Addis Ababa University

Addis Ababa Institute of Technology

School of Electrical and Computer Engineering

On the Performances of User Association Enhancements in Dense Wireless Heterogeneous Networks

A PhD Dissertation Submitted to the School of Electrical and Computer Engineering, Addis Ababa Institute of Technology, Addis Ababa University in Partial Fulfillment of the Requirements of the Degree of Doctor of Philosophy in Communication Engineering.

By:

Dinkisa Aga Bulti

Supervisors:

Prof. Jyri Hämäläinen, Aalto University, Finland

Dr. Eng. Yihenew Wondie, AAiT, Ethiopia

March 26, 2023

Approval Page

Addis Ababa University

Addis Ababa Institute of Technology

School of Electrical and Computer Engineering

This is to certify that the dissertation by Dinkisa Aga Bulti, entitled **On the Performances of User Association Enhancements in Dense Wireless Heterogeneous Networks** and submitted in partial fulfillment of the requirement for the degree of Doctor of Philosophy (PhD) in Communication Engineering complied with the regulations of the University and meets accepted standards with respect to originality and quality.

Approved and signed by board of examining committee

	Name	**Signature**	**Date**
Dean, School of Electrical and Computer Engineering	Dr. Bisrat Derebssa		
Supervisor	Prof. Jyri Hämäläinen		2023-03-27
Supervisor	Dr. Eng. Yihenew Wondie		27/03/2023
External Examiner			
Internal Examiner			

Declaration

I, the undersigned, declare that this dissertation submitted to the School of Electrical and Computer Engineering, Addis Ababa Institute of Technology, Addis Ababa University in partial fulfillment of the requirement for the degree of Doctor of Philosophy (PhD) in Communication Engineering is my original work and has not been presented for a degree in any other university and that all source materials used have been properly acknowledged.

Name: Dinkisa Aga Bulti

Signature: _____

Date: 24/03/2023

This PhD dissertation has been submitted for examination with our approval as supervisors.

Prof. Jyri Hämäläinen

Name Signature

Dr. Eng. Yihenew Wondie

Name Signature

Acknowledgment

I want to start by thanking God Almighty for His provision and guidance that I needed to complete my task and succeed in this section.

Next, I would like to extend my utmost gratitude and appreciation to my supervisors Professor Jyri Hämäläinen and Dr. Eng. Yihenew Wondie for their guidance and continuous support throughout my Ph.D study. Their guidance helped me in all the time of research and writing this thesis.

I would also like to extend my gratitude to Dr.-Ing. Dereje Hailemariam, Dr. Beneyam H. Berhanu at Addis Ababa Institute of Technology and Dr. David Gonzalez for their technical support, discussion and guidance during the early stage of my PhD study. I would also like to thank the European International credit mobilty framework for supporting my stay at Sivas Cumhuriyet University, Turkey.

Last but not least, I would like to thank my friends and family for being my main sources of inspiration in all that I do. To my wife, Tizita and my kids, Robera and Milki, you have been a great source of joy and a much-needed break from the stress of research work.

Abstract

User Association (UA) plays a significant role in radio resource management of wireless communication systems. Currently, network densification and heterogeneity have already been identified as a feasible solution for the exponentially expanding data service demand. Hence, UA methods must meet different requirements in dense and ultra dense Heterogeneous Networks (HetNets). The load-imbalance due to transmit power difference between tiers and interference coordination challenges, the effect of serving node intensity on load sharing and achievable throughputs and the effort to satisfy certain users with high data rate demands are a few problems. Furthermore, the interconnected and complicated problems of service delivery are posed by the spatio-temporal dynamics in service demand and the mobility of User Equipment (UE).

This thesis takes a step-by-step approach to solving UA problems in dense and ultra dense HetNets. This research uses stochastic geometry tools, system level simulations, and realistic test case deployment simulations. Models were created for each scenario based on the load balancing, interference coordination, varied densification levels, heterogeneity, and user mobility.

The work's first contribution is a solution to the problem of load imbalance and interference coordination. The proposed method is simple to integrate into an existing HetNets network, and the results demonstrate effective load-aware association and adaptive interference coordination. A cell clustering-based load-aware offsetting and an adaptive Low Power Subframe (LPS) approach was developed. The solution allows the separation of UA functions at the UE and network server such that users can make a simple cell-selection decision similar to that in the Maximum Received Signal Strength (max-RSS) based UA scheme, where the network server computes the load-aware offsetting and required LPS periods based on the load conditions of the system. The proposed solution was evaluated using system level simulations wherein the results correspond to performance changes in different service regions. Results show that the method effectively solves the offloading and interference coordination problems in dense HetNets.

The second contribution of the research is on the coupled and decoupled User Association. It can be used as a guide for network operators to select the appropriate UA scheme for their network. The concepts of Poisson random networks were used to analyt-

ically obtain the relative densification levels for which we need the offloading, decoupled or coupled UA and validate the analysis with numerical and system level simulation of realistic network. The association window, where users choose to use decoupled association in terms of the relative intensity, transmit powers at each tiers and the Path Loss Exponent (PLE) of the propagation environment, is derived. Further, the ergodic rate expressions in order to study throughput performances in different densification regions, which can be computed numerically, are formulated . To validate the theoretical analysis, numerical, system level simulation and realistic network analysis were used. The analytical, simulation, and realistic test case results provide insights for the operators about the densification ranges, where to use coupled or decoupled association.

Finally, the research work focused on solutions for UEs with high data rate demands and mobility management. With Multiple Association (MA), user-centric clustering, control, and user-plane split usages were designed and investigated. Mobility management approaches in Long Term Evolution Advanced (LTE-A)/Fifth Generation (5G) and MA were used. The scheme attempts to separately treat UEs based on their speed by setting some predefined thresholds. In addition, a clustering approach, which produce virtual cells with which UEs gets associated was developed. Combining of MA with clustering enhances cooperation between most appropriate cells to serve a given UE. The findings indicate that the issues were addressed in an efficient and effective manner.

Key Words: LTE/LTE-A, 5G, user association, load balancing, HetNets, cell densification, service demand, cell range expansion, almost-blank subframe, inter-cell interference, clustering, downlink and uplink decoupling, multiple association, mobility management.

Contents

List of Figures

List of Tables

Acronyms

1G	First Generation
2G	Second Generation
3G	Third Generation
4G	Forth Generation
5G	Fifth Generation
3GPP	Third Generation Partnership Project
ABS	Almost Blank Subframe
ANR	Automatic Neighbor Relation
ASE	Area Spectral Efficiency
BBU	Baseband Processing Unit
BM	Brownian Motion
BOA	Bubble Oscillation Algorithm
BS	Base Station
CA	Carrier Aggregation
CAGR	Compound Annual Growth Rate
CAPEX	CAPital EXpenditure
CDF	Commulative Distribution Function
C-RAN	Cloud-based Radio Access Network
CoMP	Coordinated Multi-point
CoV	Coefficient of Variation
CP	Critical Point
CPs	Critical Points
CRE	Cell Range Expansion
D2D	Device-to-Device
DA	Dual Association
DL	Downlink
DPM	Dominant Path Model
DUDe	Downlink and Uplink Decoupled
eCoMP	enhanced Coordinated Multi-point
EE	Energy Efficiency
eICIC	enhanced Inter-Cell Interference Coordination
eNodeB	evolved NodeB
FeICIC	Further enhanced ICIC
GE	Grammatical Evolution
HD	High Definition
HetNets	Heterogeneous Networks
HSPA	High Speed Packet Access
ICI	Inter-Cell Interference

ICIC Inter-Cell Intereference Coordination

IoT Internet of Things

ISD Inter-Site Distance

ITU International Telecommunications Union

LA-OLPS Load-aware Offsetting and Adaptive LPS Configuration

LPN Low Power Node

LPNCR Low Power Node Center Region

LPNER Low Power Node Edge Region

LPS Low Power Subframe

LT Laplace Transform

LTE Long Term Evolution

LTE-A Long Term Evolution Advanced

MA Multiple Association

M2M Machine-to-Machine

max-RSS Maximum Received Signal Strength

MC Macro Cell

MCCR Macro-cell Center Region

MCER Macro-cell Edge Region

MIMO Multiple Input Multiple Output

mmWave milli-meter Wave

OFDMA Orthogonal Frequency Division Multiple Access

OPEX OPerating EXpenditure

P2P Peer-to-Peer

pdf probability distribution function

PDL Power Density Upper Limit

PGFL Probability Generating Functional

PLE Path Loss Exponent

PPP Poisson Point Process

QoS Quality of Service

RAN Radio Access Network

RAT Radio Access Technology

RF Radio-Frequency

RFA Reverse frequency Allocation

RHS right-hand side

RR Round Rubin

RRH Remote Radio Heads

RRM Radio Resource Management

RSRP Reference Signal Received Power

RSRQ	Reference Signal Received Quality
SC	Small Cell
SE	Spectral Efficiency
SG	Stochastic Geometry
SINR	Signal to Interference Plus Noise Ratio
SIR	Signal to Interference Ratio
SNR	Signal to Noise Ratio
SON	Self-Organizing Network
TDD	Time Division Duplexing
UA	User Association
UBKCA	User-Based K-means Clustering Algorithm
UDN	Ultra-Dense Networks
UE	User Equipment
UL	Up-Link
VNI	Visual Networking Index
WCDMA	Wideband Code Division Multiple Access
Wi-Fi	Wireless Fidelty
WIGIG	Wireless Gigabit
WLAN	Wireless Local Area Network

Chapter 1

Introduction

The demand for mobile data traffic has grown at an exponential rate in the past two decades and is likely to continue to do so in the future. According to the prediction by International Telecommunications Union (ITU), the global mobile traffic per month is estimated to be 543 Exabytes in 2025 and 4,394 Exabytes in 2030 without Machine-to-Machine (M2M) traffic [1]. Including M2M traffic (Fig. 1.1), the mobile traffic will be growing at an annual rate of around 55% between 2020-2030. The global mobile traffic per month would then be estimated to reach 607 Exabytes in 2025 and 5016 Exabytes in 2030.

The increase in mobile data traffic is primarily due to behavioral and sociological shifts in mobile device usage to smart phones, which consume more data than a basic feature phone. In 2017, a smart phone generated 10 times more traffic than a basic feature cell phone, according to Cisco Visual Networking Index (VNI) [2]. Furthermore, the rapid expansion of M2M connections and Internet of Things (IoT) applications is adding to the explosive development of mobile data consumption [3].

Lower-generation network connectivity (Second Generation (2G) and Third Generation (3G)) is being phased out in favor of higher-generation network connectivity (Forth Generation (4G) or Long Term Evolution (LTE), and soon Fifth Generation (5G)). The adoption of advanced multimedia applications will be facilitated by combining device capabilities with faster, greater bandwidth, and more intelligent networks, which will contribute to increased mobile and Wireless Fidelty (Wi-Fi) traffic. According to Cisco's Annual Internet Report, 4G connections will account for 46% of overall mobile connec-

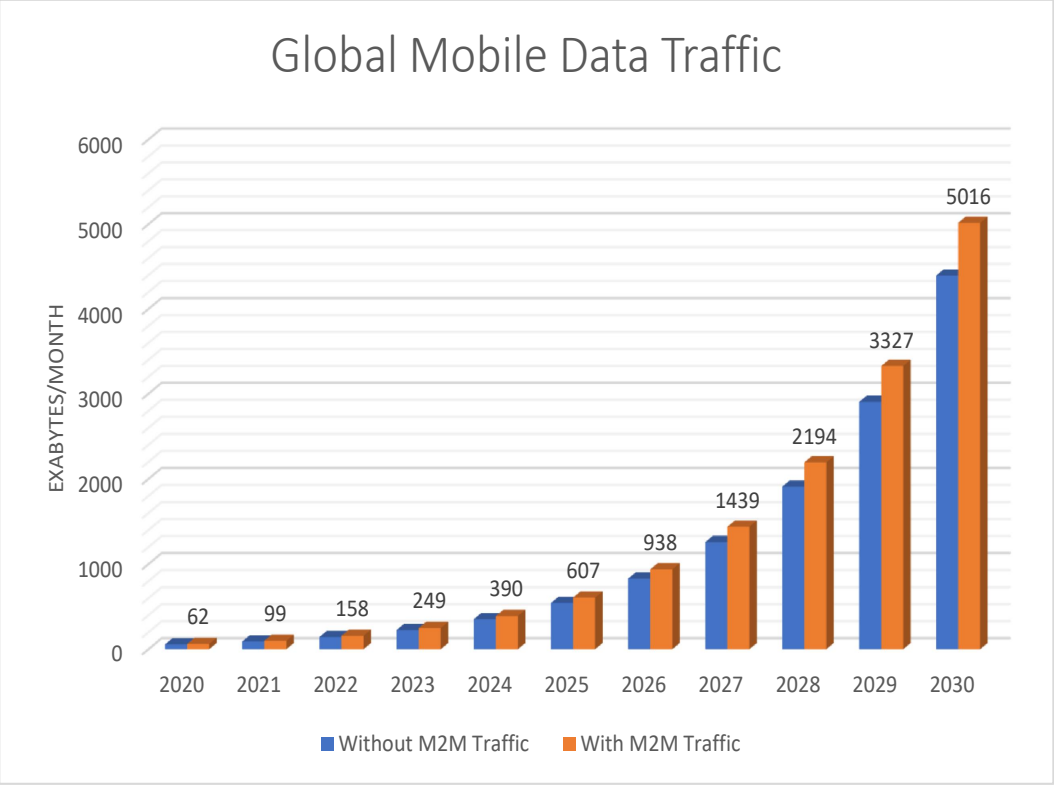


Figure 1.1: Global mobile data traffic forecast between 2020 and 2030 [1]

tions in 2023, up from 42% in 2018. At a Compound Annual Growth Rate (CAGR) of 10%, global mobile 4G connections will increase from 3.7 billion in 2018 to 6.0 billion in 2023. 5G connections, trials launched in 2019, are expected to more than triple in size from 13 million in 2019 to 1.4 billion in 2023 [4].

The forthcoming various innovative mobile applications and services (e.g. Social media, various High Definition (HD) and ultra-HD multimedia streaming services, billions of M2M connections, wearable systems, augmented and virtual reality) places diverse requirements in terms of throughput, delay, reliability, availability, energy efficiency and other diverse and critical topics which are required to be addressed by 5G and beyond-5G wireless systems [5, 6, 7]. Agiwal et al. in [6] provided an extensive survey of the various requirements and enabling technologies for 5G wireless networks. Amongst 5G technology enablers are: network densification (through dense HetNets or Ultra-Dense Networks (UDN)s), massive

Multiple Input Multiple Output (MIMO), milli-meter Wave (mmWave) communication, Cloud-based Radio Access Network (C-RAN), Self-Organizing Network (SON), and energy harvesting networks.

Therefore, operators of existing and future cellular networks rely on cell densification and heterogeneity to meet rising traffic needs. Dense HetNets comprises of Macro Cell (MC) operating with a large number of Low Power Node (LPN)s (e.g., small cells, relays, Remote Radio Heads (RRH)), Device-to-Device (D2D) and M2M communications which are used for provisioning of the Peer-to-Peer (P2P) services [7]. This issue is discussed in more detail in the following section.

1.1 Network Densification and Heterogeneity

1.1.1 Heterogeneity

As the demand for higher data rates increases, one of the solutions available to operators is to reduce cell size. By reducing cell size, the Area Spectral Efficiency (ASE) is increased through higher frequency reuse, while transmit power can be reduced such that the power lost through propagation will be lower. Additionally, coverage can be improved by deploying LPNs, where reception may not be good and offloading traffic from MCs when required. This solution has only been made possible in recent years with the advancement in hardware miniaturization and the corresponding reduction in cost.

Furthermore, changes to the functional architecture of the access network allowed data and control signals to tunnel through the Internet, enabling the LPNs to be deployed anywhere with Internet connectivity. LPNs can have different flavors, with low powered femtocells typically used in residential and enterprise deployments, and the higher powered picocells used for wider outdoor coverage or filling in MC coverage holes. The concurrent operation of different classes of base stations, macro-, micro-, pico-, and femto-cells, is known as Heterogeneous Networks [6].

Therefore, a cost effective way to handle the exponentially growing data traffic demand is the deployment of large number of LPNs overlying MC network gives rise to HetNets. From this, heterogeneity can be defined as deployments with a mixture of cells with different Downlink (DL) transmission power, operating on (all or partially) the same

set of frequencies and with overlapping geographical coverage. HetNets can also be an inter-working of different Radio Access Technology (RAT)s but operating in licensed or unlicensed frequency [8]. Hence, HetNets enable a more flexible, targeted, and economical deployment of new infrastructure versus tower-mounted macro-only systems, which are very expensive to deploy and maintain.

1.1.2 Network Densification

According to [9], *Network densification* is realized by increasing the density of Base Station (BS)s deployed in the given geographic area or the number of antennas per node, while ensuring nearly uniform distribution of users among all BSs. Network densification is quantified by the site density (sites/km^2) or Inter-Site Distance (ISD). However, how large the intensity of BSs per unit area for a network to be called *Dense*, *hyper-dense* or *Ultra-dense* remains vague.

Ge et.al., in [10] puts the density of BSs (or interchangeably called 'cells' in this thesis) in the 3G cellular networks to be about $4-5$ cells/Km^2. In the 4G cellular networks, like LTE-A, the density of MCs is approximately $8-10$ cells/Km^2. Furthermore, the density of 5G BSs is highly anticipated to come up to $40-50$ cells/km^2 in order to satisfy seamless coverage.

The deployment of LPNs has been a critical part of the 4G network upgrades and expansion. According to Nokia [11], the ongoing densification based on LTE network, cell densities of $10-30$ cells/Km^2 has already been common in many places. Nokia defines a UDN as a network with sites on every lamppost or having indoor sites less than 10m apart (see Fig. 1.2). This agrees with a definition in [12], which puts a typical average ISD to be as small as 10 - 30m. By 2025 or 2030, Nokia expects UDNs to be covering most urban indoor and outdoor areas with LPNs providing cell edge data rates of 100 Mbps to everyone.

Contrary to this, from the coverage probability and ASE analysis, Ding et al. [12] defined UDN quantitatively as measure of the density at which a network can be considered ultra-dense to be around 10^3 $cells/Km^2$ or more. Other definitions state that UDN is a network which contains more serving nodes or cells compared to the active number of users per-Km^2 area [13, 14]. In the second case, since there are more number of cells than

number of UEs in a given area, it is also expected that a significant number of cells may be turned OFF/ON at a time. However, most documents forecast that the number of users will still remain multiple of times greater than the number of cells for a foreseeable time [11]. The process of migration is taking the path of densification of existing LTE networks with 5G new radio as well as emerging shared spectrum radios [15].

In addition to the familiar horizontal densification, an increase of cell density in the vertical dimension (site infrastructures on high-rise building) will be significant deployment scenarios. The survey in [16] identifies the horizontal and vertical densifiifcation, which will be important in the realization of UDN in dense urban environment. With the believe that the demand for mobile data service continues to increase, a high-rise buildings in dense-urban centers will attract more cells and will create basis for the coming ultra-dense HetNets.

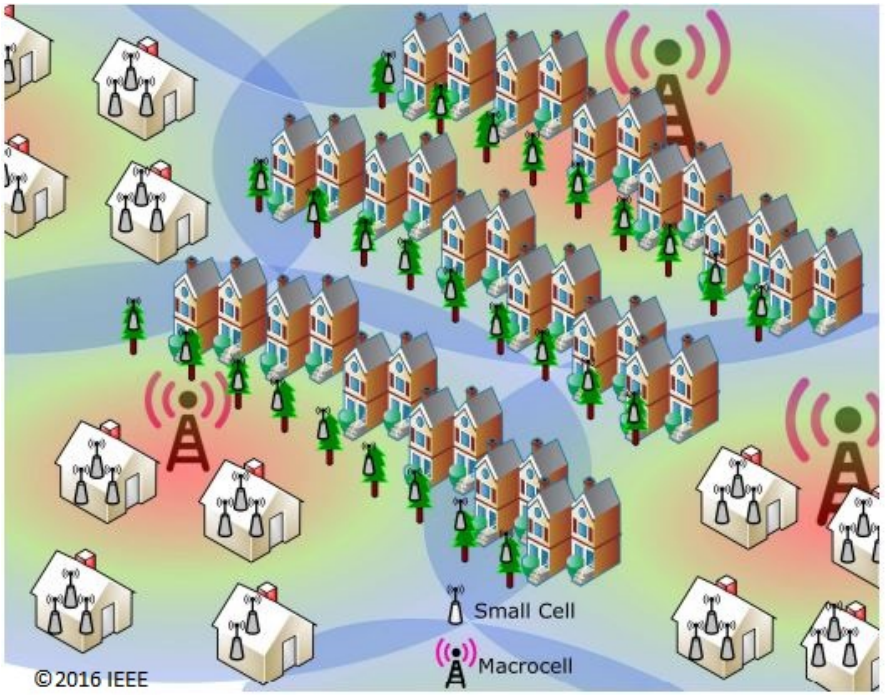


Figure 1.2: The network is densified by deploying small cells indoors in buildings and stores, and outdoors on trees, lampposts, and building walls. Small cell networks coexist with macro-cells, either in the same spectrum or on a dedicated carrier [16]

1.1.3 Densification in Different Regions

The regional variation in mobile data demand is also attributed to the technological solutions being proposed and used by respective operators in each regions or vice versa. In [17], by 2026, 5G mobile subscription is expected to be 5% in Sub-Saharan Africa while this will be 80% in North America (Fig. 1.3). While 5G and LTE subscriptions will continue to grow, High Speed Packet Access (HSPA) will remain the dominant technology with a share of over 40% in 2026, in Sub-Saharan Africa. Another forecast by Cisco's

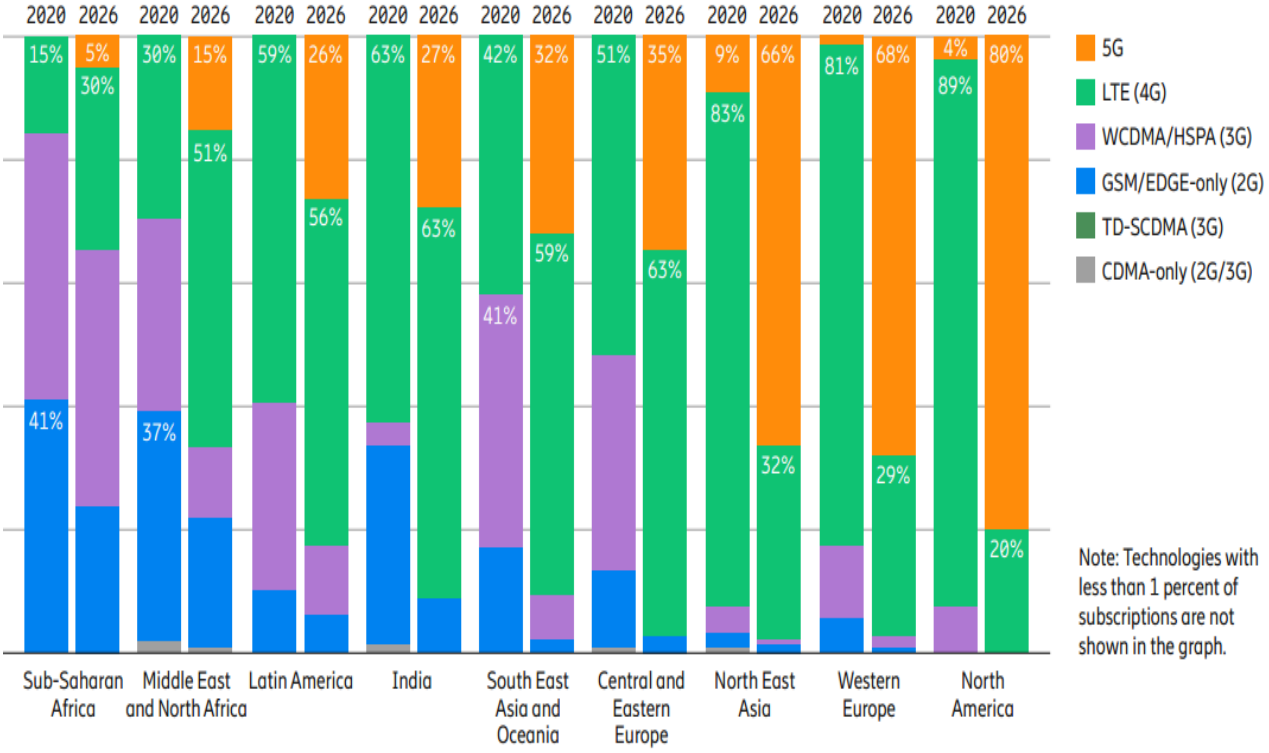


Figure 1.3: Mobile subscriptions by region and technology [17]

Annual Internet Report [4] shows that Middle East and Africa will have the highest share of its devices and connections (73%) on 3G and below in 2023. However, in 2023, North America will be the region with highest share of connections on 5G at 17%.

Furthermore, densification is to be characterized by the numerous deployment options [18] in the urban and rural environments. The urban scenario is driven by the need to spot coverage at outdoor coverage holes and capacity demand at hotspot areas. Therefore, unlike the MC that are mounted on rooftops or on radio towers, the LPNs deployments will be on building walls, lampposts and so on. In addition to this, the urban scenario is

expected to see areal wireless deployment based on drone LPNs which are also envisioned for 5G [19] and already attracted many research efforts. The rural deployment scenario typically differs in the coverage range of the LPN which is designed to be higher compared to the urban scenarios.

1.2 User Association

A *User Association (UA)* is defined as a technique to determine a cell(s) or BSs that provides service to the UE. In other words, it is a process performed when a UE joins the network (also called cell selection), or when a mobile station is on the move in idle mode (also called cell reselection) [20]. UA concerns which cell or BS a UE associates with in the Up-Link (UL) and DL directions. UA differs from *handover*, as the later happens when a mobile user in active call is crossing a cell boundary.

UA plays a vital role in implementation of performance enhancing technologies, load balancing, energy saving and latency reduction of a communication system. I.e., schemes used for UA will impact system performance by changing the load distribution among BS and affecting the interference levels.

Different wireless technologies have various standards on the UA schemes, procedures and criterion. Our discussion is limited to coupled, decoupled and multiple UA schemes in connection with network heterogeneity and densification.

1.2.1 Coupled Association

Using coupled UA, a UE is associated with one BS both in DL and UL. In MC-only or single-tier networks, the received power-based UA rule is the most prevalent one, where a UE will choose to associate with the specific BS, which provides the Maximum Received Signal Strength (max-RSS) [21]. For instance, variants of best Reference Signal Received Power (RSRP) [22] like best Reference Signal Received Quality (RSRQ) or best Signal to Interference Ratio (SIR) are used for UA in LTE network. The UA process is initialized by a local procedure by the UE according to the best detected RSRP, RSRQ or Signal to Noise Ratio (SNR) [20].

In HetNets, the transmit power disparity among serving nodes/cells needs effective in-

terference coordination schemes. Due to this, for multi-tier network like in HetNets with varying transmit power, the Cell Range Expansion (CRE) or biasing-based association, and interference coordination techniques like enhanced Inter-Cell Interference Coordination (eICIC) and Further enhanced ICIC (FeICIC) received attention [23, 24, 25, 26]. In CRE, a positive bias value is added to the received signal strength in order to offload more UEs onto the LPNs.

