (10 ms) (Fig. C.2, EURF (2)).

Step 3- Transmitting to Server: separately transmit all DRUs LTE radio frame to server (Fig. C.2, EURF (3)). Now, by knowing the uplink scheduling map obtained from EDCI and uplink LTE radio frame of each DRU obtained from EURF, the server can easily compare received power strength of different UEs based on scheduling map and make a decision whether which UE is associated to which DRU.

Since the DM-RSs are always transmitted with the same power as the corresponding physical channel, estimation channel quality between each pair of DRU and UE can be done using DM-RS symbols obtained from EURF and power control command obtained from EDCI.

It is worth mentioning that the uplink power control insures that the received power of different UEs should be almost the same at the eNB. Since the uplink signals from all the UEs at the DRUs are summed before the eNB, the ratio of the number of UEs belonging to a different DRU is the same as the ratio of uplink power strength at point C in Fig. C.2 for the different DRUs (equation (C.1)). Note that, there is no need to extract downlink control information and uplink radio frame to find this ratio.

$$\frac{N_i}{N_j} = \frac{P_i}{P_j} \tag{C.1}$$

where N_i and P_i are the number of UEs associated with DRU_i and the uplink power strength at DRU_i during at least one radio frame. Note that the number of users cannot be obtained by measuring only the received power at the DRUs. In the following section, we demonstrate an example to show how the solution method helps to monitor the traffic.

C.2 Example of Traffic Monitoring.

Fig. C.3 demonstrates an example when 4 users are distributed and supported by 1 eNB and 2 DRUs. Cell bandwidth is obtained after synchronization and decoding the PBCH (EDCI, step 1 and 2). In this example, 1.4 MHz (6 RB) bandwidth is used for transmission where 2 edge RBs are assigned to users as Physical Uplink Control Channel (PUCCH) and 4 center RBs are assigned to users as Physical Uplink Share Channel (PUSCH) at each sub-frame. The uplink scheduling map for one radio frame is shown in Fig. C.4 which is obtained from decoding the PDCCH. Before

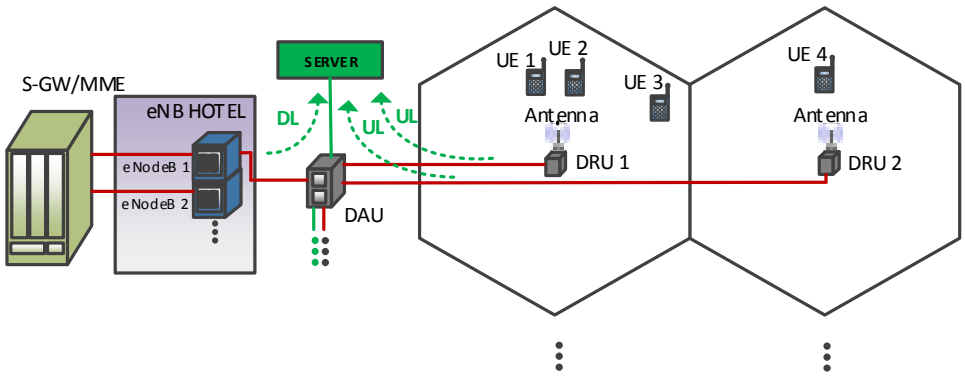


Figure C.3: Structure of Traffic Monitoring.

	SF1		SF2		SF3		SF4		SF5		SF6		SF7		SF8		SF9		SF10	
	TS1	TS2	TS1	TS2	TS1	TS2	TS1	TS2	TS1	TS2	TS1	TS2	TS1	TS2	TS1	TS2	TS1	TS2	TS1	TS2
RB1(PUCCH)																				
RB2(PUSCH)	UE3		UE2		UE3		UE4		UE1		UE4		UE3		UE2		UE1		UE1	
RB3(PUSCH)	UE2		UE1		UE1		UE4		UE3		UE2		UE4		UE3		UE2		UE4	
RB4(PUSCH)	UE1		UE4		UE4		UE2		UE3		No UE		UE4		UE4		UE3		UE2	
RB5(PUSCH)	UE1		UE4		UE4		UE3		UE2		UE1		UE1		UE4		UE3		UE2	
RB6(PUCCH)																				

Figure C.4: UL scheduling map for one LTE radio frame (SF: sub-frame, TS: time slot).