Therefore, in HetNets, user traffic is offloaded to the small cells whenever possible increasing the user throughput and capacity of the system by improving the ASE. Mobile offload exceeded cellular traffic by a significant margin in 2017. 54% of total mobile data traffic was offloaded onto the fixed network through Wi-Fi or femtocell in 2017 [2].

1.2.2 DL and UL Decoupled Association

As presented earlier, in conventional network, UEs are associated to the same BS both in UL and DL. This results in the problem of DL-UL asymmetry in coverage and capacity provisioning in HetNets deployment with different transmit powers. This becomes worse in the dense deployment because a UE may experience different propagation gains in UL and DL from nearby dense BSs, which calls for decoupled UA scheme for better performance. This scheme allows a UE to be associated to MC in the DL and to a LPN in the UL, which is called Downlink and Uplink Decoupled (DUDe) UA [25]. The need for DL and UL *Coupled* or *Decoupled* associations arise from the fact that the optimal DL associations are not necessarily optimal for UL performance.

1.2.3 Multiple Association

The association of a user to a single cell limits its maximum achievable data rate to the capacity of a cell. On the other hand, the cloud-computing trend and the bandwidth-hungry applications accelerate the need to even higher data rates than what could be offered by a single cell. Moreover, the UDN deployment scenarios introduced a different coverage environment in which a given user would be in close proximity to many cells [9]. This motivates the use of a different association scheme as a solution to distribute the traffic load of the user to multiple cells in a user's neighborhood. This is also called Multiple Association (MA) [27, 28]. In MA, a user associates itself to two or more LPNs

or MC forming what is called cluster of cells or *MultiCell* [27]. It is also defined as the ability for a UE to associate with multiple access nodes or cells, in its vicinity, at the same time [29]. This comes with the price of sophisticated UE, which must be equipped with many Radio-Frequency (RF) chains. However, in dense networks, the close proximity of the user to the serving BSs relaxes the requirements on these RF circuits [21].

1.3 Motivation

The move towards dense and ultra-dense network is not only led by the data demand but also governed by the investment potential in each countries. The pace in different nations are dissimilar because of the large digital divide. All countries by level of development have experienced significant growth in mobile broadband penetration rates since 2010 . However, large gaps remain over a decade later; the penetration rate in developed countries is double that of developing countries. In the case of Africa, mobile broadband penetration in 2020 was nearly 20 times greater than in 2010. While this allowed developing countries to narrow the gap with more advanced countries, there is still a significant mobile broadband divide [30].

In this regard, the developing economies like Ethiopia are lagging behind in embracing this huge move towards 5G with poor infrastructure. Though the developing nations' mobile subscription rate is the fastest in the world, due to other factors like investment potential and infrastructure, their move towards UDN will be slow. According to Ericsson's forecast, in Africa, only 5% 5G subscription will be made in 2026 [17]. However, the move is not something to be overlooked while it is not imminent compared the case for developed nations [31]. This shows that the evolution is going to be step-wise. Therefore, operators in this region needs cost-effective Radio Resource Management (RRM) options for their existing wireless networks while evolution is ongoing.

Furthermore, in developing economies, operators struggle with both urban and rural coverage and capacity provisioning. In particular, urban areas are with MC-only network, which will soon be overlaid with large number of LPNs [18]. This evolution might take years to advance into a full UDN and in the mean time such networks has to be cost effective and provide the expected capacity demand. The possibility of using coupled or

de-coupled UAs to solve the asymmetry problem between DL and UL performance can be harnessed as densification level continues to increase.

In other cases, operators are speedily advancing towards dense or ultra-dense networks. In dense urban (the case of developed nations) areas, the network density has already reached node intensity milestones [11]. For such scenarios, there is a need to study options for the operators on how and when to use different UA schemes and at which network intensities.

In addition, when user's demand is not to be satisfied by a single cell, options must be sought to use MA schemes. This is where HD and ultra-HD video streaming are prevalent. The realization of the capacity enhancing future technologies like Carrier Aggregation (CA) and Coordinated Multi-point (CoMP) needs to harness the benefits of MA techniques and algorithms.

Ray tracing tools are getting importance with more accurate propagation computation [32] and accurate digital maps are now a days available. This pushes to consider simulations of more realistic networks and evolution of UA schemes and other algorithms using computers with less effort. Doing this, we can get the detailed performance index of these irregular network with irregular user distribution at low cost before deployment.

1.4 Problem Statement

The dominant path of densification, as presented earlier, is the deployment of HetNets and gradual increase of LPN intensity. With this in consideration, the challenges of UA in dense, extremely dense, or ultra dense HetNets are summarized as follows.

- **Coupled UA in HetNets and irregular service demand distribution**

  A coupled UA in HetNets has been met with a challenge from two fronts: the irregular service demand distribution, and the heterogeneity of the network. Due to the transmit power differences in different tiers of HetNets and irregular service demand, there is typically a load imbalance among different serving nodes. So, the challenge in HetNets is how to dynamically select bias value which respond to data traffic distribution and heterogeneity of the network. To offload more traffic to the LPN and to coordinate the Inter-Cell Interference (ICI), the Third Generation

Partnership Project (3GPP) have developed the CRE, eICIC and FeICIC. Therefore, in HetNets, user traffic is steered towards the LPN whenever possible in order to enhance the user throughput and capacity of the system by increasing the ASE.

The static offsetting and fixed Low Power Subframe (LPS) configuration degrades performance due to the spatio-temporal load dynamics [26]. Therefore, load-awareness in offset value settings and adaptiveness in LPS configuration is very important. In addition, as a given MC can be overlying many LPNs, to which LPN that data should be offloaded has to be studied. Therefore, UA adaptability to irregularity in deployment and demand distribution needs to be investigated.

- **DUDe UA in dense HetNets under full load condition**

  The increasing intensity and heterogeneity of serving nodes drive the need for the study of coupled and decoupled associations. In the framework of HetNets, densification is heavy on the pico- or femto-tiers. Therefore, the relative intensity of nodes at each tiers impacts the network performance added to the different transmit powers. It could be asked for which densification levels and relative intensity of nodes can we use aggressive offloading with the established interference coordination techniques or DUDe association?

  In other words, under full-load condition, where every cell has at least one user to serve, the relative intensities of nodes at different tiers of the network can have impact on the load and performance. Operators need to know when to use coupled or DUDe association as their network evolve from MC-only to dense/ultra-dense network with comparably large number of LPNs.

- **UA in ultra-dense HetNets under non-full-load condition**

  Under non-full-load condition, there are more number of cells than number of UEs in a given area, and it is also expected that a significant number of cells may be turned OFF/ON at a time. To serve high data rate UEs, forming a group or cluster of serving cells is crucial and must be investigated. In UDN deployment, where cells are with smaller footprint, it is also important to consider the effect of UE mobility together with MA scheme.

- **Implementation in realistic networks**

  Real world assumptions and policies has to get due attention in order to achieve the best from UA scheme [33]. This is particularly considerable in propagation computations [34], service nodes topology and service demand modeling. Moreover, almost all UA algorithms resort to either max-RSS or simple biasing when it comes to the realization in real deployment. The reason is that most UA schemes are too complex for implementation in real network. Support for realization is crucial for operators to maximize their profit with minimum investment and upgrade their network.

1.5 Research Objectives

The objectives of the research work, categorized into four main objectives are:

Objective 1: To investigate the joint optimization of cell-specific offset values and number of LPS periods considering the load dynamics to maximize the user throughput. Specific objectives include: developing adaptive and load-aware UA scheme in HetNets which enables an easy implementation in existing system while providing effective offloading and interference coordination, developing network and link models which can be used for the performance analysis purpose of the network under consideration. The evaluation will be done on system level simulator.

Objective 2: Investigation of the critical densification levels for coupled and decoupled UA in UDNs. Specific objectives include: analysis of the densification intensity and its impact on load distribution and performance, to obtain the critical levels of densification at which cell load remains the same in the DL and UL, to drive the association window, where operators can choose between the coupled and decoupled association. Further, we formulate the ergodic rate expressions in order to study throughput performances in different densification regions.

Objective 3: Analyze performance of non-full load association or MA with mobility

support in 5G ultra-dense networks. Study the effects of MA on the user throughput and related challenges. Separate between high speed and low speed users' performances and propose solutions. Study the effects of UE mobility and cell clustering.

Objective 4: Carry out the performance analysis in realistic network. In addition to the system level simulation, UA schemes will be analyzed using ray tracing based real network simulators. The implementation of proposed UA schemes and algorithms in existing system will be discussed and recommended.

1.6 Methodology

A more general methodology is presented here, and specific and more detailed discussions are deferred until each of them are used in the respective chapters.

The foundation of this study is a desk research approach. A desk research strategy is used to obtain information from sources such as research journals, technical reports, and technological standardization documents, among others. As a result, the study is well-informed, and insight is gained through extensive desk research and collaboration with supervisors and colleagues. System level simulation, stochastic geometry tools, and test cases in more realistic networks employing deterministic propagation modeling tools are all employed side by side. Below is a general description of each of them.

1.6.1 System Level Simulation

As UA can be considered at a layer above physical layer of network protocol, a system level simulation can give sufficient performance evaluation results. This method was used in all the research works of this thesis as a result generation and visualization tool. A Matlab-based system level simulator was developed and validated for this purpose. The input to this simulator considers different spatio-temporal user demand distributions and signal propagation maps. The signal maps can be generated using different propagation models or can be imported from signal map computational tools. Typical snapshots of UE distribution in the deployment area and the deployment scenarios for different analysis and setups are provided in details in respective chapters. Algorithms and schemes were

developed as part of the simulator, where we used a modular approach in its development.

1.6.2 Stochastic Geometry

One can view a wireless communication network as a collection of nodes, which can in turn be transmitters or receivers (depending on the network considered, nodes may be UE, BSs or cells in a cellular network, or access points in a Wi-Fi network) [35]. However, we refer to nodes (in this thesis) as BSs or cells and user stations as UEs.

The hexagonal cell topology model is the most studied and more suited for MC-only networks. But, irregular deployments in HetNets whether dense or ultra-dense needs appropriate model for capturing the randomness of the network. In this regard, we made use of the following stochastic geometry tools.

Considering wireless communication network as Poisson random network, the Poisson Point Process (PPP) [35, 36, 37] was used to model cell locations (in two-dimensional space) for dense and ultra-dense HetNets in the study of coupled and decoupled association schemes. The link models and ergodic rates were hence drawn from stochastic geometry. In addition, the user demand distribution was modeled as PPP.

The visualizations of coverage shapes and topologies are also important. Poisson Voronoi diagrams or tessellations [38] were used to display cell topology in UDN networks for the study of DUDe and MA.

In modeling user mobility, Brownian motion [39] was used in connection with the study of MA. The detailed discussions are provided in respective chapters, where each of this tools are used.

1.6.3 Test Cases in Realistic Networks Simulators

Propagation computation and network planning tools enable the emulation of real networks, especially, when ray tracing [32] and accurate digital maps are used. In this regard, real deployment scenarios were considered, where the existing MC network is overlaid with LPNs. The setup and configurations, signal map and path loss computations were done using WinProp: Propagation modeling and network palnning software suite [40].

WinProp employs a simplified ray tracing model called Dominant Path Model (DPM). In standard ray tracing the contributions of all rays are superposed to obtain the received

power. But in most cases only 2 or 3 rays are contributing more than 98% of the energy [41], i.e. by focusing on these dominant rays, the accuracy can be sufficient and guaranteed.

In addition to the use of DPM, the WinProp software suite facilitates the use and edition of digital map, two- and three-dimensional antenna patterns, which makes it best emulator of the real environment. Further, there are many indoor propagation models including exponentially decreasing model with rich database for wall penetration losses.

1.7 Publications

As a result of this study, the following research outputs were published in peer-reviewed journals and conference proceedings, with parts of the findings included.

Journal Articles

P1 Dinkisa A. Bulti, Jyri Hämäläinen, Beneyam B. Haile, Yihenew Wondie, Clustering-based adaptive low-power subframe configuration with load-aware offsetting in dense heterogeneous networks, Digital Communications and Networks, Volume 8, Issue 5, Pages 843-852, October 2022.

P2 Dinkisa A. Bulti, Yihenew Wondie, On the performance of multiple association with mobility support in 5G ultra-dense networks: realistic network simulation, IET Communications, Published on-line, Jan. 28, 2023, DOI: 10.1049/CMU2.12576.

Under review

P3 Dinkisa A. Bulti, Yihenew Wondie, Investigation on the Critical Densification Levels for Coupled and Decoupled User Association in Ultra-dense Networks, International Journal of Wireless Information Networks, Springer,

Conference Papers

1. Beneyam B. Haile, Dinkisa A. Bulti, Bekele M. Zerihun, On the Relevance of Capacity Enhancing 5G Technologies for Ethiopia, Ethiopia ICT Expo 2018.

2. Dinkisa A. Bulti, Michael M. Admassu and Dereje H. Woldegebreal, "Performance of Inter-cell Interference Coordination with Frequency Reuse Techniques in Realistic LTE Network", Korea and Ethiopia International ICT Conference, May, 2016.

3. Dinkisa A. Bulti, Dereje H. Woldegebreal, David González G, Beneyam B. Haile, Jyri Hämäläinen, User Association and Load Balancing in Long Term Evolution Network in Addis Ababa, Ethiopia, IEEE Africon 2015, DOI: 10.1109/AFRCON.2015.7332007.

1.8 Thesis Outline

The thesis is organized in six chapters. The first chapter presented the background, definition and discussion of important terminologies, motivation for the research, the problem statement, objectives, and the methodology. The state-of-the-art UA techniques in dense and ultra-dense networks are presented in Chapter 2. Possible UDN deployment scenarios and related UA optimization issues are covered. Specifically, it presents different state-of-the-art UA schemes in HetNets and UDNs.

In Chapter 3, the load-aware and interference coordination efforts are presented. The joint optimization of offsetting and interference coordination schemes are the core part of the chapter. It focuses on how the clustering based adaptive LPS configuration and offsetting was developed, evaluated and discussed.

The impact of relative densification intensities of nodes at different tiers on load distribution and performance is presented in Chapter 4. It formulated the intensities at which an operator can use offsetting with LPS configuration, coupled or decoupled association. Numerical evaluations, system level simulations and realistic test case studies were carried out. The the results were discussed for different scenarios and conclusions were made.

In Chapter 5, we present the concepts and analysis of MA. The user and control plane splitting which is crucial for the dual and multiple associations are presented. Then, the utilized system model, user-centric clustering and mobility support schemes were developed. Both system level simulation and realistic test case simulations were used for the performance evaluation. Conclusions and recommendations were also provided on how the proposed UA schemes were easily implemented in the real networks. Finally, the work is summarized and recommendations are forwarded in the last chapter.

Chapter 2

State-of-the-art User Association in Dense HetNets

This Chapter deals with the state-of-the-art UA in dense HetNets and UDN deployments along with the UA concerns and optimization objectives. It first treats the densification scenarios to indicate the possible deployment strategies. Secondly, the objectives and metrics used in performance measurements are presented. Thirdly, the issues of irregular user distribution and un-planed deployments were dealt with load balancing concerns. Fourth, UA schemes in these different network evolution cases are presented. Lastly, the issues that needs consideration for intelligible, efficient and effective UA are described.

2.1 An Overview of Possible densification paths

The 5G communication system integrates various radio technologies, including LTE evolution, the 5G New Radio, and Wi-Fi-based technologies [**?**]. As a result, the new 5G radio will support low-latency, beam-based channels, massive MIMO with a large number of controllable antenna elements, scalable-width sub-channels, CA, C-RAN capability, and dynamic coexistence with LTE. This is due to LTE's ability to improve with CA, higher-order MIMO, standardized operations in licensed and unlicensed spectrum, vehicle-based communications, and LPN enhancements such as FeICIC, SON, and enhanced Coordinated Multi-point (eCoMP) capabilities.

Various technological advancements are expected to support densification, without

which capacity enhancements are not possible. As a result, network densification may be possible using the following technology combinations, [16, 42].

Densification through Small Cell (SC) deployments

Densification through use of SCs or LPNs with full functionalities and Relays. These LPN can be classified as indoor or outdoor SC based location and power levels; and also open, closed or hybrid access SC based on their access model which will be the integral components of the next 5G network. The high energy efficiency and smaller CAPital EXpenditure (CAPEX) and OPerating EXpenditure (OPEX) [43] is making this option a viable alternative suitable in the future network. This option has already been in the deployment phase and has constituted the today's HetNets [44, 21] except the level of densification or the expected inter-site distance in UDN which could be much closer than we experienced yet.

The densification is to be characterized by the numerous deployment options [18] in the urban and rural environments. The urban scenario is driven by the need to spot coverage at outdoor coverage holes and capacity demand at hotspot areas. Therefore, unlike the MC that are mounted on rooftops or on radio towers, the LPNs deployments will be on building walls, lampposts and so on. In addition to this, the urban scenario is expected to see areal wireless deployment based on drone LPNs which are also envisioned for future 5G [19] and it is already attracted many research efforts. The rural deployment scenario typically differs in the coverage range of the LPN which is designed to be higher compared to the urban scenarios.

The survey in [16] identifies the horizontal and vertical densifiifcation which will be important in the realization of UDN in dense urban environment.

Massive MIMO

Massive MIMO can potentially increase the network capacity by multitudes of magnitude [45, 46] which makes it a promising candidate technology for future wireless systems. In [47], it was shown that the energy efficiency of large scale SC Networks is higher compared with massive MIMO. Due to this, achieving energy efficiency in massive MIMO needs special arrangements. It could be massive MIMO at MC and single antenna systems at the LPNs.

The work in [42] also proposed a "Phantom cell" concept in which a non-fully functional

LPNs serves the user data communication in control- and user-plane split mode. The phantom cell should be backed/linked to the MC or backbone network through high capacity link which could possibly be massive MIMO link.

With ultra-densification and massive MIMO enabled evolved NodeB (eNodeB)s, UA needs special treatment [45].

mmWave link

The need for high bandwidth led to the use of higher frequencies in the milli-meter Wave (mmWave) spectrum range [48, 49]. This abundant spectrum has become the focus of major research efforts in recent years. It could serve the access link between UE and the LPN like in the case of Wireless Local Area Network (WLAN), as in Wireless Gigabit (WIGIG) alliance [50] or could be used for a high capacity non-ideal backhauling option [51] for the dense and fully functional SCs.

C-RAN

In the C-RAN architecture, the BS is broken into centralized signal processing and management or the Baseband Processing Unit (BBU) which is pooled at one geographic area and the analog radio access units, referred to as Remote Radio Heads (RRH)s which is spatially distributed [7, 52].

C-RAN lowers the cost of baseband processing and reduces the power consumption by performing load balancing and cooperative processing of signals from several BSs. Therefore, the BBU pool centralizes all the processing and maximizes energy efficiencies. The BBU pool are connected to the RRH with high speed fiber-optic links which is also called front-haul communication link. Densification through C-RAN architecture advantages the centralized control mechanisms' implementation of algorithms.

2.2 UA Optimization Objectives/metrics

User Association can be seen as multi-objective optimization problem due to need for load balancing, energy consumption or overcoming the backhaul constraints. In a broader category, UA goals can be divided into Spectral Efficiency (SE) and Energy Efficiency (EE) [9] optimizations. However, in the future UDN, we observe that further number of constraints and objectives involve and joint optimization needs as maximizing one objective

has a potential to affect the other objectives [53]. These objectives can be decomposed into the performance metrics or optimization criteria like fairness, data rate, aggregate throughput, delay, energy efficiency and minimum backhaul overhead as presented below.

- Fairness: Fair resource utilization has been a major research area in the past couple of decades related to the wireless network [54, 55]. Load balancing try to show some fairness index level.

- Throughput: The aggregate/system throughput can be measured in terms of the ASE and represents the system/network capacity [54, 55, 56, 57]. Many others focus on the cell-edge throughput improvements [24, 56].

- Latency: There is a stringent latency requirements in 4G, and even more in 5G, wireless systems. UA plays important role of identifying the node which could minimize the latency. The work in [56] attempted to reduce the delay along with other performance metrics.

- Energy efficiency: In the future UDN, we are to observe massive connection of devices with stand-alone batteries or connected to renewable energy supply and also energy harvesting techniques will be employed. Bottai et al. [58] studied the energy-efficient UA algorithm, i.e., an association algorithm that tries to maximize the energy saving of the LTE access network, namely the *Min-Energy association* algorithm. The ultimate goal of this technique is to switch off as many eNodeBs and switches as possible, while keeping their utilization to the highest level. The result shows that the energy saving algorithm saves energy but at a cost of reduced throughput. Hybrid energy options and optimization [59, 60] were also considered with green-energy allocations in HetNets which needs further inquiry in the UDN environment.

- Backhaul overhead: As presented in previous section, backhauling is the main challenge in UDN. Therefore, intelligible UA is needed to ease the backhauling problem and a significant number of works have attempted to solve this problem.

2.3 Demand Distribution and Load Balancing

2.3.1 Data Traffic Demand Irregularity

The prolifrating data traffic demand is accompanied by irregular demand distribution in space and time. Studies have shown that indoor traffic demand constitute up to 70% [61] of the total mobile data demand. Network users are usually concentrated at social attractions such as residential and office buildings, shopping malls, and bus stations [62] which represents the *Hotspot* areas. Also, these indoor users experience the worst signal quality due to high wall penetration loss.

Modeling the full nature of the spatio-temporal dynamics in traffic demand distribution is challenging due to the large number of parameters that involves. However, in a number of literatures (eg. [62, 63, 64, 65]), we can find models to represent a few aspects of the distribution ignoring the others. For instance, models for the request/packet arrival process, service time, the spatial distribution in urban environments, and so on.

The spatial dynamics in service demand distribution results from the diverse topography and human-social activities. α-Stable model [63] was used to represent the spatial traffic demand distribution. The spatial service demand is reported to follow Weibull distribution [65]. The PPP is also used in many literatures as a candidate to represent both the time and the spatial service demand distribution [62, 63].

Though, PPP model may be a fitting one for the BS locations, it is less adequate for the UE locations mainly due to the fact that the model is not adjustable (tunable) to represent the severity of the heterogeneity (non-uniformity) in the UE locations [62]. However, this do not disapprove that there is a positive relation between service demand irregularity and network topology as seen in [66]. If the altitude of UE locations are assumed constant or the same, the heterogeneity in the spatial distribution can be captured as two-dimensional point pattern $U \epsilon \Re^2$. The desired statistics of the traffic are the mean (μ) and the Coefficient of Variation (CoV). Based on the CoV, the two-dimensional distribution of the spatial traffic can be modeled as Poisson process (i.e., exponentially distributed inter-UE distance), sub-Poisson processes, super-Poisson processes [62].

2.3.2 Load Balancing

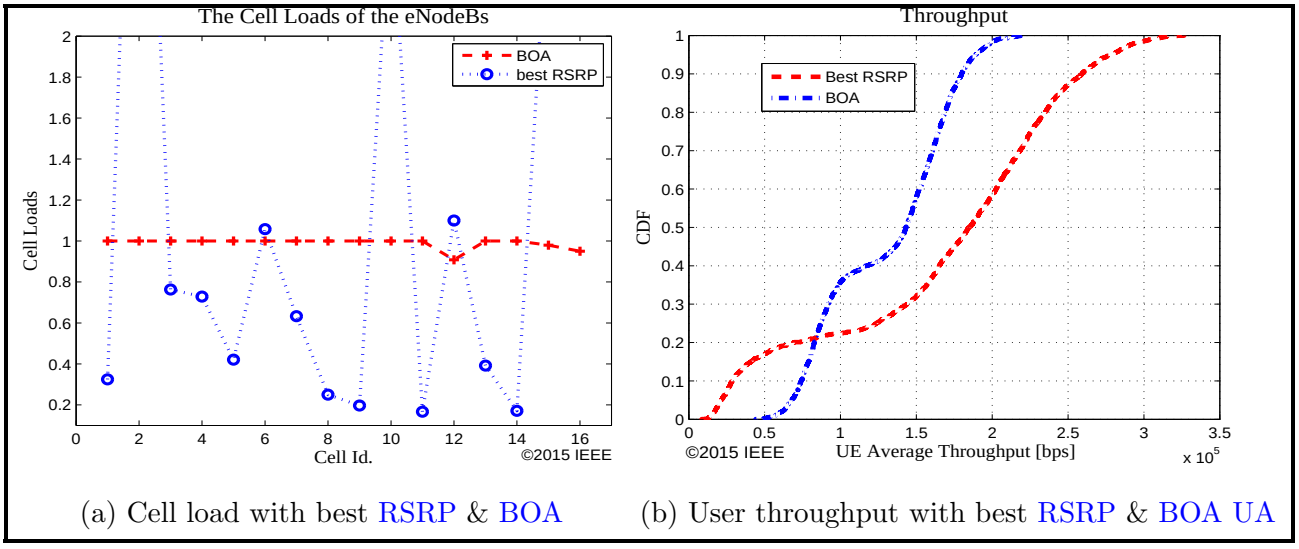
Due to the above issues related to the demand variations in time and space, the impacts of UE and traffic demand distribution on network performance is complex and interlocking as studied in [67, 68]. Even if the spatial distribution of BS locations is originally planned to achieve compatibility with the spatial variations of service demand, the variation in time dimension requires adaptable UA and load balancing with efficient interference management.

Load balancing in a wireless network determines the performance and Quality of Service (QoS) of the system. Therefore, the irregularity in service demand distribution in both space and time requires an efficient and effective load balancing among the BSs to get the most from the increasingly irregular topology of HetNets. The effect of service demand distribution was studied by Kim et al. in [67].

One of the main criteria for an adaptive load balancing algorithm is its load-awareness and dynamic response to service demand variations. A load-aware association considerably improves resource utilization [24]. A widely accepted approach for load balancing is to achieve load balancing by formulating the problem as a convex optimization problem [23]. Load balancing using cell breathing [69], Bubble Oscillation Algorithm (BOA) [70, 71] and practical algorithm with low complexity for LTE network as presented in [72] were among the attempts. The cell breathing techniques [69, 70, 71] were used in different RATs and scenarios. The authors, in [71], have compared the best RSRP with geographic load balancing BOA [70] in macro-cellular LTE network. The result shows that received signal based UA demonstrates poor performance, especially at cell-edge under irregular service demand distribution compared to BOA (see Fig.2.1a and Fig. 2.1b).

2.4 UA Control Mechanisms

The need for globally optimal [20] UA vs. the latency requirement can result in inter-cell or backhaul communication overheads. Therefore, the adopted control mechanism in the design of UA scheme heavily affect the computational complexity, the signaling overhead, and the optimality of user association algorithms. Densification by small cells, massive MIMO and mmWave or C-RAN networks, present their own different effects on the effi-



(a) Cell load with best RSRP & BOA (b) User throughput with best RSRP & BOA UA

Figure 2.1: (a) Cell load and (b) user throughput for max-RSS and BOA based UA [71]

ciency and effectiveness of UA algorithms and corresponding control mechanism. Broadly speaking, there are three different control mechanisms [53]: centralized, distributed, and hybrid.

In **Centralized UA**, decision is made at central node. So, information and decision signals has to be communicated between the ultra-dense peripheral nodes and the central node which incurs communication overhead. With its proper architecture, C-RAN lends itself for the efficient implementation of UA algorithms.

In contrary, the **Distributed UA** distrbutes the UA decision making to the peripheral access nodes. The user-centric cell selection schemes can advantage this mechanism so that latency can be reduced. In a UDN based on densification by small cells, massive MIMO and mmWave link networks, this control mechanism is preferable because of their architectural appropriateness for distributed implementation of UA algorithms.

Hybrid UA: The Hybrid control relies on a compromise approach, which combines the advantages of both centralized and distributed control. For instance, the power control at the BS may rely on using a distributed method, whereas load balancing across the entire network could be implemented in a centralized manner.

2.5 User Association Schemes

This section presents the UA schemes following the step-by-step advancements due to the growing and varying requirements of different technological progresses. It also detailed the load balancing and interference coordination efforts together with UA in HetNets, DUDe, control- and user-plane split and the need for multiple association.

2.5.1 Received Signal Based UA

Received signal strength or signal quality is used to associate UEs to a cell. In existing systems, the received power based user association rule is the most prevalent one, where a user will choose to associate with the specific BS, which provides the Maximum Received Signal Strength (max-RSS) or best signal quality in a single-tier network [21, 73]. For instance, in LTE, the best RSRP is the default cell selection criteria and it assumes that the eNodeB with the maximum RSRP provides the best performance for a UE. The Best RSRP-based UA exploits the 3GPP's LTE physical layer measurements that are available at the UE [74].

Let G_j be antenna gain of j^{th} BS, G_i be antenna gain of the i^{th} UE located at the edge of the MC or LPN, $L_j(R)$) —path-loss from j^{th} cell to i^{th} UE when R is the separation distance between the transmitter and receiver, and $\mathcal{S}$ represents the normally distributed shadow fading term. The received power, $R_{i,j}$ from transmitting j^{th} BS at the cell edge is given as:

$$R_{i,j} = 10\log[P_t(j) \cdot G_j \cdot G_i \cdot L_j(R) \cdot \mathcal{S}]. \tag{2.1}$$

Here, $P_t(j)$ is the transmit power of the eNodeBs which could assume different values in HetNets deployment at different tiers. The user association with max-RSS scheme is given by:

$$RSS_i = \arg\max_{j \epsilon \mathcal{N}} \{R_{i,j}\} \tag{2.2}$$

Variants of best RSRP [22] like best RSRQ or best SIR can also be used for UA in LTE network. The user association process is initialized by a local procedure by the UE according to the best detected RSRP, RSRQ or SINR [20].

2.5.2 Offsetting or Bias-based UA

In a dense and irregular deployment [66], the challenges in UA further increase because of the transmit power disparities between different tiers of the network. With received signal based UA scheme, a few UEs associate to the LPNs which will further exacerbate the load imbalance on top of the effect of irregular service demand distribution. Different load balancing algorithms have been proposed including the Log-linear learning algorithms used in [75] for load balancing in HetNets. Kuang in [76] developed a tabu search algorithm to solve the mixed-integer non-convex problem for the joint user association and multi-cell frequency allocation problem by treating the frequency allocation as frequency partitioning among multiple patterns. However, frequency partitioning will not continue be a viable option as Wideband Code Division Multiple Access (WCDMA) and LTE networks have already deployed with universal frequency reuse for higher spectrum utilization.

Due to this, for multi-tier network like in HetNets with varying transmit power at different tiers, the CRE-based or bias-based UA scheme (see Fig.2.2) has got greater attention [23, 24, 25]. In CRE, a positive bias value multiply the received signal strength in order to offload more UEs onto the LPNs. To formulate expression for UA scheme we use the 3GPP's definition of CRE [77, 78], in which a certain positive bias or offset value ($P_{bias}(j)$) in dB is added to the actual received signal strength of the LPNs and the user association is performed differentiating the LPNs with their non-zero bias values. So, again the max-RSS selection criterion is:

$$RSS_i = \arg\max_{j \epsilon \mathcal{N}} \{R_{i,j} + P_{bias}(j)\} \tag{2.3}$$

The challenge in HetNets is how to dynamically select the offset value, which respond to data demand distribution. In heterogeneous UDNs, irregular cell location overlaid with large number of LPNs and the irregular demand distribution further increase the challenge on how to choose the offset value, and coordinate the Inter-Cell Interference (ICI), which in turn affect the SINR distribution and hence, the performance.

Adaptive bias selection [79], Q-learning to select bias values [80], cell power and bias selection using Grammatical Evolution (GE) in order to achieve load balancing in HetNets assuming constant traffic per UE [81] were proposed approaches in order to design efficient CRE schemes. However, in [81] it is noted by the authors that realizing GE for on-line

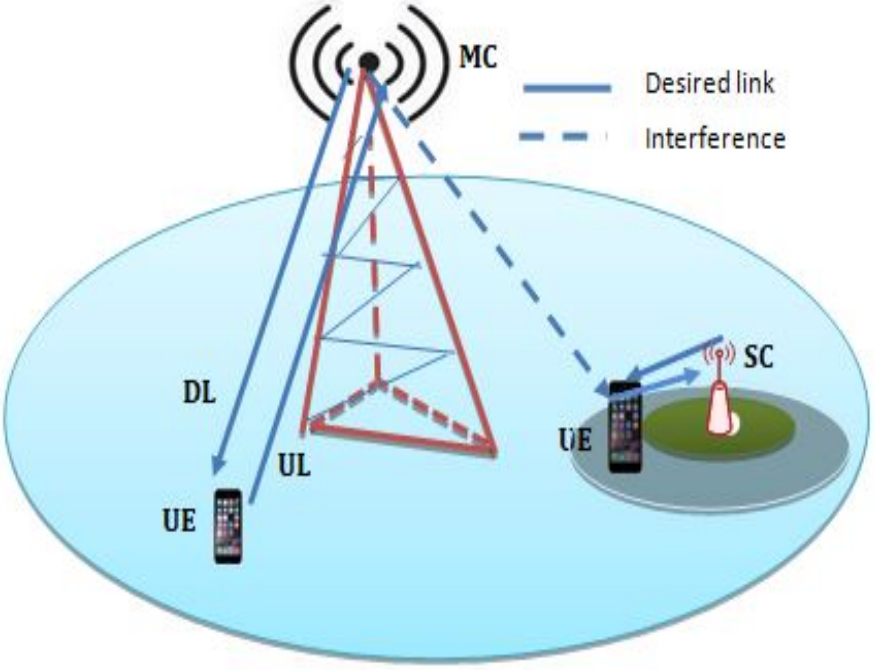


Figure 2.2: Cell Range Expansion

deployment is impractical as it is too slow to converge when used in a genetic algorithm -like way to evolve parameters.

The use of CRE to push more UEs towards the LPNs along with Inter-Cell Intereference Coordination (ICIC) schemes, found to enhance the performance especially at the cell-edge and hotspots areas [82, 83, 54] but results in further increase of the asymmetry between the DL & UL and remain in sub-optimal performance.

To reduce the inter-tier interference due to this asymmetry, time domain orthogonalization using the Almost Blank Subframe (ABS) was introduce [82]. In an ABS scheme, some sub-frames are left blank (except the power for control signaling) by the unbiased MC and off-loaded users are associated with the LPNs and served by these sub-frames. In 3GPP, ABS is enabled by the eICIC that provides means for macro- and pico-cells to time-share the radio resources for DL transmissions [83, 84].

Though, the use of ABS improves the system performance, it significantly degrades MC's total throughput. This problem is addressed through encouraging the Low Power Sub-frames to be re-used in MC center-region enabled by the FeICIC [26, 85], which was termed as the re-use of the Low Power Subframe (LPS)s. An effective combination of CRE and LPS can potentially enhance the network performance with appropriate bias

setting and dynamic LPS ratio configuration.

The joint UA and ABS proportion was devised by Jin et al. in [86]. Since this joint optimization problem is non-convex and NP-hard, decoupling it into two sub-problems: UA and ABS allocation, then the problem was solved iteratively to find the sub-optimal solution. Abbas et al. [87], employed a Reverse frequency Allocation (RFA) scheme with load balancing to reduce ICI from the unbiased macro-cell. In the RFA scheme, the UL and DL transmission spectra are reversed between the LPNs and MCs in a multi-region HetNets.

The work in [88] proposed a load-aware network selection approach applied to automated dynamic offset in HetNets. The authors investigated the properties of a hierarchical Bayesian game framework, in which the MC dynamically chooses the offset from the state of the channel in order to guide users to perform intelligent network selection decisions between MC and LPN. In [89], ABS ratio optimization is attempted iteratively and distributed UA and semi-distributed load balancing schemes between MC and LPNs were proposed. Kamel et al. [55] proposed an approach called ABS offsetting to reduce the blanking rate at the LPNs while preserving the required optimal blanking rate at the MC.

Table 2.1 summarizes literatures on max-RSS and bias-based UA schemes. In all these attempts, the joint consideration for adaptive LPS enhancement using FeICIC and load-aware offsetting is missing.

2.5.3 DUDe User Association

In the traditional network, UEs are associated to the same BS in both UL and DL. As a result, DL-UL asymmetry on coverage and capacity provisioning in HetNets deployment with different transmit powers between UEs and BSs arises. This becomes worse as the density of the LPNs increases, because an UE may experience different propagation gains in the UL and DL from nearby dense BSs, necessitating the use of an DUDe UA scheme for improved performance.

DUDe is a recently proposed disruptive technique for reducing the UL and DL imbalance problem that occurs in HetNets due to the large transmit power disparities between macro and small cells.

This scheme allows a UE to be associated to a MC in the DL and to a LPN in the UL.In

Table 2.1: Summary of max-RSS and Bias-based UA Schemes

UA Scheme	Control Mechanism	Objectives	Interference coordination	Topology	References
max-RSS	Distributed	Fairness, throughput	–	MC-only network	[70, 71, 72]
Bias based UA	Centralized	Long-term network throughput	Frequency Reuse Pattern selection	LPNs based HetNets	[76]
Bias based UA	distributed	sum rate &min rate, Fairness	–	LPNs based HetNets	[23, 24]
Bias based UA, Adaptive	Hybrid	Fairness	–	LPNs based HetNets	[79, 81]
Bias	–	Resource utilization, Fairness, Throughput, SE	eICIC	LPNs based HetNets	[84, 82, 83]
Bias	Distributed	Aggregate Throughput	FeICIC	LPNs based HetNets	[26, 85, 55]

other words, the received signal-based association scheme is employed in both directions. The work in [90] proposes and analyzes the DUDe user association framework under a two-tier heterogeneous network. The UL and DL coverage area for DUDe based UA is depicted in Fig.2.3.

The benefits of DUDe are discussed in detail by Boccardi et al. in [91]. These are: increased SNR and reduced transmit power, improved UL Interference Conditions, improved UL Data Rate, different load balancing in the UL and the DL, and low deployment costs with Radio Access Network (RAN) centralization. In [92], a coverage analysis with interference coordination was performed, and DUDe with RFA employment outperforms coupled UA.

Adapting the UL and DL wireless resources to meet the asymmetric and dynamically changing UL and DL traffic loads is becoming increasingly important. The work in [56] proposes a load-aware DUDe UA scheme under which UEs are associated to LPNs for both UL and DL based on both distance and load for Time Division Duplexing (TDD) small-cell networks. Improvements in throughput and delay were reported when compared

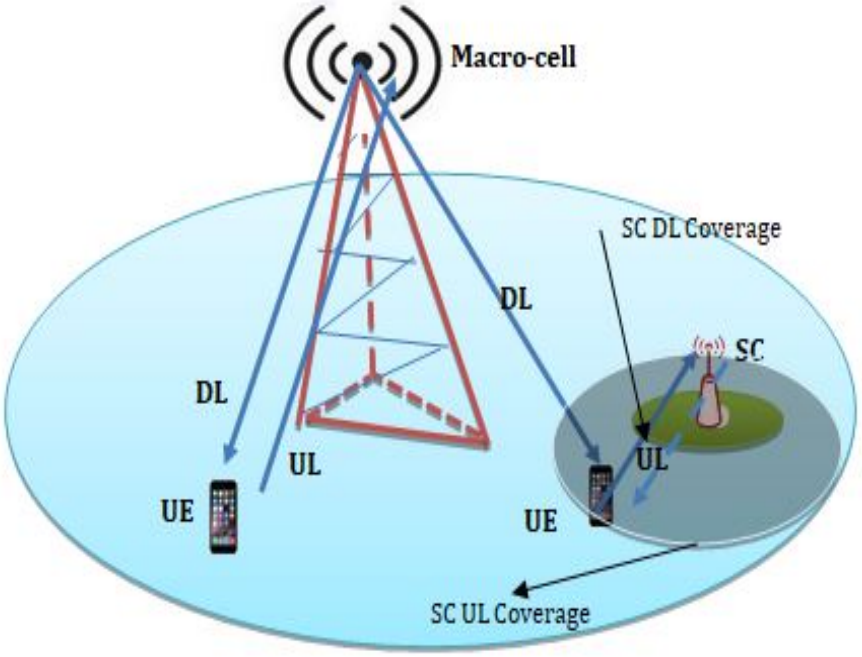


Figure 2.3: Downlink and Uplink Decoupled UA

to a number of baseline schemes, particularly when subjected to high traffic loads.

In [93], the authors proposed a UA scheme in which the UE uses a per-link max-RSS user association criteria to enable UL and DL split and achieve efficient UA in a HetNets system, where users are allowed to aggregate bandwidth with the use of Dual Association (DA), which will be discussed in the following subsection.

The work in [94] studied the dependency of DUDe performance with small cell density in two-tier network with 2x2 MIMO at each tier. The result shows that the gains of DUDe do decrease after a certain threshold density of LPNs. See Table 2.2 for literature on DUDe scheme.

2.5.4 Dual Association

One specific scenario for DA is the user association based on the decoupling of UL and DL. The DA discussed in this subsection, on the other hand, is the UA with user (U-) and control (C-) planes split.

For improved performance, an DA with a split C- and U-plane allows the user to connect to two network nodes (an MC for control signaling and an LPN for user data communication). This scheme significantly improves mobility resilience in HetNets and

Table 2.2: Literature on DUDe Scheme

UA criteria	**Control Mechanism**	**Performance metrics**	**Topology**	**References**
UL and DL max-RSS	–	average network throughput	Two-tier network	[90]
Load conditions	Distributed	throughput & service delay	TDD based LPN network	[56]
UL and DL max-RSS	Distributed	UL spectral efficiency gain	HetNets with CA	[93]
No overlap	Centralized	DL and UL throughput	Two-tier network	[95]
DL max-RSS and UL min pathloss	Distributed	SINR and throughput	Two-tier network with 2x2 MIMO at each tier	[94]

UDN environments by handling handover more efficiently [53, 96]. With increased density of LPNs in an UDN system, more frequent handovers are expected, resulting in excess delay, communication overhead, and link failure, especially in high mobility environments.

In order to avoid the frequent handover problem, Nakamura et al. in [42] proposed the macro-assisted network architecture, which includes the concept of "Phantom cells" based on a multi-layer network architecture. Using different frequency bands, the concept divides the C- and U-planes between MCs and LPNs. This architecture has been standardized by the 3GPP as a part of the Release 12 specifications [97, 98, 99] See Fig. 2.4. The LPN handles traffic for high-throughput data sessions with the U-plane, while the MC controls C-plane signaling (e.g., radio resource control (RRC)) which effectively makes the LPN non-fully functional SC; here termed as "Phantom cells".

The work in [96] presents a sub-optimal algorithm to maximize the sum-rate using dual association. It considers the two profiles standardized by the 3GPP. In the first profile (1A), a UE can have a U-plane connection only with the LPN and in the second profile (3C), a UE is allowed to associate itself to both LPN and the MC. The association decision is made at the MC in both scenarios as the C-plane connection is made always

with the MC.

Carrier Aggregation (CA) and Coordinated Multi-point (CoMP) techniques require simultaneous association to more than one BS. [93] investigated DA to aggregate bandwidth as a method of improving UL capacity and spectral efficiency of aggregated transmissions. This work also takes advantage of the benefits of the DUDe association's flexibility to improve the performance of both UL and DL communications.

Only an efficient and adaptive load balancing technique can meet the spatial and temporal load dynamics. In this regard, the work in [100] focuses on UA in DA with the 3C profile and investigates a load balancing in HetNets in terms of the QoS requirements, with the objectives of UA being to maximize the load balancing index, minimize the average load of the network, and throughput.

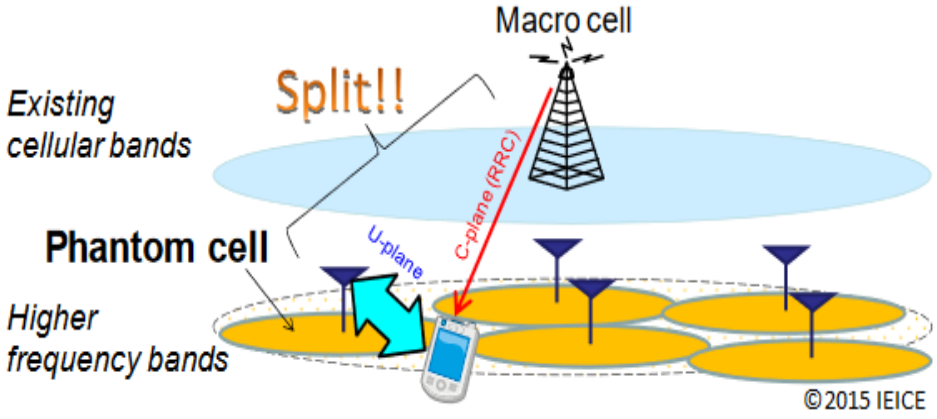


Figure 2.4: DA with Phantom cells [42]

2.5.5 Multiple Association

As discussed in the previous chapter, realizing capacity-enhancing future technologies such as CA and CoMP requires leveraging the benefits of MA techniques and algorithms.

In UDN, there is a good chance that an UE will be in close proximity to a number of LPNs. Liu et al. provided extensive simulation-based results for the probability of transmission occurring within different regions for different densification levels in [101]. The Table 2.3 shows the sharp increase in the number of transmissions for various link lengths (d in meters) and BS densities ($BS/km2$). The probability of transmission occurring within $50m$ increases up to $approx70$ times from 0.78 percent to 54.4 percent as

BS density increases from $1BS/km2$ to $100BS/km2$. This demonstrates that higher BS density indicates a higher probability of transmission points in a given region.

Table 2.3: The probability that cellular transmissions occur within different distances [101] ©2017 IEEE

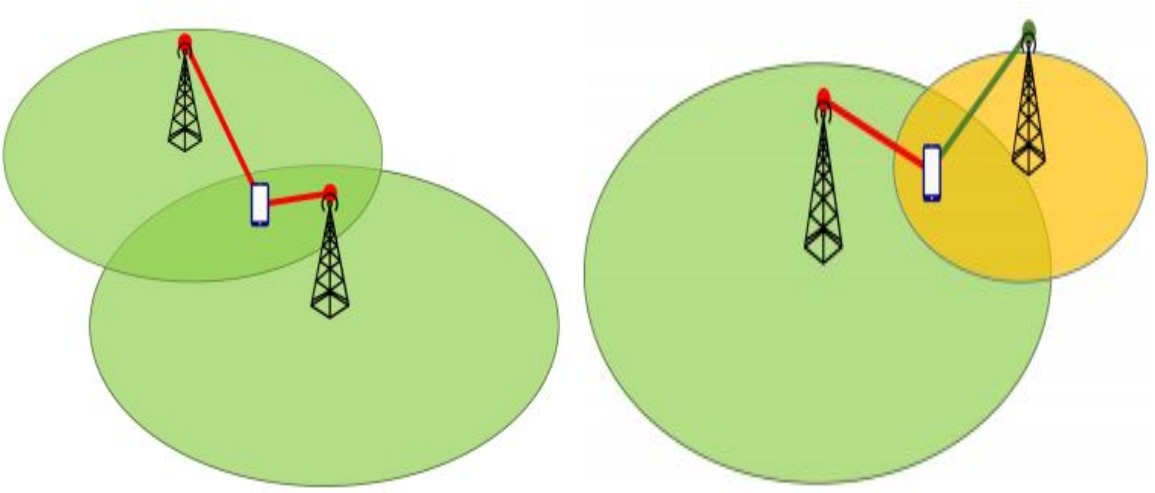
Avg. link length, d	BS **density**	**Prob.** $d <$ $50m$	**Prob.** $d <$ $20m$	**Prob.** $d <$ $10m$	**Prob.** $d <$ $5m$
500	1	0.78%	0.13%	0.03%	$3.1x10^{-6}$%
100	25	17.8%	3.1%	0.78%	0.01%
50	100	54.4%	11.8%	3.1%	0.03%
10	2500	$\approx$ 100%	95.7%	54.4%	0.8%
5	10,000	$\approx$ 100%	$\approx$ 100%	95.7%	3.1%

The work in [101] employs the PPP and assumes that the nearest neighbor association scheme connects users to the geographically closest BS. However, such a scenario is only possible in a line-of-sight dominated environment and with a short-range communication network. As multiple cells surround a given user, capacity gains in a more complex environment can only be made with appropriate and well-designed UA schemes.

As defined earlier, Multiple Association (MA) is the ability for a UE to connect to multiple access nodes at the same time. These connections can happen either within the same frequency band, which we call intra-frequency MA, or in different frequency bands, which we call inter-frequency MA [29] (see Fig. 2.5a and 2.5b).

As a result of a more complex association decision, an UE may associate with many cells and increase its aggregate throughput by communicating on multiple links at the same time. MA provides higher throughput because it allows for greater adaptability and flexibility. Furthermore, MA scheme may also improve backhaul capacity by allowing for multiple selections and easing constraints [27].

MA benefits from DUDe and C/U-plane splitting [93] and it can be seen as the extension of DA with C/U-planes are split. A UE can receive/transmit user data on multiple links based on the U-plane associations. Summary of literature on DA and MA is given



(a) Intra-frequency MA (b) Inter-frequency MA

Figure 2.5: Intra-frequency MA (a) and inter-frequency MA (b) [29]

in Table 2.4.

2.6 Prospectives and Research Challenges

This section focuses primarily on the futures and challenges of UA efforts in dense HetNets and UDN environments. In addition to existing unsolved optimization problems, UDN has posed new challenges that must be addressed immediately in order to achieve the aforementioned capacity enhancements from 5G networks. The section discusses the joint Optimization of UA and interference coordination, cell clustering challenges, mobility support mechanisms.

2.6.1 Joint Optimization of UA and interference coordination

As mentioned in Section 2.3 above, load-awareness is crucial for efficient UA in highly irregular deployment and service demand distribution scenarios. The most important goals for an effective and efficient UA are load-aware offsetting and adaptive LPS configuration.

Fair resource allocation improves the cell-edge performance but the irregular spatio-temporal service demand distribution requires dynamic load balancing algorithms. The works in [103, 104] reveals the severity of the temporal and spatial user demand variations

Table 2.4: Summary of Literature on Dual and Multiple Association Schemes

UA criteria	**Control Mechanism**	**Performance metrics**	UA Scheme	**Topology**	**References**
UL and DL max-RSS	Centralized	Sum-rate	DA	Two-tier network	[96]
UL and DL max-RSS	Distributed	UL spectral efficiency gain	DA	HetNets with CA	[93]
Load conditions	Distributed	load balancing index, spectral efficiency and throughput	DA	Two-tier network	[100]
Weighted sum-rate	Centralized	Weighted sum-rate	MA, no C/U-plane split	OFDMA based HetNets	[102]
No overlap	Centralized	DL and UL throughput	MA, C/U-plane split with DUDe	Two-tier network	[95]
Nearest cell	–	Average DL rate	MA	UDN	[27, 28]

at different social attraction sites and its effect on the user throughput. This gets worsened in ultra dense HetNets with irregular deployment topology. Transmit power difference and load imbalance needs search for paradigm shift in UA schemes.

According to the work of Abbas et al. in [87], the unplanned deployment of LPNs has a significant impact on coverage and rate. As a result, the adaptability of UA to irregularities in deployment and demand distribution must be investigated.

Techniques, schemes and any algorithm used for UA will undoubtedly have an impact system performance by changing the load distribution among BS and influencing the interference levels. In other words, as UA schemes change, so does the interference matrix. As a result, efficient interference management techniques are required for each UA scheme used.

Hamza et. al. in [105] examined various ICI coordination techniques in Orthogonal Frequency Division Multiple Access (OFDMA)-based Cellular Networks. Some of them were self-organizing strategies for eICIC [106], algorithms for eICIC [84, 107], and MC throughput enhancement using FeICIC [26, 85]. However, future wireless networks, such as heterogeneous UDNs, are expected to be multi-RAT HetNets, with high node density

per area, severe demand variation, and high energy efficiency, necessitating the use of advanced and intelligent ICI techniques. To accomplish these goals, joint optimization of adaptive UA with ICI coordination algorithms must be prioritized.

Algorithms should also be simple enough to be implemented in real-world networks. It is critical to optimize the cell-specific offset values for CRE and the adaptive number of LPS periods together. In other words, this effort should take into account the load dynamics, interference coordination, and network topology all at the same time.

2.6.2 Challenges of cell clustering

The control mechanism impacts the performance of a UA. Clustering can be used to combine the advantages of centralized and distributed control based UA. Further, cell clustering play an important role both in single and multiple association schemes.

In single association scheme, clustering can be used to facilitate resource sharing efficiency. Loads must be shared with those cells which are in a neighborhood of a given UE. Only some studies have considered both clustering and the joint optimization of UA and interference coordination. As per our knowledge, limited studies [108] and [109] have considered this important aspect for the performance enhancement of HetNets and heterogeneous ultra dense networks [110].

In MA, a cluster of LPNs form a virtual cell by multiple SCs to serve one UE. This enhanced solution is known as the soft-association control, in which each UE with multiple association capability is associated with a virtual cell [111]. This virtual cell is mapped to a set of LPNs dynamically. The effort on soft-association can be supported by the help of C-RAN as proposed in [112] using joint UA and resource allocation and the well promoted in software-defined radio. In [113] both centralized and distributed control mechanisms were presented for sum rate maximization, sum rate maximization with proportional fairness and repeated game model. However, further efforts should take into account a truly UDN environment either in the C-RAN or dense LPN topologies.

2.6.3 User mobility and UA

As presented earlier, attaining the high capacity increase and stringent latency requirement of 5G system is tempted by the UA scheme. It was shown that achievable user rate drops

dramatically as the user mobility increases [46] and increased signaling overhead [16].

When the size of a cell gets smaller in 5G networks, the traffic load balance issue emerges in contrast to macro cells that can smooth the random fluctuation in the space domain. With cell size reduced to tens of meters in 5G cellular networks, quickly moving terminals lead to frequent handovers and additional latency is inevitably added. When the handover occurs between different types of heterogeneous wireless networks, the large amount of overhead will decrease the data exchanging efficiency [114].

Therefore, associating high speed user to LPNs which has a small footprint consequently results in overhead of frequent cell re-selection. So, categorized UA based on the relative speed of UE is necessary. In this scheme, fast moving UEs chose to associate to MC while slow moving UEs associate to LPN [115]. In UDN, due to the high density of cells per area, user movement and temporal wireless channel variations could result in unstable association and hence reduce performance.