building the radio frame from received signals at each DRU, consider how the LTE frame looks like at each individual UE. Fig. C.5 shows the magnitude of mapped resource element during one radio frame for different users at point A of Fig. C.4. Note that the mapping resource elements are based on decoded UL scheduling grant from PDCCH at each user (EDCI, step 5). All mapped resource elements for the different users are concatenated in Fig.C.6 to verify the scheduling map decode in Fig. C.4. Note that the 4th and 11th OFDM symbols (red line) of each sub-frame are reserved for DM-RS which is generated by Zedoff-chu sequence with constant amplitude. Fig. C.7 shows signals at point B in Fig. C.2 after performing a 72-point FFT (when 1.4 MHz is the bandwidth), making them serial and adding CP for different users during one LTE radio frame (10 ms). The DM-RS signals with constant amplitude are shown as well. The received combination of uplink signals is converted to digital baseband in the DRU using an ADC (Analogue to Digital Convertor), and they are transmitted to the DAU by an optical fiber. Fig. C.8 shows the combination of all received users digital signals at point C in Fig. C.2 at two different DRUs. The number of UEs associated with

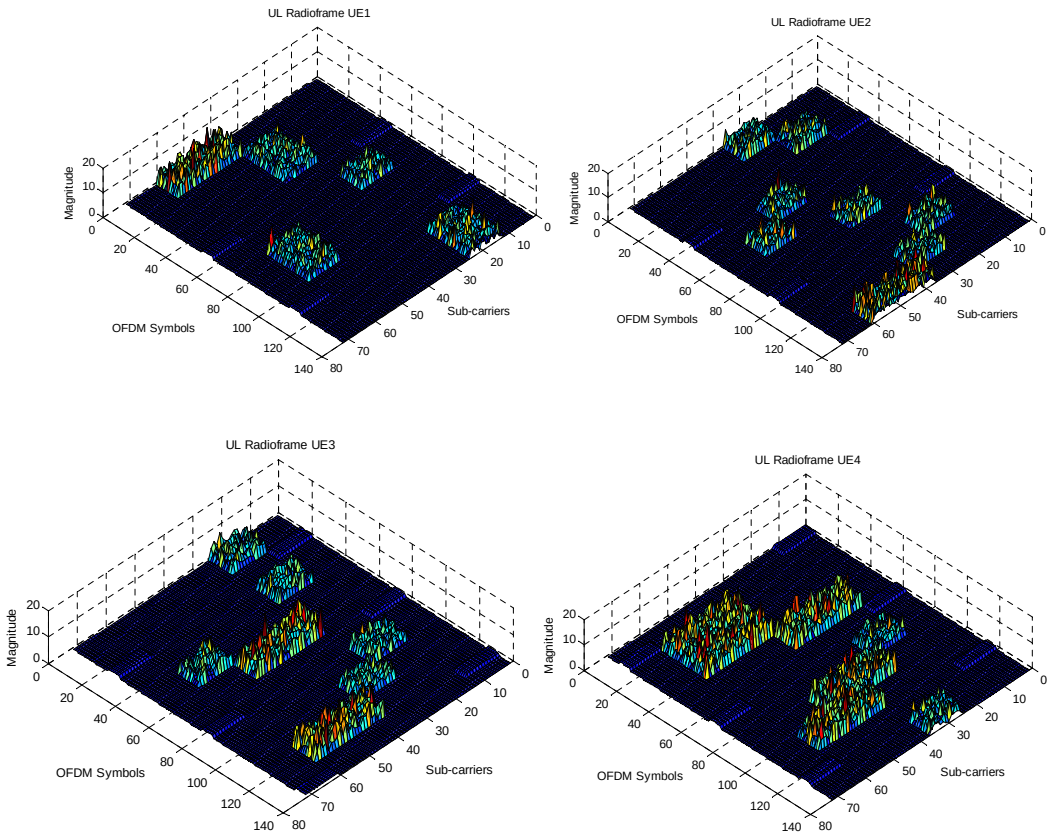


Figure C.5: Mapped resource elements of four UE1, UE2, UE3 and UE4 during one LTE radio frame.

DRU1 is approximately three times greater than DRU2 based on what is explained in section III, equation (C.1). This is because the received power strength over one radio frame (1 ms) at DRU1 is approximately three times greater than DRU2.

The magnitude of the de-mapped resource elements of the received signal at point D in Fig. C.2, for one radio frame, is shown in Fig. C.9 and Fig. C.10 for DRU1 and DRU2, respectively. Note that the de-mapped resource elements are obtained after removing CP and performing a 72-point FFT on the received signal at each DRU (EURF, step 5). The left figures of Fig. C.9 and 10 show color figures in such a way that colors identify different power levels i.e. red and green are assigned to highest and lowest power level, respectively. The right figures of Fig. C.9 and 10 distinguish each users' RBs with different color, e.g. dark blue, light blue, green and red are assigned to user 1, 2, 3 and 4, respectively. Now, by knowing the uplink scheduling map obtained from EDCI step 5, the server can easily compare received power strength of UE_i (i = 1, 2, 3 or 4) at the RBs associated to UE_i at DRU1 and DRU2 and make a decision whether UE_i is associated with DRU1 or DRU2.

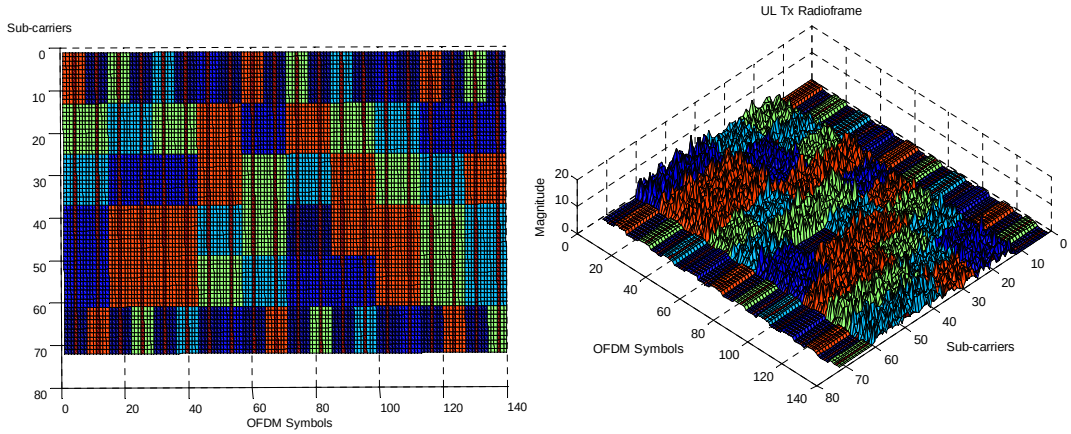


Figure C.6: Mapped resource elements of four UE1, UE2, UE3 and UE4 together.

We define UE_i associated to DRU_j when the received power of UE_i at DRU_j is greater than the UE_i received power at the other DRUs, i.e. in this example, UE1, UE2 and UE3 is associated with DRU1 and UE4 is associated with DRU2. In other words, UEi is associated with DRU_j when it is closer to DRU_j than the other DRUs. Since the DM-RS symbols are always transmitted with the same power as the corresponding physical channel, received power strength comparison can be done by de-mapping only DM-RS symbol by performing the 72-point FFT only at the 4th and 11th symbol of each sub-frame. Although, de-mapping only the DM-RS decreases the required memory size at the DAU rather than de-mapping all the symbols, it requires a complex DM-RS detector with an accurate synchronization tool to extract DM-RS signal. Fig. C.11 shows the magnitude of the de-mapped DM-RS resource elements of the received signal at DRU1 and DRU2. Different colors in the Figures demonstrate the power levels. Note that, since the DM-RS symbols were generated based on Zadoff-chu sequences with constant amplitude, all DM-RS of specific UEi have equal constant amplitude. The server can estimate the channel quality between DRUj and UEi using the de-mapped DM-RS frames (EURF step C.2) transmitted from DAU to server and UL control information (EDCI step 5) transmitted from eNB to server such as power control command and UL scheduling grants.