The intermittent UA and frequent transfer from one LPN to the other brings a challenge in UDN, which also happens in the case of mmWave communication [49]. Kela et al. [46] proposed a continuous UDN to provide 5G services to mobile UE which is found to outperform the widely accepted solution based on macro cells and massive MIMO systems. They designed a frame structure which carries UL pilot resources constituting the basis of mobility and user tracking. In their work, a continuous UDN is a network composed of a high number of LPNs providing continuous small cell coverage. Here too, there is profound motivation for intelligent clustering-based UA for improving the latency and reliability. Further studies on UA in C/U-plane split systems can play central usefulness.

2.6.4 Other real world considerations

Real world assumptions and policies has to get due attention in order to achieve the best from UA scheme [33]. In its broadest term these settings could include: power budget, stable connection guarantee and transmission cost, the channel and service demand models used, backhauling options utilization and algorithmic complexity. Energy efficiency optimization through managing the Power budget and transmission cost has already became consequential factors in 'green-network' notion of 5G wireless systems.

Operators usually favors cost-effective deployment and follows evolution paths for their

network. As pointed out in Chapter one, the choice between coupled or decoupled association depends on the service demand and densification levels of networks. Therefore, knowing the impact of cell density/intensity on UA policy and network performance during the evolution process help operators a safe and cost-effective roll-outs.

Easy to study models are often used in simulations and modeling complex systems like wireless links. However, in real world deployment all complex and interrelated factors comes into play and the simulated results may not be repeated. This is particularly considerable in pathloss [34], service nodes topology and service demand modeling.

Algorithmic Complexity is another crucial factor for realization of UA schemes in real networks. Almost all UA algorithms resort to either max-RSS or simple biasing when it comes to the realization in real deployment. This is because of the computational costs [23, 67] related to these algorithms which impacts the association time. Especially, all on-line running algorithms need to converge quickly to minimize the overhead.

Backhauling is the network that connects the access nodes or cells to the core-network and consists mostly of dedicated fiber, copper, microwave and mmWave links in the next 5G systems [116]. In UDN, backhauling poses an important challenge as the large number of LPNs involve with high data rate load and different environmental condition may exacerbate the problem of carrying high data rate. Network latency-aware UA schemes as in [117] can be useful in backhaul constrained UDNs. However, the reliability during failure of backhaul links also needs further work on resilient, and self-optimizing UA schemes. An implementation of CoMP also found to incur backhaul problem [118]. But, UDN along with its challenges associated to backhauling, has a potential to solve much of the future network problems.

2.7 Conclusion

The exponentially flaring capacity demand from mobile data users and stringent latency requiring applications are pressuring operators to consider advanced technologies proposed in the 5G realm. The most important of these are: the ultra-densification of network nodes, mmWave communication, massive MIMO and C-RAN based technologies. A UDN can be realized by combining and utilizing these different technologies in many different

scenarios: densification through use ultra-dense SC deployments, massive MIMO enabled MC and LPNs with use of ample bandwidth at mmWave spectrum and densification through ultra-dense deployment of RRH in C-RAN architecture.

Related to UA, meaningful attention is needed to consider: dense irregular deployment and demand distribution, ICIC, backhauling, control mechanism and mobility support. In addition to this, UA can be seen as multi-objective optimization problem which tries to improve the network capacity in terms of performance metrics including fairness, throughput, latency, energy efficiency and backhauling overheads.

UA schemes are explained following the step-by-step advancements due to the growing and varying requirements of different technological progresses and categorized according to the problems they are supposed to solve. This work also detailed the load balancing efforts together with UA in HetNets, DUDe, C/U-plane split and the need for MA. From the discussion on prospectives and research challenges, it is learned that there are ongoing research efforts which is answering the major problems projected in the coming 5G networks specifically in UDN environment. However, it also observed that complex and context specific UA schemes under each categories of dense HetNets scenarios has to be further examined including: adaptivity to irregular topology, clustering and multiple association, mobility support and other many real world problems.

Chapter 3

Load-aware Offsetting and Interference Coordination

This Chapter extends the work presented in [P1] by giving more details on the coupled association in HetNets. As the 4G LTE is already rolled out in many countries, one of the most effective approaches to enhance the spatial reuse of limited spectrum is through the deployment of small cells to complement existing MC networks. The problem that naturally arise from transmit power differences between different tiers, spatio-temporal demand distribution and the proposed solution are dealt with in this chapter. The concepts of load balancing, CRE, ABS and LPS will be addressed. Then, network and link models, the problem formulation and proposed solution will presented. Lastly, result analysis and conclusions will be presented.

3.1 Joint Offsetting and Interference Coordination

Adopting a multi-tier network topology does not effectively solve the problem of the surging data traffic demand. One important challenge is the existing User Association (UA) mechanism based on the Maximum Received Signal Strength (max-RSS), where some User Equipment (UE) may associate with LPNs and on the contrary MCs remain overloaded. Moreover, the load imbalance may naturally occur owing to the irregularity in the service demand distribution. Indoor traffic demand accounts for up to 70% of the total mobile data demand, and users are usually concentrated at social attraction places

such as residential and office buildings, shopping malls, and bus stations, among others, representing hotspot areas [61, 62].

In a dense and irregular deployment, the challenges in UA further increase because of the transmit power disparities between different tiers of the network. With max-RSS-based UA scheme, a few UEs associate to the LPNs which will further exacerbate the load imbalance on top of the effect of irregular service demand distribution.

Due to this, for multi-tier network like in HetNets with varying transmit power, the bias-based UA scheme has got greater attention [23, 25]. To offload more traffic to LPNs, the Cell Range Expansion (CRE) was developed by the Third Generation Partnership Project (3GPP) [44]. Using CRE, biasing extends the LPN's coverage and more users are attracted towards LPNs, thus resulting in a fair distribution of service provision. Therefore, in HetNets, user traffic is offloaded to small cells when possible to increase the user throughput and system capacity by improving the ASE.

The use of CRE to push more UEs towards the LPNs is found to enhance the performance especially at the cell-edge and *hotspots* areas [82, 83, 54] but results in further increase the asymmetry between the DL & UL and remain in sub-optimal performance. To reduce the inter-tier interference due to this asymmetry, time domain orthogonalization using the ABS was proposed [82]. The time-domain ICIC is considered a promising interference coordination scheme for 4G and 5G systems, for which a more detailed survey can be found in the literature [119].

In an ABS scheme, some sub-frames are left blank (except the power for control signaling) by the unbiased MC and off-loaded users are associated with the LPNs and served by these sub-frames. In 3GPP the ABS is enabled by the eICIC that provides means for macro- and pico-cells to time-share the radio resources for DL transmissions [83, 84]. Consequently, CRE with an ABS was developed to mitigate increased interference from unbiased MCs [77, 82]. However, static biasing and ABS could not solve the problem in case of a severely irregular service demand distribution. Therefore, integrating load awareness when selecting offset values in CRE is critical and can enhance system performance. Dynamically setting the ABS periods has also attracted research efforts because the SINR of cell-edge users may deteriorate because of a strong interference from an unmatched frame.

Though the use of ABS improves the system performance, it significantly degrades MC network total throughput. This problem is addressed through encouraging the ABS to be re-used in MC center region enabled by the FeICIC [26, 85]. In this later case, MC center users can reuse the ABS using low transmit power (i.e., using so-called Low Power Subframes (LPSs)). An effective combination of CRE and LPS can potentially enhance the network performance with appropriate bias setting and dynamic LPS ratio configuration.

When LPNs and MC co-exist, the DL received power based UA will be in-efficient as it produces load imbalance between the nodes. Load-aware UA are a must and ways of realizing in real deployments also needs a consideration and further research. Further, densification of the network together with high frequency reuse results in increased ICI. The eICIC, and FeICIC defined in 3GPP's LTE-Advanced standard is geared to solve this problem. However, the unplanned deployments of LPNs especially in customer premises could increase the ICI. A review of adaptive cell selection techniques in LTE-Advanced to attain improved balanced traffic and system capacity was presented by Gadam et al. [79]. It shows that adaptive biasing perform better compared to static biasing.

3.2 Contribution of the Work

In this chapter, we present the joint optimization of cell-specific offset values and number of LPS periods considering the load dynamics to maximize the user throughput. To steer traffic toward appropriate LPNs, we use a cell clustering approach where MCs are cluster centers. Our cell clustering approach is designed based on the interference between cluster centers (i.e., MCs) and cluster elements or LPNs. We integrate load awareness into the offset value selection and LPS period configuration processes using iterative and enumerative methods, respectively. We call this scheme a LA-OLPS configuration. Using this approach we can separate the UA functions to be performed by the network and users. The coordination of all the optimization procedures is performed on the network, and users can perform simple association procedures using the known max-RSS scheme [25]. The LA-OLPS configuration enables users to make distributed cell-selection decisions, ensuring less complexity in algorithm implementation. The contributions of this work are

summarized as follows:

1. We develop a link model that can be used to study the network performance in four service regions for both MCs and LPNs. Using this model, we can investigate the effects of load-aware dynamic offsetting and adaptive LPS configuration on key performance indicators like SINR and throughputs.

2. We formulate a clustering-based LA-OLPS configuration that employs the iterative offsetting and enumerated LPS period configuration. This work considers the joint optimization of clustering and load-aware offsetting with an adaptive LPS ratio; all functions are computed by the network. Using this scheme, users require only the received signal power and offset values for UA procedure initiation. Our approach enables an easy implementation in the existing system while providing an effective offloading and interference coordination solution.

3. The performance of the proposed scheme is evaluated using system level simulations with irregular network deployment and service demands. Our results reveal the performance changes or spatial gains corresponding to the system model developed in this work.

The rest of this chapter is organized as follows. Section 3.3 presents studies related to efforts on UA and ICIC in a single-tier network and the challenges in UA and interference coordination in HetNets environments. Section 3.4 discusses the system/link model and topology applied in this work. The problem formulation and proposed solutions are presented in Section 3.5. Section 3.6 presents performance evaluation settings and discusses the results. The conclusions are drawn in Section 3.7.

3.3 Related Works

The dynamic UA was proposed for load balancing and interference avoidance in multi-cell networks [73] to jointly optimize partial frequency reuse in ICIC and load-balancing schemes and improve the cell-edge performance. The system evaluation was limited to a single-tier network case, calling for an extension to the multitier HetNets scenario, which was treated in another study [76]. An irregular layout with LPNs and an irregular service

demand distribution further increase the challenge pertaining to the selection of a reuse pattern and UA scheme, which affects the load distribution.

Using CRE to steer additional UEs toward LPNs along with ICIC scheme enhances the performance, particularly in the cell-edge and hotspots areas [82, 83]; however, this leads to a further increase in the asymmetry between the downlink and uplink. This causes an increased interference from unbiased MCs. degrading the system performance. To reduce the intertier interference caused by this asymmetry, time-domain orthogonalization using ABS was studied [82]. In an ABS scheme, some subframes are left almost blank by the unbiased MC to serve offloaded users, which are associated with LPNs.

In [26], the coverage enhancements of LPS using FeICIC was studied. ABS outperformed LPS in terms of the rate coverage with an optimal range expansion bias but induced a heavier burden on the backhaul of the picocell. However, with a static range expansion bias, LPS achieved a better rate coverage than ABS. Umair et al. [85] studied the identification of interferes for a system that uses FeICIC in HetNets with a static CRE offsetting of 9 dB via system level simulations. For interference management, it is sufficient to identify the strongest interference levels based on the neighboring cells. Another study [88] proposed a load-aware network selection approach applied to the automated dynamic offsetting in HetNets. This work focused on the properties of a hierarchical Bayesian game framework, where MC dynamically selects the offset values to guide users when making intelligent network selection decisions between MC and LPNs. In a study [89], the ABS ratio optimization was iteratively attempted and the distributed UA and semidistributed MC–LPN load balancing schemes were proposed. This work did not consider the LPS enhancement using FeICIC. Our work proposes a novel approach using the joint optimization of load-aware offsetting and the adaptive LPS ratio. It also employs the partially distributed cell-selection, which is a user-centric UA strategy with a centralized coordination.

In a study [120], game theory was used to compute the cell specific offset and power transmission pattern in frequency and time domains. Consequently, average and cell-edge throughput improvements of 40% and 55%, respectively, could be achieved. This finding highlights the need for a nonstatic or adaptive offset value. Additionally, Deb and Monogioudis [84] reported the importance of the joint optimization of ABS and UA using

real network data. An economic UA was studied for multimedia content delivery, which was formulated as an optimization problem, to maximize the system utility, i.e., the total profit of the network operator [121]. This work considered an access cellular network of the cache-enabled heterogeneous cloud radio, contrary to our work.

Clustering facilitates efficient resource sharing among nodes in the network. Only some studies have considered both clustering and the joint optimization of UA and interference coordination. As per our knowledge, limited studies [108] and [109] have considered this important aspect for the performance enhancement of HetNets and heterogeneous ultra dense networks [110].

Chen et al. [108] built an interference graph for LPNs based on the RSRP of all users. Then, LPNs were clustered according to the interference graph to ensure a considerably small interference between LPNs within a certain cluster. However, despite reducing this intra tier interference between LPNs in the same cluster, their approach did not guarantee the inter tier interference coordination between MC and LPN tiers, which is an important requirement in HetNets. In this paper, we introduce an approach that effectively coordinates the MC interference for offloaded users.

A clustering approach for extremely imbalanced loads in a heterogeneous ultra dense network was also proposed [110]. This approach comprised two steps for joint clustering and scheduling. The first step involved load-aware clustering, and the second step involved graph-coloring-based scheduling. Moreover, a User-Based K-means Clustering Algorithm (UBKCA) was proposed to solve the offloading problem in HetNets and, more importantly, identify cluster centers [109]. This work used UBKCA to randomly select elements, assume them to be the center of each group, and iteratively tune the cluster centers. However, because users were to be offloaded from tier-1 to tier-2, in our work, we believed that cluster centers for efficient offloading should involve MCs and it was sufficient to consider LPNs as cluster elements.

3.4 System Model and Assumptions

This section presents the network topology and service demand distribution as well as the network and link models.

3.4.1 Network topology

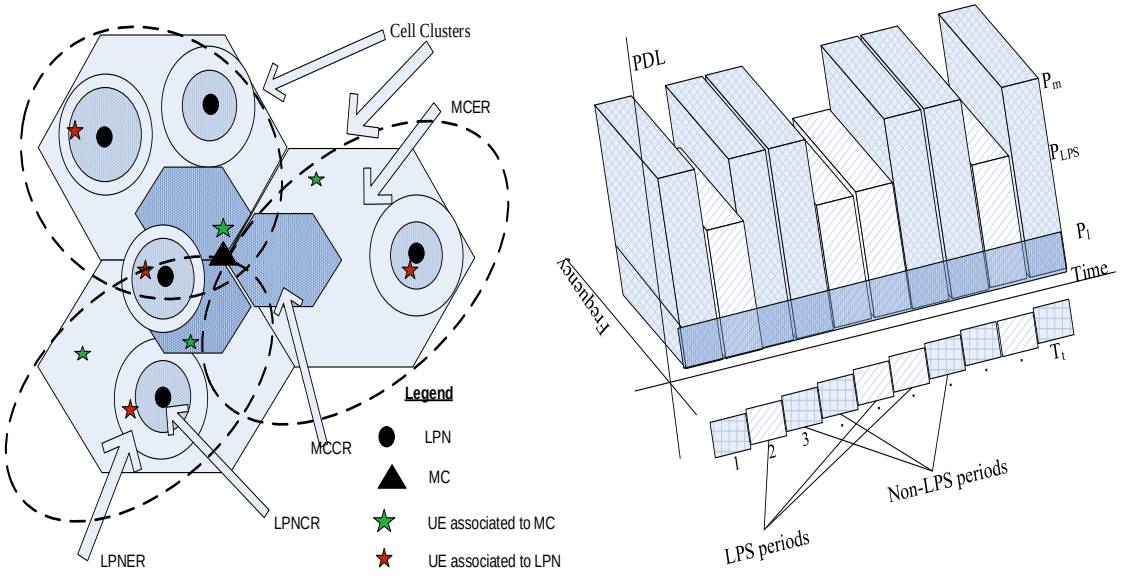
We consider three-sector macrosites overlying a large number of LPNs both in the cell-edge and cell-center regions, as shown in Fig. 3.1a for one of the macrosites. We deploy cell-edge LPNs for improving the performance of users with bad signal conditions or poor SINR, while the cell-center LPNs are used to serve the hotspots areas with high data demands. The boundary of the Low Power Node Center Region (LPNCR) is determined using the max-RSS, while that of the Low Power Node Edge Region (LPNER) is determined using the cell-specific offset value. Similarly, the boundary of the Macro-cell Center Region (MCCR) and the Macro-cell Edge Region (MCER) is determined using the channel conditions characterized by employing RSRP, SINR or other parameters.

Fig. 3.1b shows the resource orthogonalization in MCCR, MCER, LPNCR, and LPNER according to FeICIC [77, 84]. We denote the Power Density Upper Limit (PDL) [122] of MC, LPN, and LPS periods as P_m, P_l, and P_{LPS}, respectively. The resource scheduling in MCCR, MCER, LPNCR, and LPNER generally follows Round Rubin (RR). However, MCCR and LPNER users are also allocated resource blocks with an LPS period, while MCER and LPNCR users are only allocated nonLPS resource blocks. The description of the main parameters and notations is summarized in Table 3.1.

Cell clustering is used to group cells that can potentially share loads with MCs. As shown in Fig. 3.1a, LPNs are clustered under the MC or serving sector of the micro-sites. In 3GPP's LTE, the eNodeB's neighbor list is maintained at each eNodeB using the Automatic Neighbor Relation (ANR) scheme [123]. The ANR scheme enables eNodeB to detect surrounding neighbor relations based on UE measurements. Thus, the clustering approach can use this automatic neighbor list management for its implementation in a real system. More detailed discussion on clustering is presented in Section 3.5.

Another important aspects of next-generation networks is the irregular user service demand distribution, which directly affects load sharing among cells. Here, we consider only UE distribution assuming that a similar service demand exists among all users and the irregularity is reflected using the spatial and temporal variations in the geographic UE distribution. We use tools from Stochastic Geometry (SG) to define the spatial UE distribution and hence the service demand distribution.

Assume that UEs, who are supposed to represent the service demand, arrive at some



(a) Network topology (b) Resource orthogonalization using FeICIC

Figure 3.1: Network topology (a) and resource orthogonalization (b)

given region A (a subset of the plane) and independently take their locations in A at random according to some probability distribution $a(.)$. This means that each user chooses location dx with probability $a(dx)$; the uniform distribution corresponds to a *homogeneous* situation and non-uniform distributions allow us to model various *hotspots* [35].

If we denote a set of all UEs as $\mathcal{U} = \{U_i : i = 1, 2, 3, \cdots\}$ with the maximum cardinality of U, the spatial distribution can be modeled using the PPP to determine the probability that k numbers of UE exist in a two-dimensional area A:

$$\mathbf{P}(|\mathcal{U}| = k) = \frac{e^{-\lambda|A|}}{k!}(\lambda|A|)^k, \ k \in Z^+ \tag{3.1}$$

where λ is the user density per unit area. Then, the expected number of UEs in A is obtained as

$$\mathbf{E}(|\mathcal{U}|) = \sum_{k=1}^{U} k \cdot \mathbf{P}(|\mathcal{U}| = k) \tag{3.2}$$

The sets of UEs associated with MCs and LPNs are denoted by $\mathcal{U}_m \subseteq \mathcal{U}$ and $\mathcal{U}_l \subseteq \mathcal{U}$, respectively. UEs can be associated with any type of BS but not more than one simultaneously. The definitions presented in (1) and (2) and notations are used for developing

the link model and establishing a relationship between cell loads and the required offset value (Section 3.5).

Universal frequency reuse is assumed for both MC and LPN tiers. Furthermore, we assume that there is a communication mechanism such that the number of LPS needed for LPNER users can be informed to BSs and UE in the system.

Table 3.1: Notations and list of parameters in Chapter 3

Notations	Descriptions
N_m, N_l, and N	Number of MCs, number of LPNs, and total number of BSs, respectively
$\mathcal{N}$, $\mathcal{N}_m$, and $\mathcal{N}_l$	Set of all BSs, MCs, and LPNs, respectively
A	Two-dimensional area under considerations
$\mathcal{U}_m$, $\mathcal{U}_l$, and $\mathcal{U}$	Set of users associated to MC, LPN, and set of all UEs, respectively
P_m, P_l, and P_{LPS}	Transmit power of MC, LPN, and during the LPS period, respectively
α	Ratio between P_{LPS} and P_m
$P_{i,j}^{rx}(b)$	Received signal power at i^{th} UE location from j^{th} BS in dB scale
$d_{i,j}$	Separation distance between transmitter and receiver
γ	Path loss exponent
g	Exponential channel gain
$P_n(b)$	Noise power
$X_{i,j}$	UE association indicator
δ_j and δ_{max}	Cell load and maximum allowed cell load, respectively
L	Vector containing all the cell loads
λ	User density per unit area
$P_i^{rx}(b)$	Received power from the serving cell for sub-carrier b
P_j^b	j^{th} cell offset value
P_{max}^b	Maximum allowed offset value
Q_{rx}	Required minimum received power
T^{LPS}, τ, and τ_{max}	Number of LPS, LPS ratio, and maximum LPS ratio, respectively.
$\psi_{i\in\{\mathcal{U}_l,\mathcal{U}_m\}}$	Instantaneous SINR experienced by a user
r_i	User data rate
r_i^+ and r_i^-	Data rate during LPS and nonLPS periods
C_T	Aggregate throughput
η	ASE
B	System bandwidth
CC_j	List of the j^{th} cluster
I_{th}	Interference threshold for BS clustering

3.4.2 Link model

Consider two-tier HetNets comprising N_m MCs overlying N_l cochannel LPNs. The sets of MCs and LPNs are denoted by $\mathcal{N}_m$ and $\mathcal{N}_l$, respectively. We define $N = N_m + N_l$ and denote the set of BSs as $\mathcal{N} = \{BS_j : j = 1, 2, 3, \cdots, N\}$. The system bandwidth B is

partitioned into a set of subbands, each with size b, and the total time frame T is divided into t subframes. Let the number of LPS be an integral multiple of t, denoted by T^{LPS}. The noise power for a subcarrier b is denoted as $P_n(b)$. Furthermore, $X_{i,j} \in \{0, 1\}$ is a boolean variable that indicates whether the i^{th} UE is associated with a j^{th} BS. δ_{max} is the maximum cell load, and δ_j is the cell load of the j^{th} BS.

We apply the conventional propagation loss model with a path loss exponent, i.e., $\gamma > 2$, and describe the multipath fading effect using a normalized exponentially distributed gain factor g. Assuming the transmit power to be $P_j \in \{P_l, P_m, P_{LPS}\}$, the received signal power (in dB) for the subcarrier b using the i^{th} UE at a distance $d_{i,j}$ is

$$P^{rx}_{i,j}(b) = 10 \log[P_j g d^{-\gamma}_{i,j}] \tag{3.3}$$

UA is based on the received signal strength, which is a wideband received power, including noise and interference. However, because an accurate interference estimation is not possible before UA, we consider UA based on narrowband received power and treat interference coordination later in Section 3.5.

In LTE networks, the RSRP of a particular cell is the average of the received power measured for the resource elements containing cell-specific reference signals. In other words, it is a linear average of the received power from preallocated resource elements for a subcarrier [77, 84]. Therefore, without interference and noise, the received signal strength scale up or down based on RSRP; hence, we can directly use the RSRP-based association scheme. With this assumption, the max-RSS-based UA selects the candidate serving cell with the largest long-term-averaged received power $P^{rx}_i(b)$ for the i^{th} UE as

$$P^{rx}_i(b) = \arg\max_{j \in \mathcal{N}} \{P^{rx}_{i,j}(b)\} \tag{3.4}$$

CRE affects the UA decision by adding a virtual power (in dB) to the received power from LPNs in HetNets environments. In other words, in CRE, a certain positive bias value or offset (P^b_j) is added to the received power from LPNs [77, 78]. Therefore, the load imbalance situation can be mitigated by employing CRE [109]. The biasing values for LPNs range from 0 to 24 dB [120]. P^*_i represents the virtual maximum received power together with the CRE bias. Hence, a CRE based UA can be obtained by updating (3.4):

$$P^*_i(b) = \arg\max_{j \in \mathcal{N}} \{P^{rx}_{i,j}(b) + P^b_j\} \tag{3.5}$$

From (3.5), we obtain the corresponding actual received power of the candidate serving cell as $P_i^{rx}(b) = P_i^*(b) - P_j^b$. Then, using the serving cell received power (3.4) and (3.5) and the required minimum received power Q_{rx}, we obtain the UE cell association decision as

$$X_{i,j} = \begin{cases} 1, & \text{if } P_i^*(b) \geq Q_{rx} \\ 0, & otherwise \end{cases} \tag{3.6}$$

Generally, the UA topology is represented by the matrix X of size $U \times N$, with element values obtained using (3.6).

In the forthcoming discussion, $\alpha = \frac{P_{LPS}}{P_m}$ represents the ratio of the transmit power between LPS and nonLPS periods. Subcarriers with LPS are scheduled only for MCCR and LPNER users, resulting in the reduced received signal strength and ICI during LPS periods. Therefore, we obtain four SINR expressions corresponding to MCCR users during the LPS period (3.7a), MCCR and MCER users during the nonLPS period (3.7b), LPNER users during the LPS period (3.7c), and LPNER and LPNCR users during the nonLPS period (3.7d). For the i^{th} UE associated with the LPN or MC, the SINR denoted by $\psi_{i\in\{\mathcal{U}_l,\mathcal{U}_m\}}$:

$$\psi_{i\in\{\mathcal{U}_m\}}^{MCCR,LPS} = \frac{\alpha P_i^{rx}(b)}{\sum_{j\in\mathcal{N}_m} \alpha P_{i,j}^{rx}(b) + \sum_{j\in\mathcal{N}_l} P_{i,j}^{rx}(b) - \alpha P_i^{rx}(b) + P_n(b)} \tag{3.7a}$$

$$\psi_{i\in\{\mathcal{U}_m\}}^{MCER} = \frac{P_i^{rx}(b)}{\sum_{j\in\mathcal{N}_m} P_{i,j}^{rx}(b) + \sum_{j\in\mathcal{N}_l} P_{i,j}^{rx}(b) - P_i^{rx}(b) + P_n(b)} \tag{3.7b}$$

$$\psi_{i\in\{\mathcal{U}_l\}}^{LPNER,LPS} = \frac{P_i^{rx}(b)}{\sum_{j\in\mathcal{N}_m} \alpha P_{i,j}^{rx}(b) + \sum_{j\in\mathcal{N}_l} P_{i,j}^{rx}(b) - P_i^{rx}(b) + P_n(b)} \tag{3.7c}$$

$$\psi_{i\in\{\mathcal{U}_l\}}^{LPNCR} = \frac{P_i^{rx}(b)}{\sum_{j\in\mathcal{N}_m} P_{i,j}^{rx}(b) + \sum_{j\in\mathcal{N}_l} P_{i,j}^{rx}(b) - P_i^{rx}(b) + P_n(b)} \tag{3.7d}$$

The channel link rate $r_{i\in\{\mathcal{U}_l,\mathcal{U}_m\}}$ for a given UE associated with the MC or LPN is estimated using the modified Shannon formula:

$$r_{i\in\{\mathcal{U}_l,\mathcal{U}_m\}} = b \cdot A \log_2(1 + B\psi_{i\in\{\mathcal{U}_l,\mathcal{U}_m\}}) \tag{3.8}$$

where A and B are the link bandwidth and SINR efficiency, respectively, which are technology-dependent link factors. For LTE, the A and B values of 0.83 and 1.10, respectively, are used as in the literature [124].

The time orthogonalization used in FeICIC also affects the average data rate. Let the link rates obtained using (3.8) during LPS and nonLPS periods be r_i^+ and r_i^-, respectively.

Further, let the ratio of the LPS period to the total frame length be $\tau = \frac{T^{LPS}}{T}$. Then, the average data rate achieved by the i^{th} UE can be approximated using a weighted average as

$$r_i = \tau r_i^+ + (1 - \tau) r_i^- \tag{3.9}$$

For users scheduled for subcarriers without LPS periods, $\tau = 0$ and the data rate is the same as that in (3.8).

3.5 Problem Formulation and Proposed Solution

In this Section, we develop mathematical expressions for a resource-sharing network while attempting to optimize the aggregate system throughput. First, we formulate the joint optimization problem for UA and interference coordination. Then, a solution is proposed based on cell clustering. The iterative load-aware offsetting along with the adaptive LPS period configuration is developed. Lastly, we present the algorithm for load-aware offsetting and LPS configuration.

3.5.1 Problem formulation

Using the definitions discussed in the previous section, we can obtain the aggregate system throughput as

$$C_T = \sum_{j \in \mathcal{N}} \sum_{i \in \mathcal{U}} r_i \cdot X_{i,j} = \sum_{j \in \mathcal{N}} \sum_{i \in \mathcal{U}} [\tau r_i^+ + (1 - \tau) r_i^-] \cdot X_{i,j} \tag{3.10}$$

We measure the cell load in terms of the bandwidth usage or the ratio of the utilized resource to the total resource of the cell. In LTE, a smallest physical resource block that can be allocated to a user shows a 180-kHz bandwidth. Therefore, for the total available bandwidth B (in MHz) at a cell and n resource blocks allocated per user, the j^{th} cell load δ_j can be expressed as

$$\delta_j = 0.18/B \cdot \sum_{i \in \mathcal{U}} n \cdot X_{i,j} \tag{3.11}$$

From (3.11), we can observe that the cell load depends on the association variable $X_{i,j}$, which depends on the received power and offset P_j^b, as presented in Sub-section 3.4.2. Therefore, we obtain two independent variables, namely, the offset value (P_j^b) and the

LPS ratio (τ), in our optimization problem (3.12). Our objective is to maximize the aggregate throughput C_T in (3.10) as

$$\max_{\{P_j^b,\tau\}} \sum_{j\in\mathcal{N}} \sum_{i\in\mathcal{U}} r_i \cdot X_{i,j} \tag{3.12}$$

with the following constraints:

(a) $\delta_j \leq \delta_{max}; \forall j \in \mathcal{N}$

(b) $\sum_{j=1}^{N} X_{i,j} = 1; \forall i \in \mathcal{U}$

(c) $0 \leq P_j^b \leq P_{max}^b$ and

(d) $0 \leq \tau \leq \tau_{max}$

A brief explanation of the constraints is presented. Constraint (a) sets the maximum limit on the cell load determined using the available resource and control overheads. Constraint (b) shows that a user can only be associated with one BS, where $X_{i,j} \in \{0, 1\}$. Constraint (c) sets the lower and upper bounds for the offset values (in dB). The association variable in constraint (b) is mainly affected by the P_j^b value. Constraint (d) sets the lower and upper bounds for the LPS ratio. A typical value of $\tau_{max} = 0.5$ is used in [108].

Here, (3.12) is a Nondeterministic Polynomial time (NP)-hard problem because of the boolean variable $X_{i,j}$. However, the potential gains from load-aware associations are large and the load sharing proportional to the resource at each serving node is optimal [24]. Therefore, most UA and resource allocation problems are formulated as utility optimization. Typically, logarithmic utility functions are used to guarantee a fair load and resource distribution. In solving such problems, MCCR users experience little to no rate gain. However, in practice, these users are already served well and the approach still ensures a near-optimal performance.

In our approach to solve (3.12), we consider the fair load distribution by designing load-aware offsetting and adaptive LPS configuration as well as a low complexity algorithm. To achieve this, we separate activities at the network server or centralized controller and UE. The network server performs offset value computations and sets the LPS configuration to ensure load awareness and effective interference coordination, respectively. Signal measurements are performed and reported by the UE. Furthermore, UE sends association

requests similar to the max-RSS association, where easy implementation is guaranteed in existing networks. For activities at the network server, we first develop a dynamic and load-aware offsetting for proportional load sharing among clustered cells. Then, we design an adaptive interference coordination using FeICIC at the LPNER and MCCR.

3.5.2 Load-aware offset

In this subsection, we develop a relation between the demand by UE in the LPN coverage area and the offset required for effective load balancing between MCs and LPNs. Our approach uses cell clustering for designing an effective load transfer mechanism between MCs and their clustered LPNs. This will ensure load sharing control within the neighboring LPNs or cluster elements.

Let $\mathbf{E}(d_c)$ and $\mathbf{E}(d_e)$ be the expected number of UE in LPNCR and LPNER, respectively, where d_c and d_e are the corresponding radii of the two regions. Here, the inter-UE distance is assumed to be exponentially distributed (i.e., the spatial UE locations follow the PPP with density λ, as discussed in Section 3.4). The two expressions can be expressed as follows:

$$\mathbf{E}(d_c) = \sum_{k=1}^{U} k \cdot \frac{e^{-\lambda\pi d_c^2}}{k!}(\lambda\pi d_c^2)^k \tag{3.13}$$

and

$$\mathbf{E}(d_e) = \sum_{k=1}^{U} k \cdot \frac{e^{-\lambda\pi d_e^2}}{k!}(\lambda\pi d_e^2)^k \tag{3.14}$$

The difference between $\mathbf{E}(d_e)$ and $\mathbf{E}(d_c)$ corresponds to the estimated number of UE that needs to be offloaded toward the LPN. Determining the d_e and d_c values yields the offset value required to offload $\mathbf{E}(d_e) - \mathbf{E}(d_c)$ number of UE. We obtain the required offset by equating the long-term-averaged received power at the LPNCR edge with that of LPNER with the required bias:

$$\begin{aligned} 10\log[\mathbf{E}_g\{P_j g d_c^{-\gamma}\}] = 10\log[\mathbf{E}_g\{P_j g d_e^{-\gamma}\}] \\ + P_j^b|_{j\in\{1,2,\cdots,N_l\}} \end{aligned} \tag{3.15}$$

where $\mathbf{E}_g\{\cdot\}$ denotes the expectation of the received power with respect to the variable g and $P_j^b|_{j\in\{1,2,\cdots,N_l\}} \geq 0$ dB. Expression (3.15) is a direct result of the biased max-RSS UA (3.6). Assuming that the channel gain g follows an exponential distribution with mean

one, we can solve (3.15) for the required offset value as

$$P_j^b|_{j\in\{1,2,\cdots,N_l\}} = 10\log\frac{d_c^{-\gamma}}{d_e^{-\gamma}} \tag{3.16}$$

The offset value in (3.16) has a physical implication that an LPN needs P_j^b offset value or bias to extend its coverage from d_c to d_e. Based on this observation, the expected number of offloaded UE is obtained from the mean number of rejected service requests using the MC of the same cluster. This translates to the load at each MC providing the cell radius required for the LPN ((3.13) and (3.14)), which can be used to obtain the required offset value (3.16). However, as the service demand distribution is irregular and variations occur in both temporal and spatial domains, the offset value cannot be computed in a single step. The clustering-based traffic steering presented in the next subsection is used to compute the required offset value in each cluster using the iterative method.

3.5.3 Clustering and load transfer

As discussed in Section 3.4, for offloading, we assume a cluster comprising at least one MC that serves as a cluster center. In our study, clustering depends on the interference strength from MCs steered toward LPNs. This approach assumes that there is a mechanism for measuring the accurate received signal levels using the LPNs and communication scheme for resource sharing coordination. Thus, this approach ensures that our system adapts to changes in the locations of LPNs and possible ON/OFF operations in dense HetNets. Our clustering approach can be explained using the following steps.

- Let the neighbor list of the j^{th} cell cluster be contained in CC_j with cluster center MC_j.
- Each LPN ($LPN_k, \forall k \in \{1, 2, \cdots, N_l\}$) measures the interference in terms of the power received from the neighboring MCs at its own location $P_{k,j}^{rx}$.
- For an interference threshold I_{th}, select all MCs with an interference greater than the threshold by a certain amount β. In other words, $LPN_k \in CC_j$, $\forall j$, where $P_{k,j}^{rx} - \beta \geq I_{th}$.

The matrix representation of clustering facilitates the updating process of load-aware offsetting. For a matrix M with size $N_m \times N_l$, the rows indicate MCs and the columns indicate

the cluster elements or LPNs. If $LPN_k \in \{CC_j\}$, then the corresponding $M(j,k) = 1$; otherwise, $M(j,k) = 0$.

The cluster elements and centers are populated every time the algorithm runs. However, because there is no change in the intertier interference for a short period, the frequency of running the clustering procedure can be reduced to improve the time complexity, which has not been considered here. Once the clustering is completed, the load-aware offsetting can be performed using iterative updates. We denote a vector containing the cell loads of MCs as $L_m = [\delta_1, \delta_2, \cdots, \delta_{N_m}]$. The average load among BSs is obtained as $L_{av} = \frac{\sum_{j=1}^{N} \delta_j}{N}$. Then, the weight for the offset value update is calculated as:

$$\omega = L_m M \tag{3.17}$$

The iterative updates on the bias values transfer loads from the highly loaded MC to its neighbor LPNs in the same cluster. Accordingly, the offset value update with a step size h is computed as

$$\Delta P^b_{j,j\in\mathcal{N}_l} = h \cdot \omega \cdot \frac{L_{av} - \delta_j}{L_{av}} \tag{3.18}$$

The Jain's fairness index [53] is used to evaluate the fairness of load distribution in the system, which is defined in terms of the cell loads as

$$J = \frac{(\sum_{j=1}^{N} \delta_j)^2}{N \sum_{j=1}^{N} \delta_j^2} \tag{3.19}$$

We identify two fairness levels: J_l, the one that can be achieved using a given UA scheme, and J_r, the required fairness level. We use these parameters to control the algorithm convergence in Subsection 3.5.5.

3.5.4 Adaptive LPS

For effective FeICIC and resource orthogonalization, the LPS ratio needs to adapt to the service demand in LPNER. In this subsection, we develop the enumeration-based LPS ratio update.

The maximum available resource at the LPN is proportionally allocated to LPNCR and LPNER users. Based on the expression of the expected number of UE in (3.13) and (3.14), the maximum ratio of LPS to the total frame is $\tau_{max} = \frac{\mathbf{E}(d_e) - \mathbf{E}(d_c)}{\mathbf{E}(d_e)}$. The demand served in MCCR and LPNER must potentially be fairly equal to facilitate FeICIC and

resource allocation. Based on this observation, τ can be expressed as $\tau = \frac{\mathbf{E}(d_e)-\mathbf{E}(d_c)}{\mathbf{E}(|\mathcal{U}_m|)}$, where the numerator represents the load rise in the respective LPN and $\mathbf{E}(|\mathcal{U}_m|)$ is the expected number of UE associated with MC_j. Therefore, for different load conditions, τ_j^C in the cluster CC_j is updated using the load in LPNER as follows:

$$T * \tau_j^C = \lfloor (\frac{\Delta\delta_j^C}{\sum_{j=1}^{N_m} \delta_j})T \rfloor \tag{3.20}$$

where $\Delta\delta_j^C$ is the average load rise in LPN owing to the CRE in cluster $C \in \{CC_j, j = 1, 2, \cdots, N_m\}$ and δ_j is the cell load of MC_j. Here, T^{LPS} assumes a smallest integer bound because the number of subframes can only be an integer value.

Thus, when the offset value is set to zero, the initial estimate of τ will be zero and using the largest CRE offset, this value assumes τ_{max}. Therefore, in our proposed algorithm, the enumerated values start form zero and increase/decrease with a step value of τ^C, depending on the load at each cell. Using a fixed step size for τ can minimize the time complexity; however, the loss of dependency on the cell load can affect the convergence. Based on this discussion, we enumerate the value of τ for each cluster as

$$\tau = [0 : \tau_j^C : \tau_{max}] \tag{3.21}$$

3.5.5 Algorithm

The proposed algorithm is presented in Algorithm 1, which lists general routines but not a specific implementation. A typical implementation can be when clustering, offsetting, and LPS computations are separated and performed by the network coordinator and cluster centers. The general routines involve initialization, clustering subroutine, cell load computing, offsetting, and LPS updating steps. As presented earlier, the clustering subroutine involves measuring signals, building intertier interference relations, and populating cluster elements. The cell load computation depends on UE measurement reports and the max-RSS-based UA scheme. Using the iterative load-aware offsetting and the enumerated LPS ratio, the algorithm attempts to optimize the aggregate throughput.

The algorithm performance depends on a given implementation scenario. For the described typical implementation, the clustering subroutine is performed periodically or can be triggered owing to deployment changes in the LPN network. As clustering divides the problem space, the problem size is reduced to the size of each cluster (N_c). Hence, we

Algorithm 1 Procedure for load-aware offsetting and LPS configuration

1: Initializations$\{N_l, N_m, \mathcal{N}, \mathcal{U}, P_m, P_l, P_{LPS}\}$
2: $CC_j+ \leftarrow LPN_k$, if $R_{k,j} - \beta \geq I_{th}$ {cluster the LPNs using interference relations between MCs and LPNs}
3: **for all** j in $\mathcal{N}$ **do**
4: UE measures and reports RSRP values to the neighboring cells
5: Compute the load at each cell, δ_j, $\forall j$ using step 4 and max-RSS
6: **end for**
7: Compute fainess level (J_l) {use the Jain's fainess index}
8: **while** $iter < iter_{max}$ and $J_l < J_r$ **do**
9: $\Delta P^b_{j,j\in\mathcal{N}_l} \leftarrow h \cdot \omega \cdot \frac{L_{av}-\delta_j}{L_{av}}$ {update the offset value}
10: $\tau^C \leftarrow \frac{\Delta\delta^C_j}{\sum_{j=1}^{N_m}\delta_j}$ {update the LPS period}
11: Compute δ_j
12: Compute J_l
13: $iter \leftarrow iter + 1$
14: **end while**
15: Associate UE using the biased max-RSS scheme
16: **End**

achieve the time complexity of $\Theta(N_c \cdot Iter_{max})$ for offset value computation. In the LPS configuration, the selection of τ is performed using some enumerated values and the time complexity is $O(1)$. All communication overheads, including UE measurement reports, are the same based on the max-RSS UA, except for the clustering subroutine.

3.6 Performance Analysis

In this section, we first briefly present the approach using which the system level simulator parameters are configured and used. Then, the simulation results are presented and discussed.

3.6.1 Simulation settings

The MATLAB-based system-level simulator generates a signal map using the selected path loss model. It also generates a user location map and computes the received signal strength at each UE location. The main parameters and deployment scenarios are listed in Table 3.2.

We consider two deployment scenarios in an area of 4.0 km^2. In the first scenario, the deployment assumes 4 three-sectored macrosites and 30 LPNs. In the second scenario, we assume 5 three-sectored macrosites and 60 LPNs. In the simulation, an urban deployment is considered where the LPN locations are selected to be both at MCCR to serve hotspot areas and MCER for investigating the worst conditions. Fig. 3.2 shows the MC and LPN locations in deployment scenario 1 with a typical snapshot of the UE distribution. As presented in previous section, the spatial user distribution in each grid element follows PPP. A random number of hotspot grid elements are selected and user density in the hospot areas is considered to be larger than that in the non-hotspot areas.

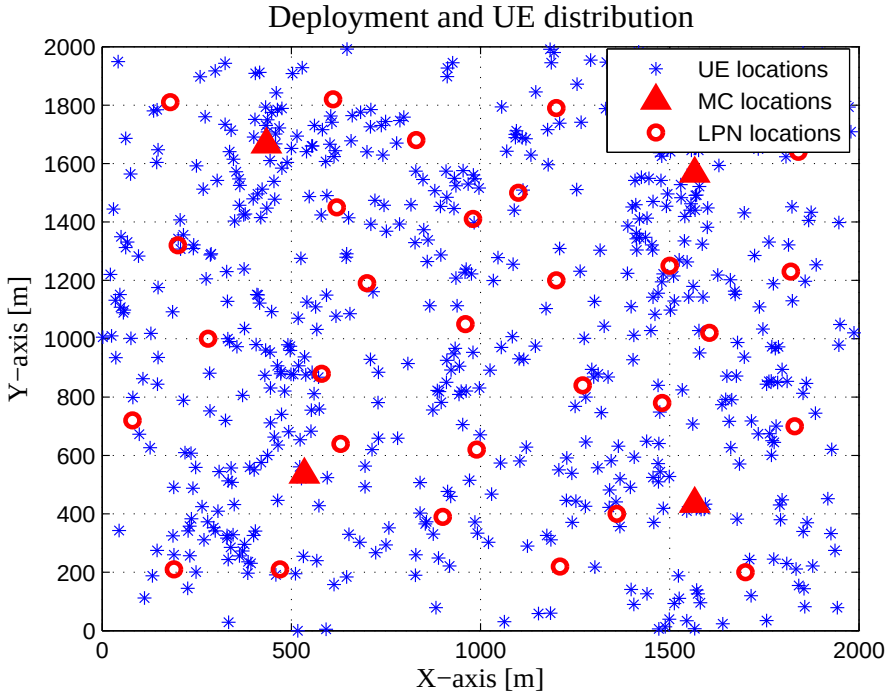


Figure 3.2: Deployment scenario 1 and a snapshot of UE distribution

Table 3.2: Simulation parameters and their values in Chapter 3

Parameters/descriptions	**Values**
Downlink transmit power	MCER: 43 dBm, MCCR: 30 dBm, and LPN: 27 dBm
Center frequency	2600 MHz
Bandwidth	20 MHz
eNodeB antenna height	MC:30 m and LPN: 5 m
UE antenna height	1.5 m
Simulation area	2000×2000 m^2
Number of UEs	Scenario 1: 588 (70% average load with $n = 5$) Scenario 2: 1050 (70% average load with $n = 5$)
Number of snapshots	300
Spatial UE distribution models	Uniform distribution with hotspot areas Typical value: $\lambda^{'}$ =1.4 λ
Tx antenna	MC: Cosine-squared power pattern, gain = 18 dBi LPN: Omni-directional, gain = 5 dBi
Rx antenna gain	0 dBi
Path loss model	From MC to UE: $128.1 + 37.6 \times Log10(R)$ From LPN to UE: $140.7 + 36.7 \times Log10(R)$ [125]
Noise power	-173 dBm/Hz
Deployment scenario	Scenario 1: 4 macrosites, 30 LPNs Scenario 2: 5 macrosites, 60 LPNs

3.6.2 Results and Discussions

Cell loads

Four schemes, namely, max-RSS, static CRE with offset values of 9 and 15 dB, and our proposed LA-OLPS, are evaluated with respect to their efficiency in load balancing. We consider both the deployment scenarios discussed in the previous subsection for comparing these UA schemes. As discussed earlier, the cell load is affected by the user distribution and transmit power of the serving nodes. LPNs attract an insignificant number of UE compared to MCs when the max-RSS-based UA scheme is used. Using the static offset values of 9 and 15 dB, a proportionally fair load distribution cannot be achieved among

the nodes. Figs. 3.3 and 3.4 show that the load variation among LPNs is large when CRE is adopted with offset values of 9 and 15 dB. If the overall system load increases, the high variation in the cell load attributed to CRE may lead to a situation where some LPNs run out of resources and LPN-neighboring UE may face outage. The clustering-

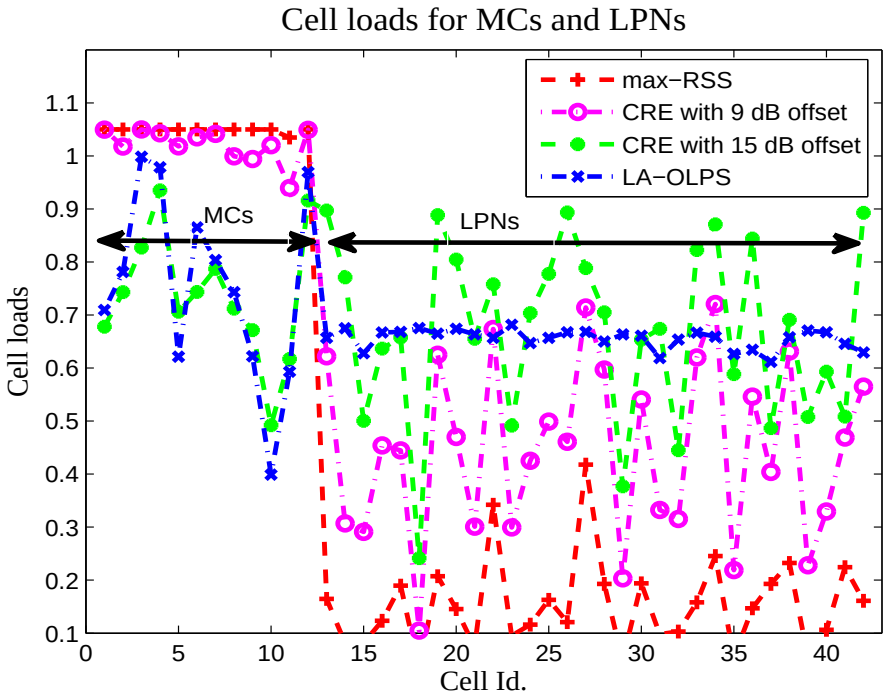


Figure 3.3: Cell loads for different schemes: scenario 1

based LA-OLPS can effectively offload traffic from the MC tier toward LPN and ensure a considerably fair load distribution. The evaluation of the two deployment scenarios show that the power disparity impose more challenge on ensuring a fair load distribution compared to the LPN density.

A comparison between the four schemes, namely, max-RSS, CRE with offset values of 9 and 15 dB, and our proposed LA-OLPS, presented in Table 3.3. The parameters of interest are the average cell load L_{av}, average cell load of the LPN tier L_{av}^{LPN}, and Jain's fairness index J_l. A fair load distribution facilitates fair resource utilization in the system. Without such load balancing efforts, average served demand is considerably less than actual demand because of high outage conditions. Using the LA-OLPS method, we observe better load sharing compared with other methods, which is also supported by a large Jain's fairness index.

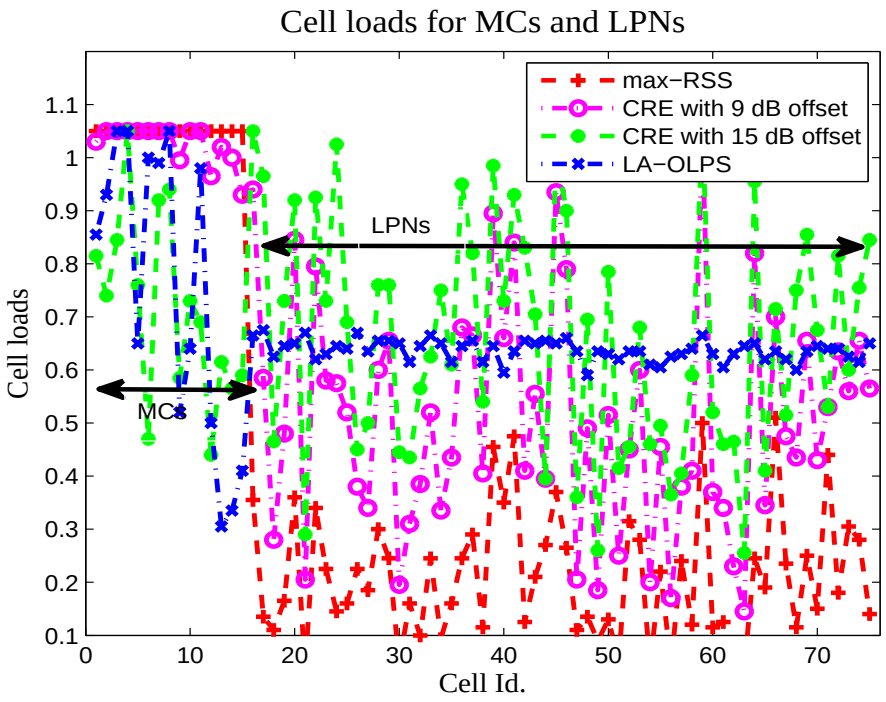


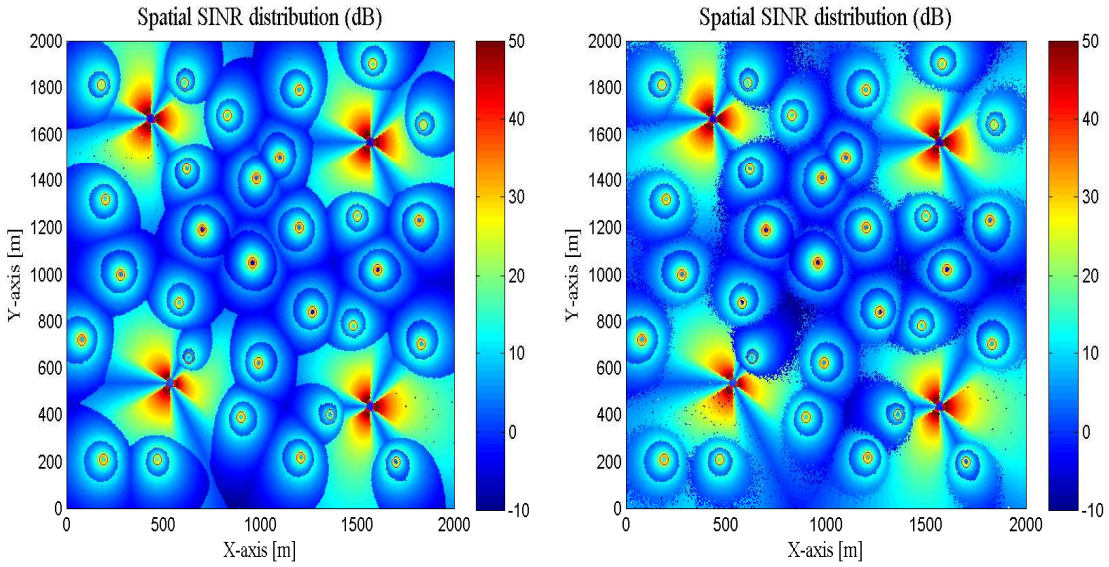
Figure 3.4: Cell loads for different schemes: scenario 2

Table 3.3: Achieved Jain's fairness index and average cell loads for different schemes

Scenarios	Parameters /schemes	max-RSS	CRE with 9 dB offset value	CRE with 15 dB offset value	LA-OPLS
Scenario 1	L_{av}/%	41.3	62.4	70.2	70.1
	L_{av}^{LPN}/%	15.9	46.5	67.9	67.8
	J_l	0.51	0.83	0.95	0.98
Scenario 2	L_{av}/%	33.9	55.2	66.9	65.2
	L_{av}^{LPN}/%	16.1	42.9	61.9	61.1
	J_l	0.46	0.77	0.90	0.96

SINR distribution

As shown in the preceding subsection, an LPN may attract less or more numbers of users depending on its offset value, which can induce varying degrees of outage events. This condition directly affects the SINR at a given location despite the availability of resources. As FeICIC or ABS/LPS is designed to reduce interference from MCs, it can be used to enhance the SINR of LPNER users. To visualize and analyze such conditions, we use the plots of the SINR spatial distribution.