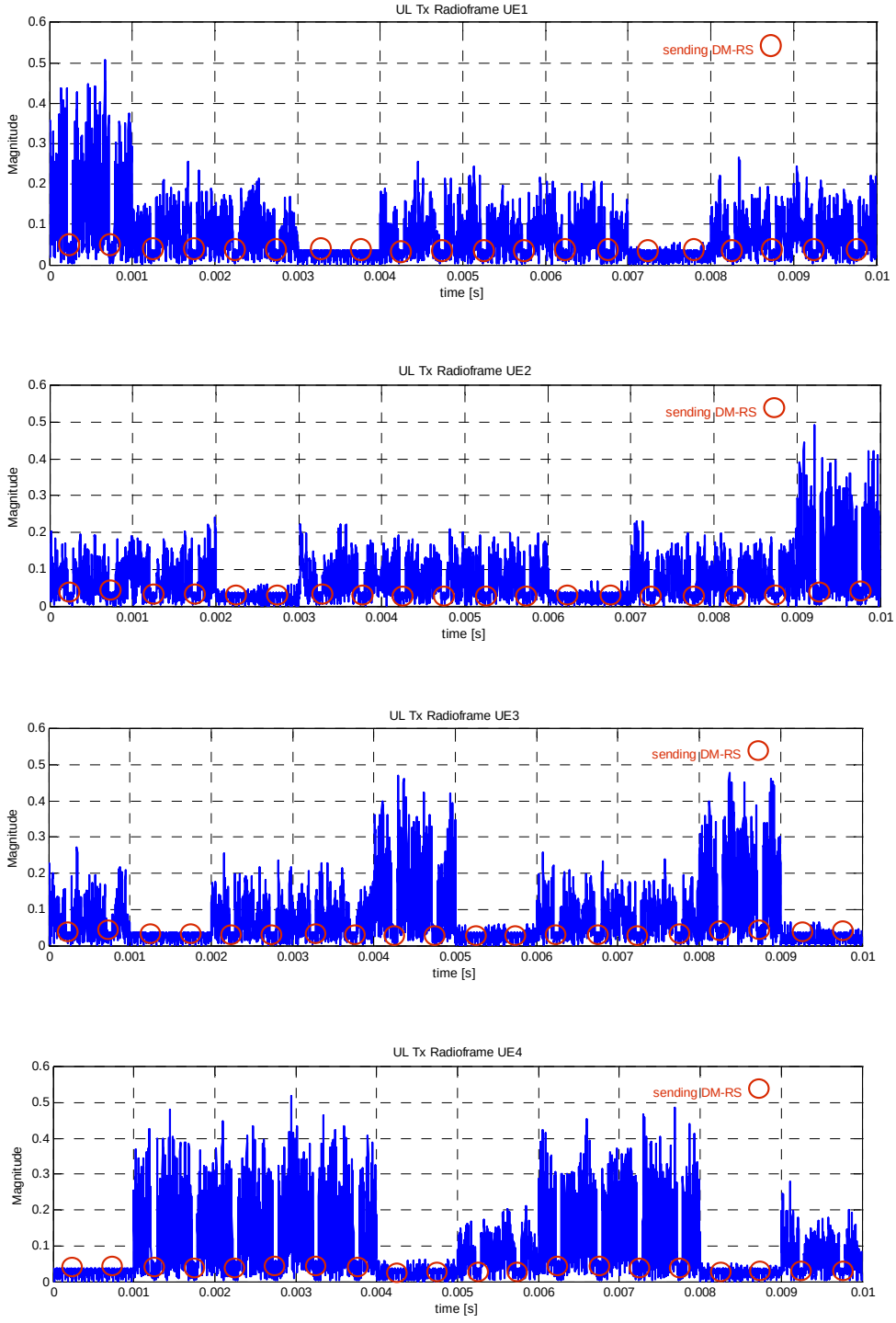


Figure C.7: The signals of point B in Fig. C.2 for UE1, UE2, UE3 and UE4 during one radio frame.

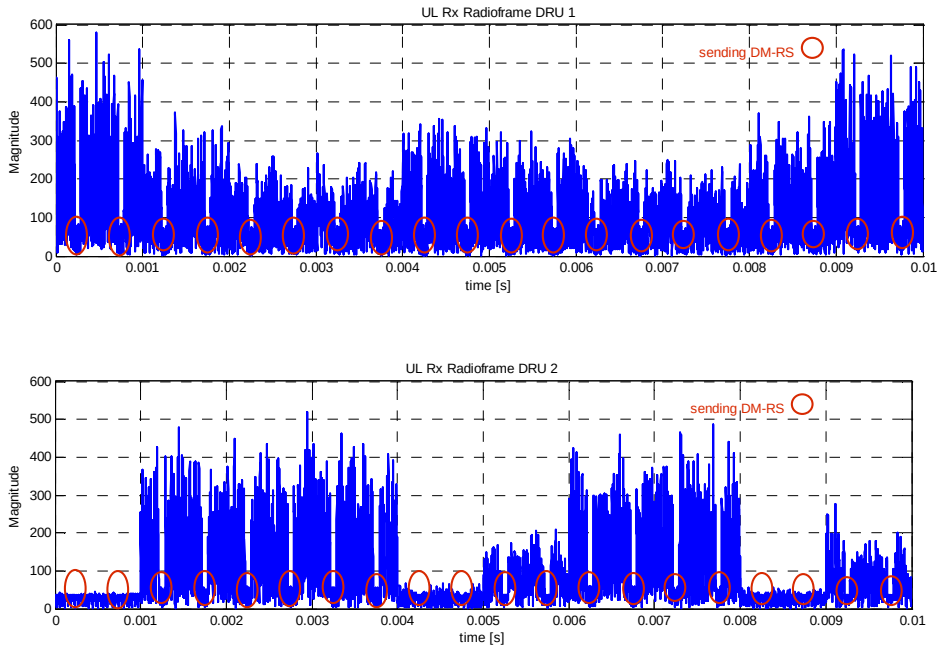


Figure C.8: The received signals at point C in Fig. 3 at DRU1 and DRU2 during one radio frame.