(a) LPS with 15 dB biased UA

(b) LA-OLPS method

Figure 3.5: Spatial SINR distribution: static offsetting CRE with nonadaptive LPS (a) and LA-OLPS (b)

The spatial SINR distributions obtained using the static offsetting with a nonadaptive LPS configuration (Fig. 3.5a) and proposed LA-OLPS (Fig. 3.5b) method are shown. The figures are typically depicted during the LPS period because the SINR distribution is the same for both methods during the nonLPS period. In these spatial plots, improvements are observed around the LPNER of LPNs using the proposed scheme. Owing to load balancing and subframe matching efforts in LA-OLPS, the cell-edge SINR is smoother relative to the sharp decrease in the static offsetting. By closely observing the cell boundaries, we found more users with a better SINR in Fig. 3.5b compared with Fig. 3.5a. Because these observations may not be easily visible, we measured the SINR quantitatively in Fig. 3.6. The measurement result shows that LA-OLPS improves the cell-average (or 50^{th} percentile) users' SINR from 4.206 to 5.052 dB (or 20% increase) compared with the static offsetting and nonadaptive LPS configuration. The SINR measurement results also show improvements when LA-OLPS is used instead of the static offsetting and nonadaptive LPS configuration for the cell-center and cell-edge users. In conclusion, the LA-OLPS scheme enhances the LPNER user's SINR by adapting and matching the LPS to the load dynamics without affecting the gains of MCCR users.

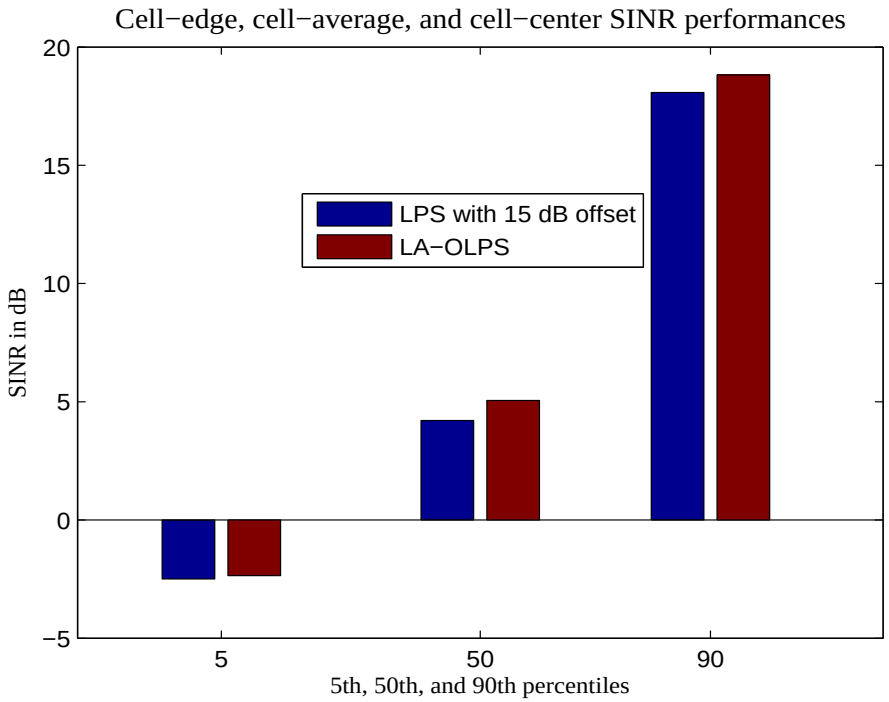


Figure 3.6: SINR performances of cell-edge, cell-average and cell-center users

User throughput and spatial distribution

For consistency in comparison, the average user and spatial throughputs are computed for a single resource block. One comparison is based on the Commulative Distribution Function (CDF), and another uses spatial throughput distribution plots. In our CDF evaluation and analysis, we consider the aforementioned two deployment scenarios and four schemes: CRE with an offset value of 9 dB, CRE with an offset value of 15 dB, LPS with an offset value of 15 dB, and LA-OLPS.

Figs. 3.7 and 3.8 show that although the CRE with an offset value of 15 dB enhances the cell-edge throughput, it degrades the cell-center and cell-average throughputs owing to the large interference from the MC tier. A comparison between CRE with an offset value of 9 dB and that with an offset value of 15 dB confirms this finding. Using the LPS configuration with offsetting improves both the cell-center and cell-edge user throughputs. The use of high offset values to offload more UE must match with that of the adaptive LPS configuration. In this regard, the LA-OLPS approach improves both cell-center and cell-edge users throughputs. Owing to the increased LPN density and subsequently the increased interference, a reduced average user throughput is achieved. The cell-average user throughputs of 180 and 160 Kbps are obtained in scenarios 1 and 2, respectively.

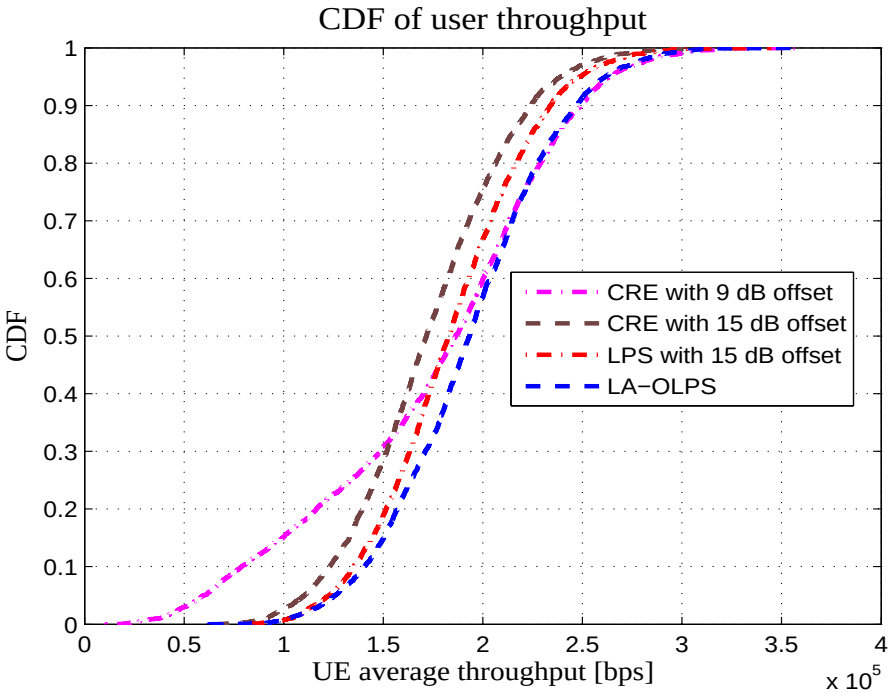


Figure 3.7: CDF of average user throughput: scenario 1

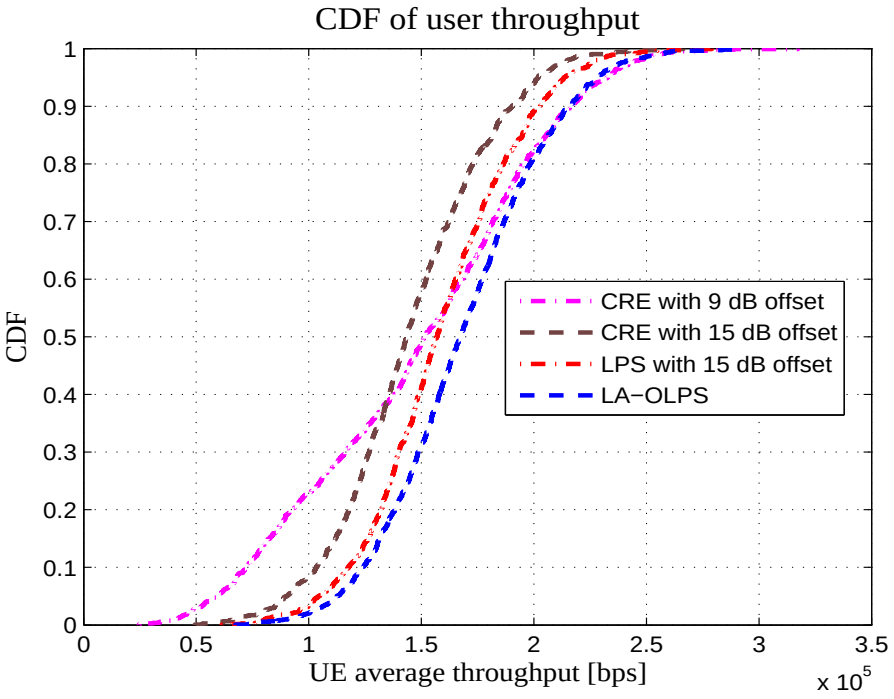
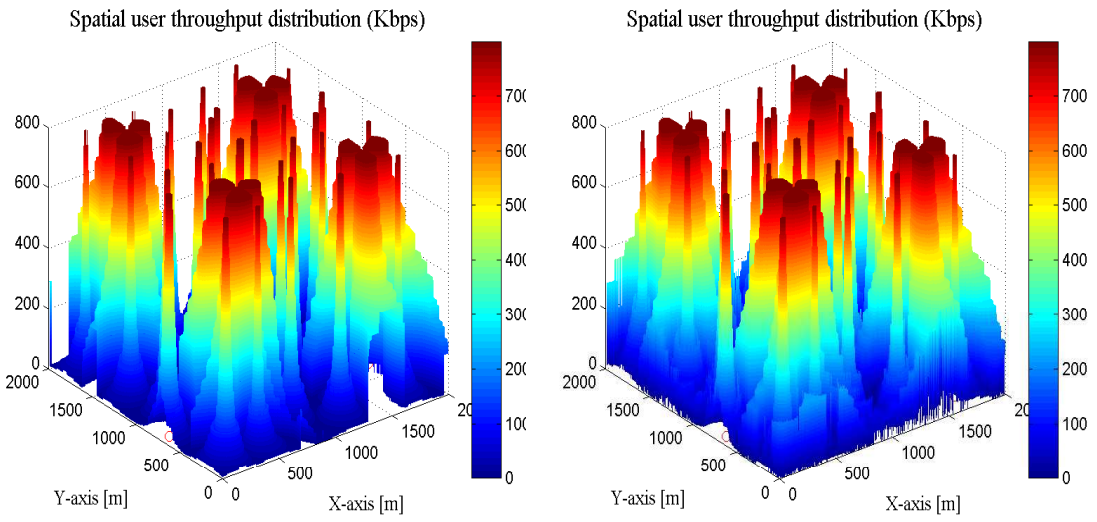


Figure 3.8: CDF of average user throughput: scenario 2

In Figs. 3.9a and 3.9b, we compare the spatial user throughput distribution for two schemes: LPS with an offset value of 15 dB and clustering-based LA-OLPS. Using LA-OLPS, enhancements can be observed in terms of LPNER (Fig. 3.7). Using LPS



(a) Spatial user throughput distribution using LPS with an offset value of 15 dB

(b) Spatial user throughput distribution using LA-OLPS

Figure 3.9: Spatial user throughput distribution with static offsetting (a) and LA-OLPS (b)

with an offset value of 15 dB, LPNER users achieve a lesser throughput and sharp decrease in throughput are clearly observed. Conversely, using LA-OLPS, LPNER users obtain a higher throughput because both the LPS configuration and offsetting adapt to the load dynamics. Thus, the enhancement in the LPNER users' throughput using LA-OLPS is achieved without reducing the center-cell users' throughput.

3.7 Conclusions

We introduce clustering-based LA-OLPS in dense HetNets. An iterative load-aware offsetting and an enumeration-based LPS configuration are formulated and evaluated using system-level simulations. The proposed scheme balances load among cells and converges after some iterations. Using clustering-based LA-OLPS, the LPN cell loads are more uniformly distributed than those in the case of the static offsetting and LPS configuration. The average user throughput is enhanced, particularly for cell-edge and cell-average users, without reducing the cell-center users' performance compared with LPS with static offsetting. The users require only the received signal strength and cell-specific offset values to initiate UA. Moreover, all optimization computations are performed on the network.

This enables the easy implementation of the clustering-based LA-OLPS method in existing systems with neighbor-list management schemes. Our system modeling approach can also be used in further studies that aim at reducing the gap between cell-edge and cell-center users' performances and achieving a fair distribution of service provision. In addition, the results demonstrate performance changes in the four service regions.

Chapter 4

Performance Analysis of Coupled and Decoupled User Association

For the full-load case, where at least one UE is associated with a cell, UL and DL decoupling can be used in order to balance the asymmetry between the two links. The question here is that at what cell density we consider decoupled association. In this chapter, The work in [P3] is extended and elaborated. The concepts of UL and DL coupled and decoupled association schemes is presented. Then, by using tools from stochastic geometry, we formulate association probabilities and define critical desification levels. Then, achievable ergodic rates are formulated. Finally, perform numerical and realistic simulations.

4.1 Background

In the last four decades, mobile communications have evolved from the First Generation (1G) to the 4G [126], where traditional communication networks, which mainly focus on voice services, have been gradually revolutionized into multi-functional systems that provide high speed mobile data and other services. In this network evolution, a UE is associated to the same BS both in UL and DL [91]. This results in the problem of DL-UL asymmetry in coverage and capacity provisioning in HetNets deployment with different transmit powers between different tiers and UEs. The problem becomes worse in the 5G UDN [127, 10] deployment because a UE may experience different propagation gains in UL and DL from nearby ultra-dense pico- or femto- BSs. Here, we relate an UDN to an

extremely large node density or intensity, which in turn defined as the number of BSs per unit area.

As a promising solution to the aforementioned problem, DUDe UA scheme has long been area of research. The DUDe UA allows a UE to be associated to MC in the DL and to the a LPN in the UL. The flexibility offered with DL and UL decoupling can be used to reduce interference and improve the throughput performance [128, 129].

On the other hand, aggressive offloading [88] of UEs from MC to LPN is the well established approach for load balancing in HetNets with coupled sub-optimal UA. However, as node intensity increases, the serving node becomes much closer to the UE. I.e., the offsetting required would also naturally decrease [92]. Therefore, the relative intensity at each tier can be used to parallel the effect of transmit power differences and to ease the required load balancing effort. The load imbalance in HetNets can also be addressed through flexible DUDe associations [130] that allows to load-balance the DL and UL separately.

Boccardi et al. [91] presented five reasons for why the UL and DL should be decoupled. These reasons are: different load balancing in the UL and DL, low deployment cost, enhanced UL data rate, reduced UL interference and reduced transmit power. Furthermore, Sial and Ahmed in [129] have shown that as BS intensities at lower tiers increase, more users prefer DUDe user association. However, in the decoupled scenario, they concluded that there is an upper bound on rise of user performance with respect to node intensity. This is also reported in [94] that as the LPN tier intensity continues to increase, the gains of DUDe do decrease after a certain threshold [94]. Above a certain densification level, users start associating their UL and DL with the same BS due to their availability in their closer vicinity.

Therefore, the effect of relative intensity ratios between HetNets tiers on the rate performance of coupled and decoupled associations needs further investigation. I.e., detailed analysis of ranges of densification levels where we can use coupled association to MC with offloading, decoupled association and coupled association to LPN tier is necessary. In this regard, this paper identifies the densificationn levels at which traffic loads in the UL and DL are independent of the network parameters like transmit power. We call this intensity ratios as *Critical Densification Levels* or Critical Points (CPs), at which we observe fair/equal load distribution among network tiers. Further, we define range of node

intensity ratios for which DUDe can be used as decoupling association window.

4.2 Contribution of the Work

In this Chapter, the Poisson random network was used to analytically obtain the relative densification levels for which we need the coupled associtions with offloading, decoupled and coupled UA. To validate our analysis, numerical and realistic network evaluations are used. We make use of Mathematica and Matlab software tools to compute the closed triple integrals and system level simulations, respectively. Then, using the WinProp software suite, we perform evaluations in a more realistic network environment. Specifically, the work has the following contributions.

- We use a Poisson random network to analytically obtain densification levels at which fair load share exists between tier-1 and tier-2 nodes in the UL and DL for randomly distributed UEs. We refer this densification levels as CPs.
- We derive the association windows, where users choose to use the decoupling association, coupled association with MC or LPN in terms of the relative intensity, transmit powers at each tiers and the PLE of the propagation environment.
- We formulate the ergodic rate expressions in order to study throughput performances in different densification regions which can be computed numerically.
- To validate the theoretical analysis, numerical, system level simulation and realistic network analysis are used. Our analytical, simulation, and realistic test case results provide insights for the operators about the densification ranges, where to use coupled or decoupled association.

4.3 Related Work

An important factor that restricts the UL capacity in dense HetNets is the problem of UL and DL imbalance. As there is a clear disparity between the transmission powers of the MCs and LPNs, the best serving cell for a user may be different in the UL and DL directions; hence, if the UL and DL associations are coupled, the UL capacity may

be severely limited, and this problem will become even worse in UDN and mmWave communications [131].

The work in [90] proposed the DUDe UA framework under a two-tier HetNets, in which the LPNs are randomly located over an MC's coverage area. DUDe framework makes it possible for an UE to select different optimal BSs in DL and UL according to its transmission requirements, which realizes a simultaneous optimal throughput over the two directions. Similarly, Feng et.al., in [132] developed a joint UA and resource partition framework for DUDe in a multi-tiered HetNets. Different from the traditional association rules such as max-RSS and CRE, a coalition game based scheme was used for the optimal UA with DUDe. However, these works fail to include all interference contributing factors and overlooks the effect of relative node intensities with respect to different UA schemes.

The derivation of association probabilities is used to calculate how the capacity is affected when the association is made either with LPN or MC in the UL or DL direction. In [133], the evaluation and comparison of the potential capacity gains of decoupled association of the UL to the LPN with respect to the MC, association that follows classical DL received power rule was performed. Smiljkovikj et al. [128] reported that as the density of the LPNs increases compared to the density of the MCs, a large fraction of UEs chooses to receive from a MC in the DL and transmit to a LPN in the UL. This clearly shows that the effect of further increase in relative node density between different tiers needs investigation.

Sial and Ahmed in [129] and [134], analyzed a UA technique for multi-tier 5G HetNets having dual connectivity and decoupled access or joint DUDe and dual association for spectrum aggregation in UL and DL. They have developed closed form solutions for association, coverage and outage probabilities along with average throughput by considering UL power control, receiver noise and multi-tiers of HetNets. The result shows that with the increase of LPN densities, more UEs prefer decoupled association. However, this preference may reduce in a highly dense HetNets where LPNs density is much more than MCs. They also found that the LPN densities and number of HetNets tiers play a significant role in improving the user performance in joint DUDe and dual association scheme. However, at what LPN density that UE performance starts to decrease was not answered.

A realistic scenario of a cellular network with different classes of real-world environ-

ments was used to analyze performance of a three-tier hybrid mmWave and ultra-high frequency network [135]. The authors investigated gains of DUDe technique. The real-world environments consists of two blockage environments: a sub-urban and a denser setting. However, the propagation model used may not give accurate result compared to the ray-tracing method.

To differentiate between deployment scenarios for which we can use coupled or de-coupled association, investigation of the relation between performance and cell density is important. The authors in [94] studied the dependency of DUDe performance with LPN density in two-tier network with 2x2 MIMO at each tier. The result shows that increasing the number of LPNs largely improves the performance of UEs initially but the gains are marginal after a certain density of LPNs. The question "What is this critical density level that marginal effect on throughput occurs?" must be investigated.

4.4 System Model

4.4.1 Stochastic Geometry Tools

The notion of irregularity and network topology is well studied and presented in [66] as associated with either the geometry or architecture. The BSs deployment and traffic spatial distribution are dually coupled [63] since BSs are built up to fulfill the traffic demand while data traffic is transmitted to mobile users through BSs. Therefore, the heterogeneity of the network infrastructure is the result of the irregularity in service demand distribution. However, the spatio-temporal dynamics in service demand cannot be fully satisfied with the fixed network infrastructure except for the mobility support.

It has been proved that stochastic geometry is the most convenient method to model communication systems. Traditionally, cellular networks have been modeled by using hexagonal grid model, with UEs either randomly scattered or placed deterministically [133]. However, for irregular deployments, like in HetNets, stochastic geometry appears to be a more realistic approach to evaluate the network performance. Especially, spatial point processes are used to account for stochastic properties of connectivity in wireless mobile systems such as cellular, ad-hoc and sensor networks from random locations.

By treating wireless communication network as Poisson random network, the PPP

has been widely used in modeling cell locations (in two-dimensional space) for dense and ultra-dense HetNets. Here, to provide a general definition for PPP as given in [35], let us consider a d-dimensional euclidean space, $\mathbb{R}^d$, with $d \geq 1$.

Definition 1. *The PPP Φ of intensity measure Λ is defined by means of its finite-dimensional distributions:*

$$\mathbf{P}\{\Phi(A_1) = n_1, ..., \Phi(A_k) = n_k\} = \Pi_{i=1}^{k}(e^{-\Lambda(A_i)}\frac{[\Lambda(A_i)]^{n_i}}{n_i!})$$

for every $k = 1, 2, ...$ and all bounded, mutually disjoint sets A_i for $i = 1, ..., k$.

The measure $\Lambda(\cdot)$ is called the *parameter measure* of the process. For the homogeneous process, $\Lambda(A) = \lambda(A)$, and for the inhomogeneous process, $\Lambda(A) = \int_A \lambda(x)dx$. Further, if Φ is a homogeneous PPP, λ is the intensity parameter. We directly use this definition for spatial point process, to model locations of BSs and UEs, on a two-dimensional plane, $\mathbb{R}^2$.

The visualizations of coverage shapes and topologies are also important. Poisson Voronoi tessellations (or also called diagrams) are useful for modeling and describing various natural patterns and for generating random lattices. In this work, we use Poisson Voronoi diagrams to display cell topology in dense and ultra-dense HetNets for the study of DUDe and MA.

By definition, a tessellation is a collection of open, pairwise disjoint polyhedra (polygons in the case of $\mathbb{R}^2$) whose closures cover the space, and which is locally finite (i.e., the number of polyhedra intersecting any given compact set is finite) [136].

In the two-dimensional case, the Voronoi diagram can be obtained from Delaunay triangulation. The Delaunay triangulation of a point set is a collection of edges satisfying an empty-circle criteria, which means that for each edge we can find a circle containing the edge's endpoints but not containing any other point from the initial set [38].

4.4.2 Network Topology

We consider a two-tier network where BSs at each tiers are located according to the homogeneous PPP represented by Φ_m and Φ_l with intensities (equivalently, densities) of λ_m and λ_l respectively for MC-tier and LPN-tier. A typical spatial coverage layout of the two-tier network deployment under consideration is shown using a Voronoi tessellation

with a normalized scale in Figure 4.1. We let the set of MCs denoted by $\mathcal{N}_m = \{BS_j :$

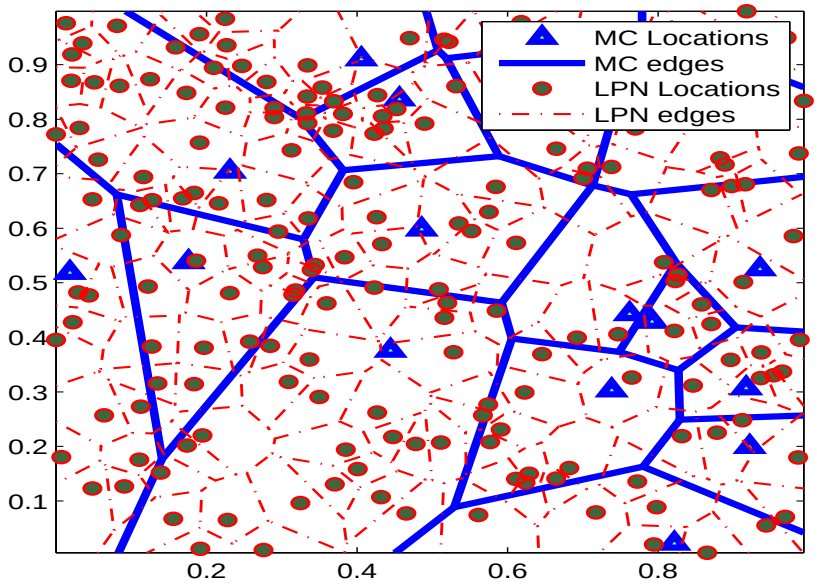


Figure 4.1: A view of Two-tier Poisson Random Network Deployment with cell boundaries corresponding to a Voronoi Tessellation with Normalized Dimensions.

$j = 1, 2, 3, \cdots, N_m\}$, set of LPNs denoted by $\mathcal{N}_l = \{lpn_j : j = 1, 2, 3, \cdots, N_l\}$ and a typical UE located at the center of the region $\mathcal{A}$ under consideration denoted by u. We also assume the users are located in the region according to the PPP denoted by Φ_u with intensity λ_u. The list and descriptions of the notations and parameters are provided in Table 4.1.

4.4.3 Link Model

For the link model, we assume that there is no intra-cell interference between users within the same cell as they can be assigned non-interfering set of resource blocks. However, users could suffer from inter-cell interference. Let the transmit power is denoted by P_k where k can be either the UE, MCs or LPNs, i.e., $k \in \{u, \mathcal{N}_m, \mathcal{N}_l\}$. In this case, the received power, P_k^{rx} at u or BS location in DL/UL at distance r_k from the serving BS is $P_k h_k r_k^{-\gamma_k}$, where h_k is a random variable that follows an exponential distribution with mean $1/\mu$, i.e., $h_k \sim \exp(\mu)$ and γ_k is path loss exponent.

The probability distribution function (pdf) of the distance $f(r, n)dr$ from an arbitrarily chosen origin (where a typical user u is supposed to be placed) to the n^{th} nearest neighbor

in the case of PPP is expressed as in (4.1) [37]:

$$f(r,n)dr = \frac{2(\pi\lambda_k)^n}{(n-1)!} r^{2n-1} e^{-\pi\lambda_k r^2} dr; \quad r > 0; n = 1, 2, 3, \cdots \tag{4.1}$$

Using the same expressions in [90], but considering a large number of MC and the interfering UE transmissions in the UL; both distributed according to the independent PPP, the SINR, ψ expressions from the UE at the center to the serving MC or LPN in the DL and UL at a distance r is given as in (4.2a) - (4.2d). Here, since the network will be interference limited the noise power can be neglected.

$$\psi^m_{UL}(r) = \frac{P_u h_m r^{-\gamma_m}}{\sum_{k\in\Phi_u\setminus u} P_u h_k r^{-\gamma_k}} \tag{4.2a}$$

$$\psi^l_{UL}(r) = \frac{P_u h_l r^{-\gamma_l}}{\sum_{k\in\Phi_u\setminus u} P_u h_k r^{-\gamma_k}} \tag{4.2b}$$

$$\psi^m_{DL}(r) = \frac{P_m h_m r^{-\gamma_m}}{\sum_{k\in\Phi_m\setminus m} P_m h_k r^{-\gamma_k} + \sum_{k\in\Phi_l} P_l h_k r^{-\gamma_k}} \tag{4.2c}$$

$$\psi^l_{DL}(r) = \frac{P_l h_l r^{-\gamma_l}}{\sum_{k\in\Phi_m} P_m h_k r^{-\gamma_k} + \sum_{k\in\Phi_l\setminus l} P_l h_k r^{-\gamma_k}} \tag{4.2d}$$

4.5 User Association and Critical Levels of Densification

In this section, we derive expressions for the UL and DL association probabilities and joint association probabilities. We make use of similar analytical derivation approaches using Poisson random network as in [128, 129, 132] and other literature for illustration and completeness of our discussion and provide tractable procedure for the readers. Our steps clearly shows the approach to identify critical densification levels and intensity ranges for decoupled and coupled user associations, which makes it different from aforementioned references. Also, we present the definitions and expressions for cell loads to be used later in the numerical evaluation.

4.5.1 DL and UL Association Probabilities

We begin with the UL association probabilities. The UL association probability of a user to an MC can be obtained considering the long term average received power based

Table 4.1: Notations and list of parameters in Chapter 4

Notations	Descriptions
N_m, N_l, and U	Number of MCs, LPNs, and UEs, respectively
$\mathcal{N}_m$, and $\mathcal{N}_l$	Set of MCs, and LPNs, respectively
Φ_m, Φ_l, and Φ_u	The PPP of MC, LPN, and UE locations, respectively
λ_m, λ_l, λ_u	Intensity of MCs, LPNs, and UEs, respectively
A	Two-dimensional area under considerations
u	A UE at the center
P_k	Transmit power for $k \in (u, \mathcal{N}_m, \mathcal{N}_l)$
P_k^{rx}	Received power at UE or BS locations
α	Intensity ratio between λ_l and λ_m
r_k	Distance of BS from the center
γ_k	Path loss exponent for $k \in (u, \mathcal{N}_m, \mathcal{N}_l)$
h_k	Exponential channel gain with mean $1/\mu$
$\overline{\gamma}$	Ratio of PLE
$\overline{P}$	Transmit power ratio
ψ	Instantaneous SINR
$\mathbf{P}$	Probability
$\mathbf{E}$	Expectation of a random variable
δ	Cell load
S^k	Resource of the k^{th} cell
s^k	Allocated resource units
n^k	Number of associated UEs to the k^{th} BS
R_z	Ergodic rate for $z \in (DL, UL, DL/UL)$
v	Association variable
$\mathcal{L}$	Laplace Transform

association as in (4.3).

$$\begin{aligned}
\mathbf{P}_{UL}^m &= \mathbf{E}_{r_m}[\mathbf{P}\{\mathbf{E}_h[P_u h_l r_l^{-\gamma_l}] < \mathbf{E}_h[P_u h_m r_m^{-\gamma_m}]\}] \\
&\overset{a}{=} \mathbf{E}_{r_m}[\mathbf{P}\{r_l^{-\gamma_l} < r_m^{-\gamma_m}\}] \\
&= \mathbf{E}_{r_m}[\mathbf{P}\{r_l > r_m^{\overline{\gamma}}\}] \\
&\overset{b}{=} \mathbf{E}_{r_m}[\exp\{-\pi\lambda_l r^{2\overline{\gamma}}\}] \\
&= \int_0^\infty \exp\{-\pi\lambda_l r^{2\overline{\gamma}}\} f_{r_m}(r,1) dr \\
&= \int_0^\infty 2\pi\lambda_m r \exp\{-\pi\lambda_l r^{2\overline{\gamma}}\} \exp\{-\pi\lambda_m r^2\} dr,
\end{aligned} \tag{4.3}$$

where, $\overline{\gamma} = \frac{\gamma_m}{\gamma_l}$. Here, the best serving MC is at distance r_m from the user and the nearest LPN is located at a distance of r_l. The $f_{r_m}(r,1)$ is the pdf of the distance between a UE and the serving MC. (a) follows from the exponential distributed h_k with mean $1/\mu$ and

the same UL transmit power of a user and (b) follows from the probability that no particle is found in a disk of area πr^2 in a two-dimensional PPP with intensity λ is $\exp\{-\pi\lambda r^2\}$.

For $\overline{\gamma} = 1$, the probability that a UE at the origin is associated to the MC-tier is

$$\mathbf{P}_{UL}^{m} = \frac{\lambda_m}{\lambda_m + \lambda_l} = \frac{1}{1+\alpha}, \tag{4.4}$$

where, $\alpha = \frac{\lambda_l}{\lambda_m}$ and the proof is as follows. Substituting the integration variable with $x = -\pi\lambda_m r^2$ and $dx = -2\pi\lambda_m r dr$ then, integrating and re-substitution in (4.5) gives the result in (4.4).

$$\begin{aligned} \mathbf{P}_{UL}^{m} &= -\int \exp\{x\frac{\lambda_m + \lambda_l}{\lambda_m}\}dx \\ &= \frac{-\lambda_m}{\lambda_m + \lambda_l}\exp\{-\pi(\lambda_m + \lambda_l)r^2\}|_0^\infty \end{aligned} \tag{4.5}$$

From (4.4), the UL association probability of a user to a LPN can be obtained as:

$$\begin{aligned} \mathbf{P}_{UL}^{l} &= 1 - \mathbf{P}_{UL}^{m} \\ &= \frac{\alpha}{1+\alpha} \end{aligned} \tag{4.6}$$

In a similar process, in the DL the probability that a UE is associated to the MC or LPN can be expressed as:

$$\begin{aligned} \mathbf{P}_{DL}^{m} = \int_0^\infty 2\pi\lambda_m r \\ \exp\{-\pi\lambda_l(\overline{P}^{\frac{-2}{\gamma_l}} r^{2\overline{\gamma}})\}\exp\{-\pi\lambda_m r^2\}dr \\ \text{and} \\ \mathbf{P}_{DL}^{l} = \int_0^\infty 2\pi\lambda_m r(1 - \exp\{-\pi\lambda_l(\overline{P}^{\frac{-2}{\gamma_l}} r^{2\overline{\gamma}})\} \\ \exp\{-\pi\lambda_m r^2\})dr, \end{aligned} \tag{4.7}$$

where $\overline{P} = \frac{P_m}{P_l}$ and integrating over the interval for $\overline{\gamma} = 1$ gives:

$$\begin{aligned} \mathbf{P}_{DL}^{m} &= \frac{1}{\alpha\overline{P}^{\frac{-2}{\gamma_l}} + 1} \\ &\text{and} \\ \mathbf{P}_{DL}^{l} &= \frac{\alpha\overline{P}^{\frac{-2}{\gamma_l}}}{\alpha\overline{P}^{\frac{-2}{\gamma_l}} + 1} \end{aligned} \tag{4.8}$$

Using (4.4), (4.6) and (4.8), we can state Lemma 4.5.1 as follows for equal load distribution between MC- and LPN-tiers.

Lemma 4.5.1. *(Points of Fair Load Distribution): For a given PLE $\gamma_1 = \gamma_m$ and transmit power ratio $\overline{P}$, the DL equal load share is obtained at $\alpha^* = \overline{P}^{\frac{2}{\gamma_l}}$ while that of the UL is obtained at $\alpha^* = 1$.*

Proof. From (4.4) and (4.6), the UL equal association probability to the MC and LPN is obtained when $\mathbf{P}_{UL}^m = \mathbf{P}_{UL}^l = 0.5$ which is for $\alpha^* = 1$. For the DL equal association probability to the MC and LPN, we equate $\mathbf{P}_{DL}^m$ and $\mathbf{P}_{DL}^l$ of (4.8) and solving for α gives $\alpha^* = \overline{P}^{\frac{2}{\gamma_l}}$. □

The relative node intensity range between $\alpha^* = 1$ and $\alpha^* = \overline{P}^{\frac{2}{\gamma_l}}$ represents the region of DL and UL load imbalance. Further, it can be observed that as tier-2 intensity increases or for $\alpha >> \alpha^*$, the UE tends to attach itself to the LPN-tier. Also note that the UL association probability is only affected by the relative density of nodes at both tiers not by the transmit power. In addition to the relative density of nodes, in the DL the association probability is also affected by both the transmit power at each tier and the PLE. Consider two PLEs $\gamma_l = 2$ and $\gamma_l = 4$ corresponding to rural and dense urban propagation environments, respectively. For a given transmit power ratio $\overline{P} = \frac{P_m}{P_l}$ (assuming $P_m > P_l$), the critical point of equal probability of association (α^*) is smaller when γ is larger. I.e., the load imbalance due to the large transmit power of the MC-tier can get better of at a smaller node intensity ratio α as shown in Figure 4.2. In the figure, we lebel CP1, CP2 or CP3 to indicate critical points of equal association probability in the DL for $\gamma =$ 4, 3, or 2, respectively.

4.5.2 Joint User Association Probabilities

The joint probability of UA in the UL and DL to the MC-tier or LPN-tier offers the opportunity to define the coupled and decoupled association regions. We identify three scenarios as in [128]:

- Case 1: User associated to MC both in the DL and UL or coupled association with MC (called Coupled-MC afterwards).
- Case 2: User associated to LPN both in the DL and UL or coupled association with LPN (called Coupled-LPN afterwards).

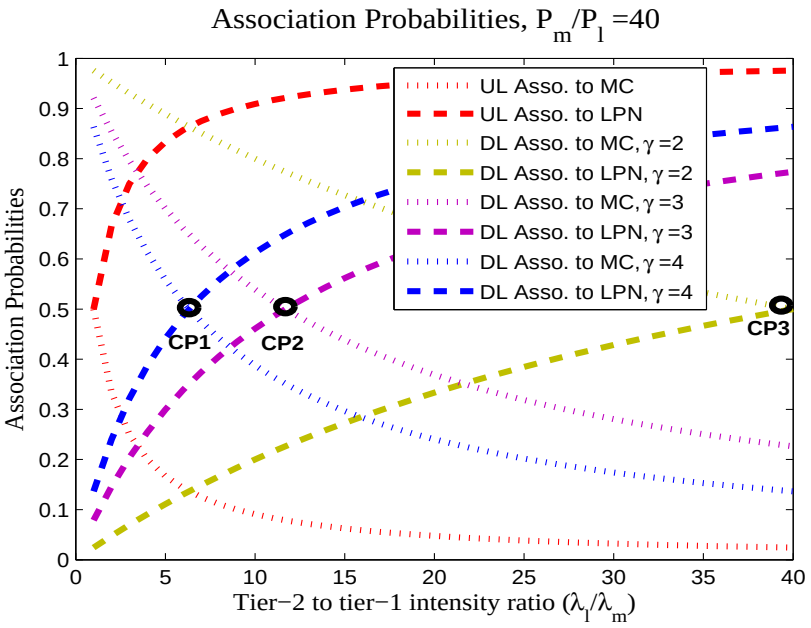


Figure 4.2: User Association Probabilities in the UL and DL for Different path loss Exponents

- Case 3: User associated to MC in the DL and to the LPN in the UL or decoupled association.

The association to the LPN in the DL and to the MC in the UL will not happen since user always tends to attach itself to the LPN in the UL as far as $\lambda_l \gg \lambda_m$ and to the MC in the DL for $P_m \gg P_l$. Here, Case 1 & 2 define the coupled association while Case 3 is for UL and DL decoupled association.

Case 1: Coupled Association (Coupled-MC)

The probability that a user will be associated to the MC-tier in both DL and UL is obtained from:

$$\begin{aligned}
\mathbf{P}^{m}_{DL/UL} = \\
\mathbf{E}_{r_m}[\mathbf{P}\{\mathbf{E}_{h_l}[P_u h_l r_l^{-\gamma_l}] \leq \mathbf{E}_{h_m}[P_u h_m r_m^{-\gamma_m}] \\
\bigcap \mathbf{E}_{h_m}[P_m h_m r_m^{-\gamma_m}] \geq \mathbf{E}_{h_l}[P_l h_l r_l^{-\gamma_l}]\}] \\
\stackrel{a}{=} \mathbf{E}_{r_m}[\mathbf{P}\{r_l^{-\gamma_l} \leq r_m^{-\gamma_m} \bigcap P_m r_m^{-\gamma_m} \geq P_l r_l^{-\gamma_l}\}] \\
= \mathbf{E}_{r_m}[\mathbf{P}\{r_l^{-\gamma_l} \leq r_m^{-\gamma_m}\}] \\
= \frac{1}{1+\alpha}.
\end{aligned} \tag{4.9}$$

In (4.9), (a) follows from the exponentially distributed gain h and taking the intersection completed the derivation.

Case 2: Coupled Association (Coupled-LPN)

Similarly, for the Case 2, the probability that a user will be associated to the LPN-tier in both DL and UL is obtained from the condition in (4.10) and is given in (4.11):

$$\begin{aligned}\mathbf{P}^{l}_{DL/UL} = \mathbf{E}_{r_m}[\mathbf{P}\{\mathbf{E}_{h_l}[P_u h_l r_l^{-\gamma_l}] \geq \mathbf{E}_{h_m}[P_u h_m r_m^{-\gamma_m}] \\ \bigcap \mathbf{E}_{h_m}[P_m h_m r_m^{-\gamma_m}] \leq \mathbf{E}_{h_l}[P_l h_l r_l^{-\gamma_l}]\}]\end{aligned} \tag{4.10}$$

$$\mathbf{P}^{l}_{DL/UL} = \frac{\alpha \overline{P}^{\frac{-2}{\gamma_l}}}{\alpha \overline{P}^{\frac{-2}{\gamma_l}} + 1} \tag{4.11}$$

From (4.9), and (4.11), it can be observed that the coupled association to the MC-tier is dominated by the UL long-term averaged received power while the coupled association to the LPN-tier is dictated by the DL long-term averaged received power.

Case 3: Decoupled Association

For the decoupled association, we consider UL association to the LPN-tier and DL association to the MC-tier. The association probability is obtained from the condition in (4.12) and is given in (4.13).

$$\begin{aligned}\mathbf{P}^{m/l}_{DL/UL} = \mathbf{E}_{r_m}[\mathbf{P}\{\mathbf{E}_{h_l}[P_u h_l r_l^{-\gamma_l}] \geq \mathbf{E}_{h_m}[P_u h_m r_m^{-\gamma_m}] \\ \bigcap \mathbf{E}_{h_m}[P_m h_m r_m^{-\gamma_m}] \geq \mathbf{E}_{h_l}[P_l h_l r_l^{-\gamma_l}]\}]\end{aligned} \tag{4.12}$$

$$\mathbf{P}^{m/l}_{DL/UL} = \frac{\alpha}{1+\alpha} - \frac{\alpha \overline{P}^{\frac{-2}{\gamma_l}}}{\alpha \overline{P}^{\frac{-2}{\gamma_l}} + 1} \tag{4.13}$$

Lemma 4.5.2. *(Decoupling Association Window): The maximum probability for decoupled association is found at $\alpha = \overline{P}^{\frac{1}{\gamma_l}}$ and the maximum decoupling association window is between $\frac{\overline{P}^{\frac{2}{\gamma_l}}}{\overline{P}^{\frac{2}{\gamma_l}}-2} < \alpha \leq \overline{P}^{\frac{2}{\gamma_l}} - 2$.*

Proof. The maximum probability for decoupled association can be readily obtained by taking the first derivative of (4.13). Then, equating to zero and solving for α gives the result. Since (4.13) is a concave function of α, the decoupling association window can be proved by evaluating the inequality $\mathbf{P}^{m}_{DL/UL} < \mathbf{P}^{m/l}_{DL/UL} \leq \mathbf{P}^{l}_{DL/UL}$. Substituting from (4.9), (4.11), and (4.13), and solving for α readily gives $\frac{\overline{P}^{\frac{2}{\gamma_l}}}{\overline{P}^{\frac{2}{\gamma_l}}-2} < \alpha \leq \overline{P}^{\frac{2}{\gamma_l}} - 2$. □

Observe that the decoupling association window is lower and upper bounded by points of UL and DL fair load distributions of Lemma 4.5.1, respectively. I.e., $\alpha^* = 1 < \frac{\overline{P}^{\frac{2}{\gamma_l}}}{\overline{P}^{\frac{2}{\gamma_l}}-2}$ and $\overline{P}^{\frac{2}{\gamma_l}} - 2 < \overline{P}^{\frac{2}{\gamma_l}} = \alpha^*$. Furthermore, for $\gamma_l = \gamma_m = 4$ and $\overline{P} = 40$, the maximum probability for decoupled association is found approximately at $\alpha = 2.5$.

Therefore, from preceding Lemma 4.5.1 and 4.5.2, we can state the main result as in the Theorem 4.5.1 without proof as it is a direct consequence of the previous discussions.

Theorem 4.5.1. *(Coupled and Decoupled Association Regions): Based on the critical densification levels and decoupling association window given above, we have the following three regions for flexible user association.*

1. $0 < \alpha \leq \frac{\overline{P}^{\frac{2}{\gamma_l}}}{\overline{P}^{\frac{2}{\gamma_l}}-2}$ – *Coupled association to MC; possibily with offloading*
2. $\frac{\overline{P}^{\frac{2}{\gamma_l}}}{\overline{P}^{\frac{2}{\gamma_l}}-2} < \alpha \leq \overline{P}^{\frac{2}{\gamma_l}} - 2$ – *Decoupled association*
3. $\alpha > \overline{P}^{\frac{2}{\gamma_l}} - 2$–*Coupled association to the LPN*

4.5.3 Number of users per cell and Cell loads

The number of users associated to the MC-tier in the DL is $|U|^m_{DL} = \mathbf{P}^m_{UL} \cdot |U|$, where $|U|$ is the total number of users. If we denote N_m as number of nodes in the MC-tier on area of $\mathcal{A}$, the number of users per cell can be obtained as:

$$n^m_{DL} = \frac{|U|^m_{DL}}{N_m} = \frac{\mathbf{P}^m_{DL} \cdot |U|}{\lambda_m \mathcal{A}} = \frac{\mathbf{P}^m_{DL} \cdot \lambda_u}{\lambda_m} \tag{4.14}$$

The number of users associated to a cell in MC-tier in the UL is given as:

$$n^m_{UL} = \frac{|U|^m_{UL}}{N_m} = \frac{\mathbf{P}^m_{UL} \cdot |U|}{\lambda_m \mathcal{A}} = \frac{\mathbf{P}^m_{UL} \cdot \lambda_u}{\lambda_m} \tag{4.15}$$

Similarly, the number of users per cell associated to LPN in the DL and UL are respectively given by:

$$n_{DL}^{l} = \frac{\mathbf{P}_{DL}^{l} \cdot \lambda_u}{\lambda_l}$$
$$\text{and} \tag{4.16}$$
$$n_{UL}^{l} = \frac{\mathbf{P}_{UL}^{l} \cdot \lambda_u}{\lambda_l}.$$

The cell load at each tier can be estimated assuming a general resource definition as follows. Let us denote a resource of a cell as S^k, where $k \in \{m, l\}$ from which s^k-units can be allocated to a user using RR scheduling. Therefore, the average cell load at the j^{th} BS of k^{th}-tier is given by

$$\delta_j^k = \frac{s^k \cdot n_{DL/UL}^k}{S^k}. \tag{4.17}$$

4.6 DL and UL Ergodic Rates

The achievable rate in the UL and DL for UEs associated with the MC or LPN can be obtained as a product of the association probabilities for the three cases and the achievable rate according to the Shannon's formula. Let v denote association to MC or LPN, i.e., $v \in \{m, l, m/l\}$ and z denote the direction, i.e., $z \in \{DL, UL, DL/UL\}$. Then, the ergodic rate R_z is obtained as follows:

$$R_z = R_z^v \cdot \mathbf{P}_z^v = \frac{1}{\ln(2)} \mathbf{E}_{r,\psi}[\ln(1 + \psi(r)_z^v)] \cdot \mathbf{P}_z^v \tag{4.18}$$

Here, we state the ergodic rates when a typical UE is associated to the MC or LPN in the UL or DL. To obtain the ergodic rates, we assume the interference in the UL is from all UEs transmitting to the LPNs or MCs except the UE at the origin; all of them scheduled on the same resource blocks. In the worst case, the number of interfering UEs scheduled on the same resource blocks becomes $N_m + N_l - 1$. We model this number of interfering UE as thinning of the original PPP with intensity λ_u^* and Lemma 4.6.1 gives the ergodic rates when the user is associated to the MC or LPN in the UL and the proof is provided in Appendix A.1.

Lemma 4.6.1. *(The UL ergodic rates)*

$$
\begin{aligned}
R_{UL}^{m} = \frac{1}{\ln(2)} \int_0^{\infty} \int_0^{\infty} 2\pi\lambda_m r \exp\{-2\pi\lambda_u^* \int_r^{\infty} (1- \\
\frac{1}{r^{\gamma_m}(e^y-1)x^{-\gamma_k}+1})\} \\
\exp(-\pi\lambda_m r^2) x dx dr dy \\
R_{UL}^{l} = \frac{1}{\ln(2)} \int_0^{\infty} \int_0^{\infty} 2\pi\lambda_l r \exp\{-2\pi\lambda_u^* \int_r^{\infty} (1- \\
\frac{1}{r^{\gamma_l}(e^y-1)x^{-\gamma_k}+1})\} \\
\exp(-\pi\lambda_l r^2) x dx dr dy
\end{aligned}
\tag{4.19}
$$

For the DL ergodic rate, we assume the interference is caused by all nodes of both tiers except the serving node. The expression is stated in Lemma 4.6.2 and the proof is provided in Appendix A.2.

Lemma 4.6.2. *(The DL ergodic rates)*

$$
\begin{aligned}
R_{DL}^{m} = \frac{1}{\ln(2)} \int_0^{\infty} \int_0^{\infty} \{\exp\{-2\pi\lambda_m \int_r^{\infty} (1- \\
\frac{1}{r^{\gamma_m}(e^y-1)x^{-\gamma_k}+1}) x dx\} \\
\times \exp\{-2\pi\lambda_l \int_r^{\infty} (1 - \frac{1}{\overline{P}^{-1} r^{\gamma_m}(e^y-1)z^{-\gamma_k}+1}) \\
z dz\}\} f(r,1) dr dy \\
R_{DL}^{l} = \frac{1}{\ln(2)} \int_0^{\infty} \int_0^{\infty} \{\exp\{-2\pi\lambda_l \int_r^{\infty} (1- \\
\frac{1}{r^{\gamma_l}(e^y-1)x^{-\gamma_k}+1}) x dx\} \\
\times \exp\{-2\pi\lambda_m \int_r^{\infty} (1 - \frac{1}{\overline{P} r^{\gamma_l}(e^y-1)z^{-\gamma_k}+1}) \\
z dz\}\} f(r,1) dr dy
\end{aligned}
\tag{4.20}
$$

From the discussion in Section 4.3 and (4.18), the UL and DL throughput performances depend on the association probabilities of UE to the LPN and MC and link rates. Based on the three user association cases in Section 4.5.2, we define the average/ergodic user rates in Table 4.2.

Table 4.2: Ergodic Rates with User Association Choices

Cases/ Direction of Connection	**Case 1**	**case 2**	**Case 3**
UL ($R_{UL} =$)	$\mathbf{P}^{m}_{DL/UL} \cdot R^{m}_{UL}$	$\mathbf{P}^{l}_{DL/UL} \cdot R^{l}_{UL}$	$\mathbf{P}^{m/l}_{DL/UL} \cdot R^{l}_{UL}$
DL ($R_{DL} =$)	$\mathbf{P}^{m}_{DL/UL} \cdot R^{m}_{DL}$	$\mathbf{P}^{l}_{DL/UL} \cdot R^{l}_{DL}$	$\mathbf{P}^{m/l}_{DL/UL} \cdot R^{m}_{DL}$

4.7 System Level Simulation and Numerical Evaluations

4.7.1 Cell Loads

In this section, we perform numerical evaluations and system level simulation on the densification level with respect to cell load distribution among BSs of the network.

The cell loads with respect to the intensity ratio between λ_l and λ_m for DL and UL associations to the MC and LPN, where $\gamma = 4$ were considered. From Figure 4.2, a DL equal load distribution is obtained at the Critical Point (CP) of $\alpha^* = \overline{P}^{\frac{2}{\gamma_l}}$. It was observed that an UL equal load share point appears before α^* because of the asymmetry between DL and UL. It was also shown that the DL CP shifts to the right as we decrease the PLE. With larger PLE, the case in dense urban deployment scenario, more UEs tend to associate with LPNs and the CP shift to the left.

In Figure 4.3, 4.4 and 4.5, we present the average cell load distribution with respect to the ratio of user to tier-1 and tier-2 intensity considering three densification levels: before, at the CP and after CP. Here, both tier-1 and tier-2 intensities are kept constant and the user intensity is varied.

At densification levels before the CP, cell loads on the MC-tier is higher in the DL. This densification level is where the DL transmit power of the MC-tier dominates and UEs tend to associate with tier-1. Therefore, it is where we need load-aware and cell-specific offsetting and adaptive inter-cell interference coordination.

In the case of densification level at the CP1 of Figure 4.2, we observe from Figure 4.4

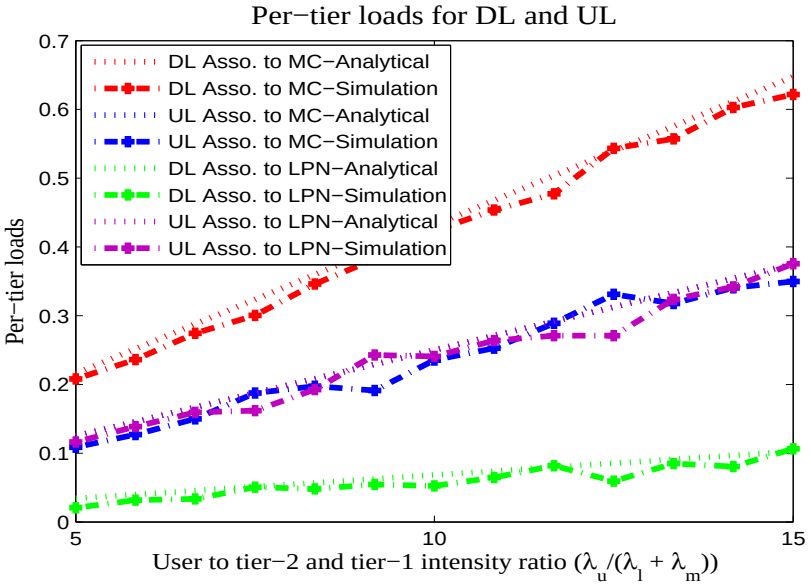


Figure 4.3: Per-tier Loads with respect to the ratio of User to tier-2 or tier-1 intensity, at UL-DL equal prob. (before CP1)

that although UE intensity λ_u grows with regard to the nodes intensity λ_l and λ_m, the DL load on MC-tier and LPN-tier increase at the same rate and remains fairly equal.

After CP, the reverse happens, where the DL cell loads on the LPN-tier becomes significant. Here, the number of UEs, which prefer to associate with LPNs, is much higher compared to the number of UEs which prefer to associate with MCs, both in the DL and UL. Hence, users tend to associate with LPN when $\alpha \gg \alpha^*$ and teir-2 BSs become loaded.

4.7.2 Average User rate

The triple integrals in (4.19) and (4.20) were integrated using the software tool Mathematica. Then, (4.18) was evaluated and the result was linked with Matlab, using Mathematica's 'matlink' application for further inquiry.

The UL link rate for the three cases with respect to the tier-2 intensity is shown in Figure 4.6. The result is plotted for $\gamma_l = 4$, $\lambda_m = 4/Km^2$, link rate threshold $y = 5bps/Hz$ and a channel bandwidth $s^k = 180KHz$.