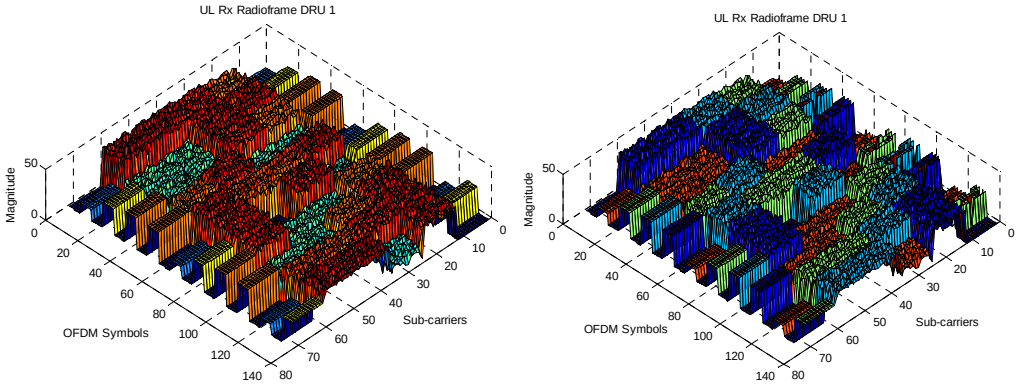


Figure C.9: The magnitude of de-mapped resource elements of received signal at point D in Fig. C.2 for DRU1.

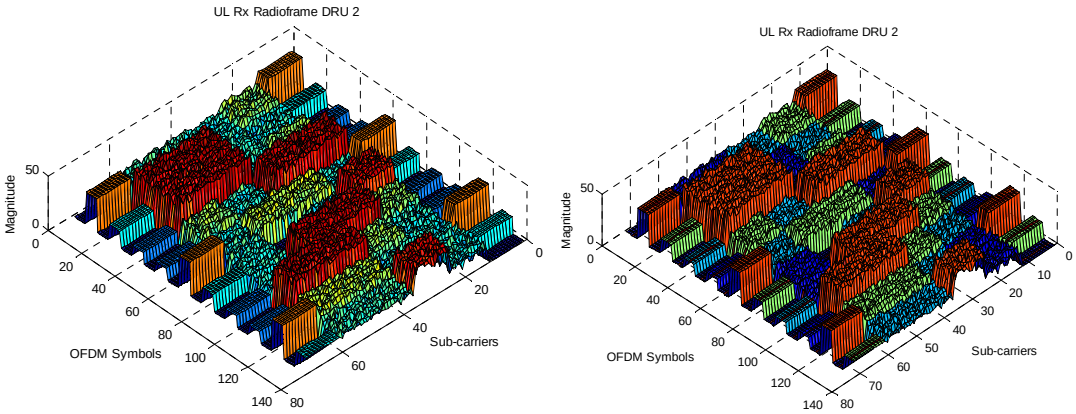


Figure C.10: The magnitude of de-mapped resource elements of received signal at point D of Fig. C.2 for DRU2.

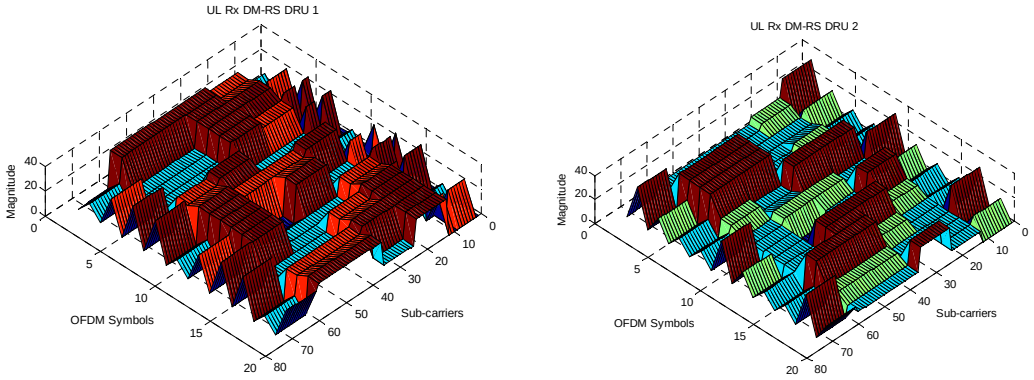


Figure C.11: The magnitude of de-mapped DM-RS resource elements of received signal at DRU1 and DRU2.