As can be seen, for smaller relative intensity $\alpha \leq \frac{\overline{P}^{\frac{2}{\gamma_l}}}{\overline{P}^{\frac{2}{\gamma_l}}-2}$ coupled association to the MC gives better user throughput. As tier-2 intensity starts to increase, i.e., $\frac{\overline{P}^{\frac{2}{\gamma_l}}}{\overline{P}^{\frac{2}{\gamma_l}}-2} < \alpha \leq$

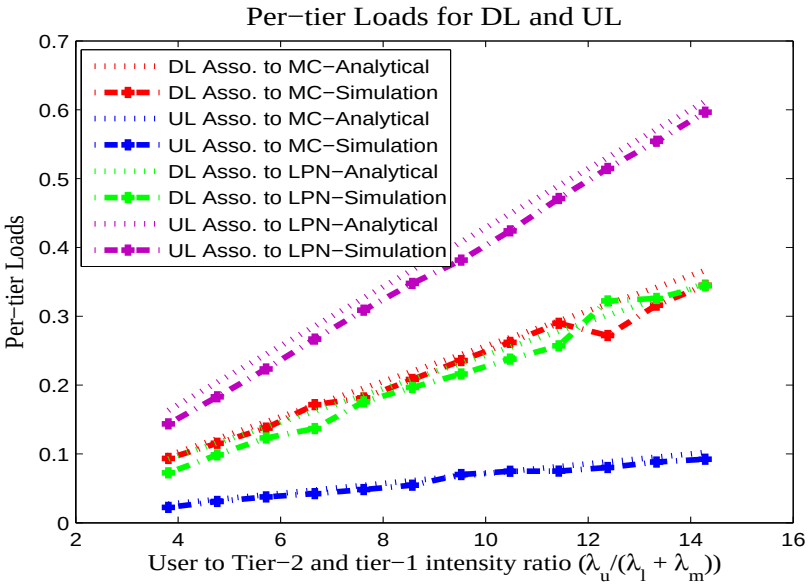


Figure 4.4: Per-tier Loads with respect to the ratio of User to tier-2 or tier-1 intensity, at the CP1

$\overline{P}^{\frac{2}{\gamma_l}} - 2$, UEs with decoupled association (case 3) gets better throughput. However, for $\alpha > \overline{P}^{\frac{2}{\gamma_l}} - 2$ coupled association to the tier-2 gives higher user throughput compared the other scheme. Also, notice that the decoupled association window gets decreased with increase of the PLE.

Figure 4.7 shows the DL link rate for the three cases with respect to the tier-2 intensity. The result is again plotted for $\gamma_l = 4$, $\lambda_m = 4/Km^2$, link rate threshold $y = 5bps/Hz$ and a channel bandwidth $s^k = 180KHz$. Similar to the UL, coupled association to the MC gives better DL user throughput for smaller relative intensity $\alpha < \frac{\overline{P}^{\frac{2}{\gamma_l}}}{\overline{P}^{\frac{2}{\gamma_l}} - 2}$ compared to case 2 and case 3. As tier-2 intensity starts to increase, i.e., $\frac{\overline{P}^{\frac{2}{\gamma_l}}}{\overline{P}^{\frac{2}{\gamma_l}} - 2} < \alpha \leq \overline{P}^{\frac{2}{\gamma_l}} - 2$, UEs with decoupled association (case 3) receive better throughput. But, for $\alpha > \overline{P}^{\frac{2}{\gamma_l}} - 2$ coupled association to the tier-2 gives higher user throughput. The numerical results for both the DL and UL throughput performances support Theorem 4.5.1 for a user association policy that may be applied for various densification intensities.

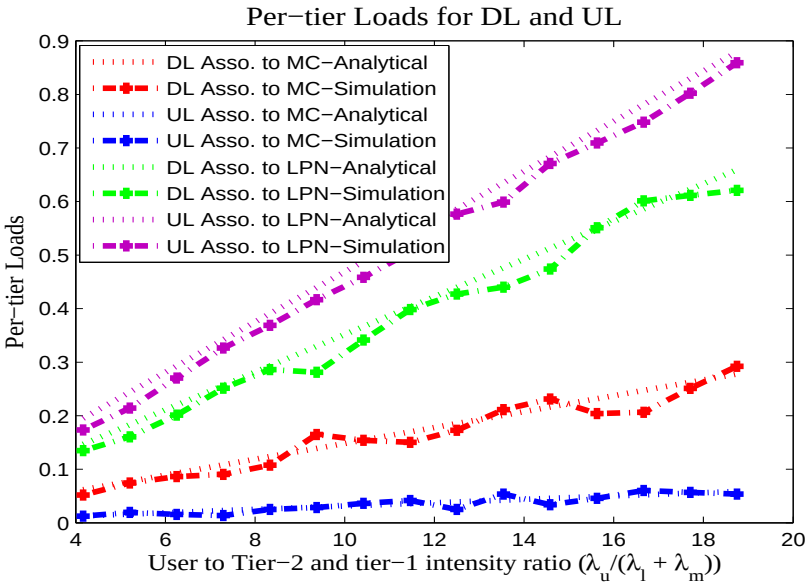


Figure 4.5: Per-tier Loads with respect to the ratio of User to tier-2 or tier-1 intensity, after CP1

4.8 Test cases in Realistic Scenario

In this section, we will go over how the system level simulator parameters are configured and used first. The use of a digital map is discussed. The deployment scenario is described in detail. The application of a ray-tracing-based signal map generating tool is also discussed. After that, the simulation results are presented and discussed.

4.8.1 Simulator Settings

A realistic scenario in Addis Ababa is considered (particularly, the Arat-kilo and Amist-kilo areas). This permits a comparison to be made between the numerical evaluation findings and the output that an operator could receive during deployment.

This scenario is used for comparison purpose with the previous numerical results in Figure 4.6 and 4.7. Hence, this scenario was designed to be analogous to the realistic one in its dimensions and number of MCs or LPNs. There are four MCs and a configurable number of LPNs in the realistic scenario. The locations of the MCs are taken from the existing deployment. As a densification layer, the LPNs are used. They are stationed on street corners and in strategic locations as hotspot service areas. All simulation parameters and values are listed in Table 4.3.

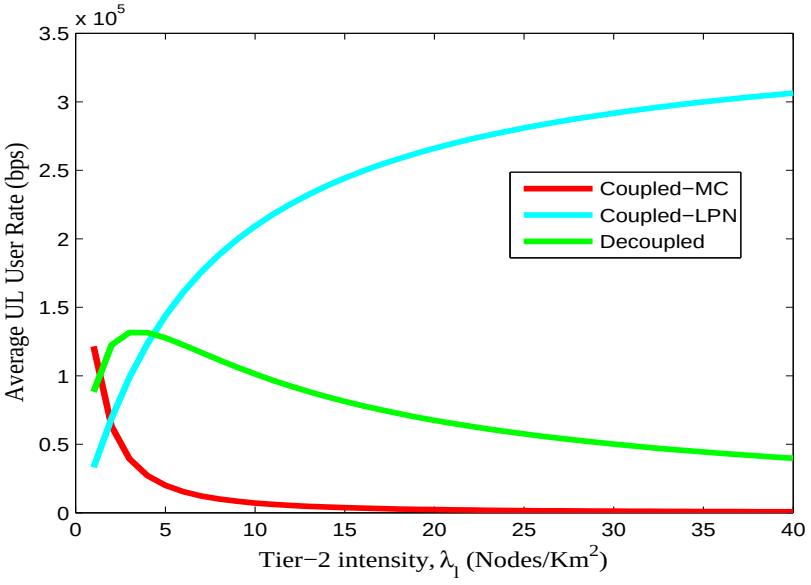


Figure 4.6: The UL User Rate for: Case 1(BLUE), Case 2(CYAN) and Case 3(GREEN)

Map and Transmitter Descriptions

The digital map for Arat-kilo and Amist-kilo area is used covering an area of around 1000m by 1000m. The satellite image of the deployment area is shown in Figure 4.8a with the topogray and building map shown in Figure 4.8b.

The terrain elevation of the area varies from 2430m - 2490m and the building heights vary from 4m-45m. We do not consider trees as its effect is assumed to be insignificant. A typical deployment is shown in Figure 4.9.

Path loss and simulator

As the realistic scenario represent the existing deployment, for the path loss and signal map computation, we consider both sectored and omni-directional transmitters. The received power and path losses are predicted at a receiver height of 1:5m from the ground for all transmitters. From the radio propagation and network planning tool WinProp, we used the DPM to generate the signal map, which guarantees accurate and confident results. Other important information on the transmitter settings are given in Table 4.3.

The Matlab based system level simulator generates signal map using selected path loss model. It also generates user location map and compute the received signal strength at each UE locations. The path loss at each pixel in the considered computation area generated from DPM is further processed using Matlab based simulator which is also used to implement the other empirical models discussed in the previous section.

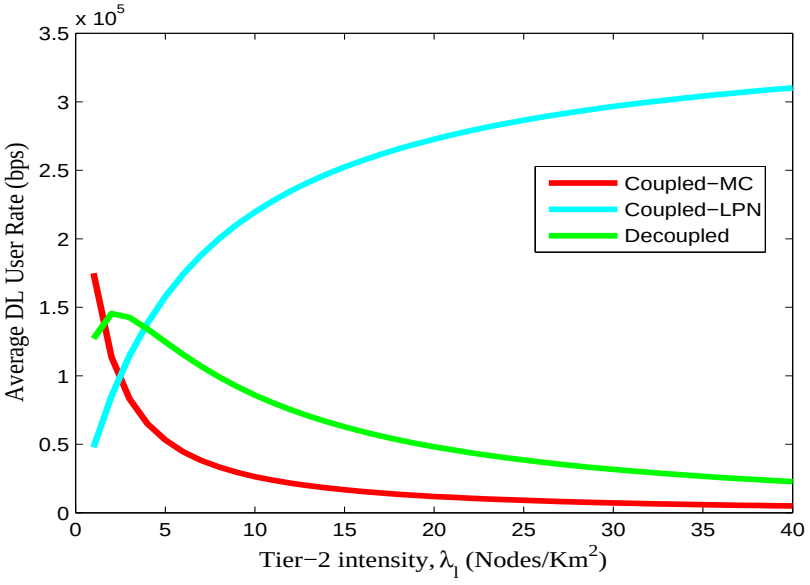
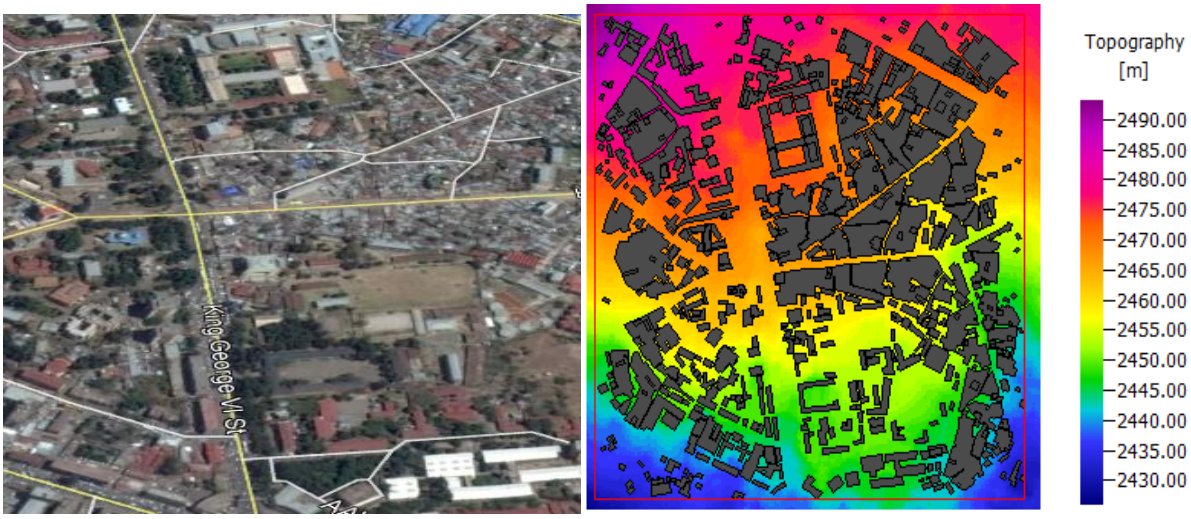


Figure 4.7: The DL User Rate for: Case 1(BLUE), Case 2(CYAN) and Case 3(GREEN)



(a) Satellite-view of the deployment area

(b) Topography and building map

Figure 4.8: Satellite image (a) and topography and building map(b) of the deployment area

4.8.2 Results and Analysis

We consider two performance metrics: cell-average and cell-edge user data rates. For both metrics, we consider different densification intensities such that a comparison can be drawn with the numerical and system level evaluation results presented earlier.

Cell-average user data rates

Table 4.3: Simulation Parameters and Values in Chapter 4

Parameters/descriptions	**Values**
Downlink transmit power	MC: 46dBm, LPN: 30dBm
Uplink transmit power	20dBm
Center frequency	1800Mhz
Bandwidth	20Mhz
Simulation area	1000 x 1000 m^2
Number of Snapshots	300
Spatial UE distribution models	Uniform Distribution
Antenna Heights	MC: 30m, LPN: 5m, UE:1.5m
Tx antenna	MC: Sectored, gain = 0dBi, LPN: Omni-directional, gain= 0dBi
Rx antenna gain	0dBi
Path loss model	DPM for realistic deployment $PL = A + 10\gamma \log(d/d_0)$, Fixed reference model, [137] $d_0 = 1m, \gamma = 3 - 4$ for Matlab-based simulator only
Noise power	-173dBm/hz
Deployment Scenario	4 Macro-sites, Variable number of LPNs

Since the test for exhaustive range of intensity level is difficult in the realistic scenario, strategic study was employed and four densification levels are considered (see Figure 4.10 and 4.11). These are intensity before the critical point, CP1 (or $\alpha = 1$), intensity at maximum probability of DUDe association (or approximately $\alpha = 2.5$), around the edge of the decoupled association window (or $\alpha = 4$) and beyond CP1 (or $\alpha = 7$). These intensity ranges are obtained by varying LPN deployments.

In the DL, the test case realistic network evaluation supports the numerical result. The best user throughput for DUDe is obtained when $\alpha = 2.5$ at which the decoupled association has the highest probability. Generally, the coupled-MC, coupled-LPN and decoupled associations gives performances that somehow goes with Theorem 4.5.1 in the DL. However, the UL throughput performance is in disagreement with the numerical

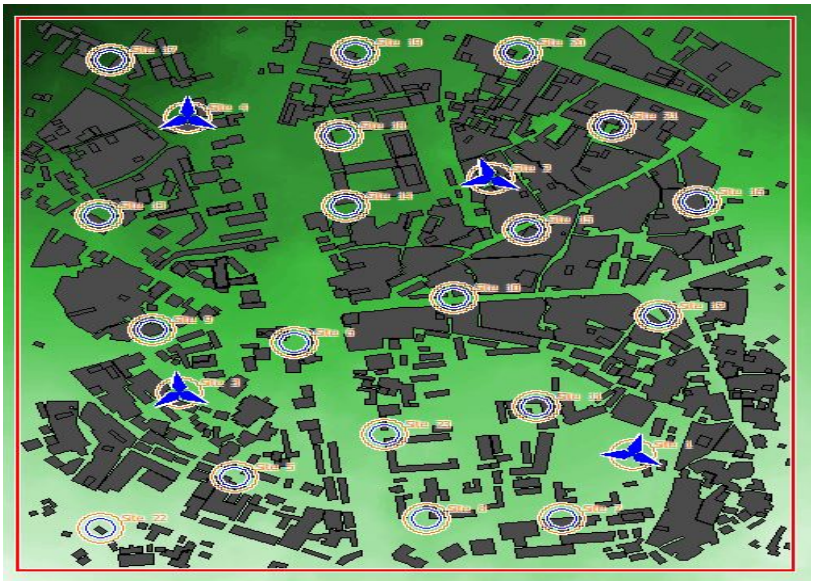


Figure 4.9: Deployment Scenarion of Test case area

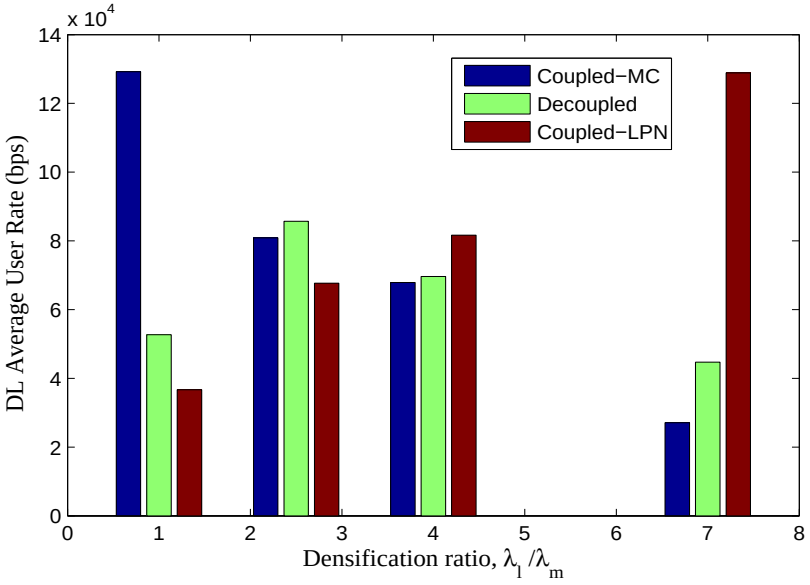


Figure 4.10: DL cell-average user rate at different densification ratios, $\lambda_m = 4/Km^2$

evaluation. The coupled association to the LPN tier offer best performance for $\alpha > 1$. We attribute the reason for the UL performance disagreement with Theorem 1 to the propagation environment (differences in PLE from the assumption) and test case deployment scenario (which may not exactly represent Poisson random network).

Cell-edge user data rates

Similar setting with the above is considered for both DL and UL cell-edge performance evaluations (see Figure 4.12 and 4.13). Again, the DL throughput performance from the

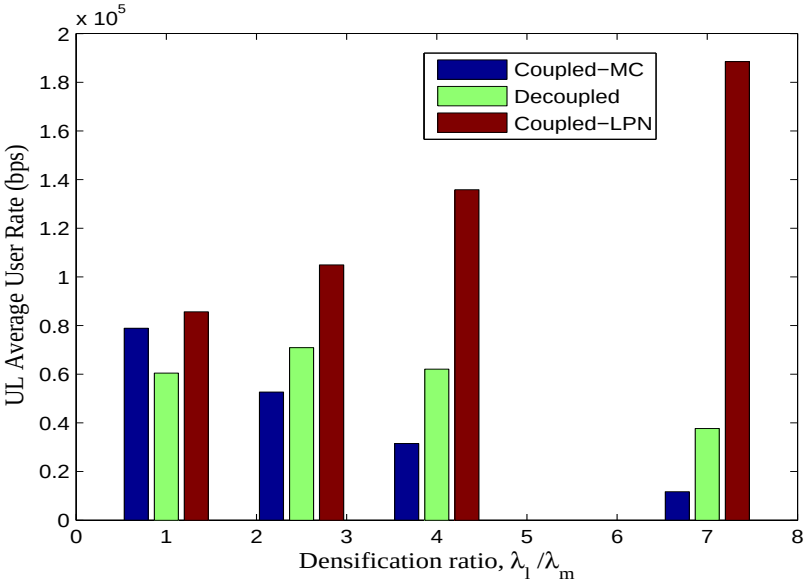


Figure 4.11: UL cell-average user rate at different densification ratios, $\lambda_m = 4/Km^2$

realistic network evaluation is somehow in agreement with Theorem 4.5.1, while the UL performance is in disagreement.

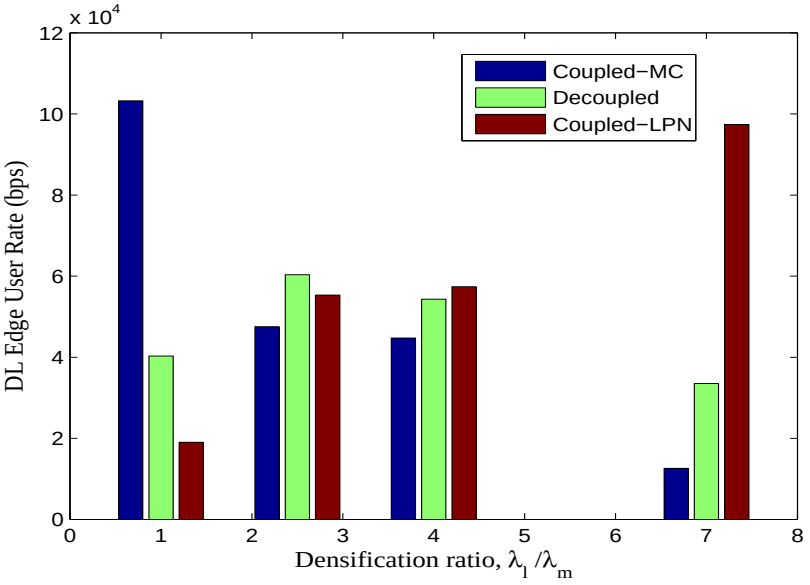


Figure 4.12: DL cell-edge user rate at different densification ratios, $\lambda_m = 4/Km^2$

4.9 Conclusions

Aggressive offloading of UEs from MC to LPN is the well established approach for load balancing in HetNets with coupled sub-optimal user association. However, as node inten-

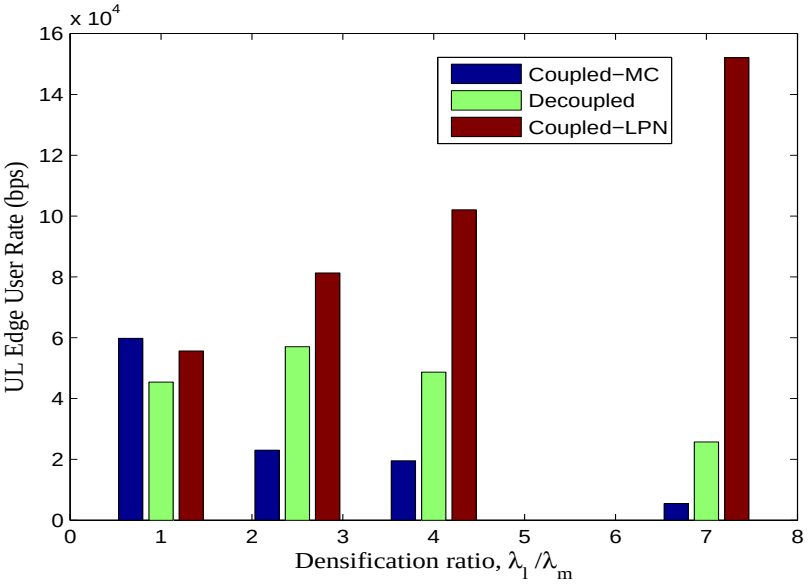


Figure 4.13: UL cell-edge user rate at different densification ratios, $\lambda_m = 4/Km^2$

sity increases the serving node becomes much closer to the UE. Therefore, the relative intensity at each tier can be used to parallel the effect of transmit power differences and to reduce the required load balancing effort. In ultra-dense heterogeneous wireless networks, the load distribution among different tiers changes with relative intensity ratios where different load balancing and interference coordination is required. Also, the user association choices impact the network performance differently for different relative intensities of nodes at each tiers.

In this chapter, we presented different critical densification levels at which fair load distributions are obtained at different tiers. The main result is that different relative node intensity can be considered for the choices of flexible user association schemes. For smaller relative intensity (or $\alpha \leq \frac{\overline{P}^{\frac{2}{\gamma_l}}}{\overline{P}^{\frac{2}{\gamma_l}}-2}$), coupled association to the tier-1 is preferred by users. In this sub-optimal association and offloading with appropriate interference coordination can be used to enhance capacity. For a medium relative intensity level ($\frac{\overline{P}^{\frac{2}{\gamma_l}}}{\overline{P}^{\frac{2}{\gamma_l}}-2} < \alpha \leq \overline{P}^{\frac{2}{\gamma_l}} - 2$), users prefer the decoupled association. At higher relative intensity ($\alpha > \overline{P}^{\frac{2}{\gamma_l}} - 2$), users choose to associate to the tier-2 both in UL and DL. The result has shown that there are cases for large PLE, like in dense urban deployment, where the decoupled association window becomes narrow and coupled association to LPNs gives the best user throughput. In this case, other capacity enhancement and mobility support approaches are required which will be part of the future work.

Chapter 5

Multiple association and mobility in ultra-dense networks

This chapter deals with the non-full-load case based on the work and contribution in [P2], where a cell may be idle or switched off at a time. From previous chapter, the critical densification level for decoupled association with LPN is obtained. For the cell density beyond this densification level, we can consider the gains from MA. So, we first present the backgrounds for MA considering the current trends in wireless network. Then, clustering and mobility models are discussed. Finally, evaluation in real network are presented with conclusion and recommendations.

5.1 Background

In UDNs, where many cells are found in a close-proximity, a UE may associate with multiple cells and increase its aggregate throughput by communicating on multiple links at the same time [138]. Because it allows for greater adaptability and flexibility, MA will deliver high throughput. MA schemes may also improve backhaul capacity by allowing for multiple selections and easing constraints [139, 27]. MA is sometimes called virtual cell association [140]. The idea of MA stems from DA that allows UEs to transmit/receive data simultaneously to/from two eNodeBs in order to boost the performance in HetNets with dedicated carrier deployment [141]. The DA scheme was proposed in the LTE Release 12 specification [142] by 3GPP as one of the most relevant technologies to accomplish even

higher per-user throughput and mobility robustness, and load balancing.

Given that a UE is configured with DA, it can be connected simultaneously to two eNodeBs: a Master eNodeB and a Secondary eNodeB, which operate on different carrier frequencies and are interconnected by traditional backhaul links (known as X2 interface) [143]. In DA network architecture [144], the Control (C)-plane is in charge of transmitting system information and controlling UE connectivity, while the User (U)-plane is in charge of UE-specific data. Considering this separation, the C- and U- Planes might not be transmitted by the same network node, which brings important new features that enable the DA, as explained in more details in [143].

In other schemes, the MA benefits from both DUDe architecture [93] and C- and U-plane splitting. When C/U-planes are split, the user can receive/transmit user data on multiple links based on the U-plane associations in the UL and DL. Therefore, as a result of utilizing the MA scheme and C/U-plane splitting, the traffic load of an UE can be spread over several BSs nearby (possibly serving cells at various network tiers).

Moreover, in a heterogeneous UDN, the handover between various RAT will produce a new issue as the UE moves from one location to another. This implies that there is a need for inter-RAT coordination for mobility support and also for authentication, authorization and accounting. Also, as cell densification increases, the serving node footprint decreases and the number of cell re-association increases, which can create overhead to the achievable capacity of the system. Therefore, UE mobility and temporal wireless channel fluctuations may cause unstable association in UDN due to the high cell density, which would lower performance. This intermittent UA and frequent transfer from one LPN to the other brings a challenge in UDN, which is a dominant problem in the case of mmWave communication. I.e., with a reduced cell sizes to tens of meters in 5G cellular networks, quickly moving terminals lead to frequent cell re-associations and additional latency is inevitably added [49, 114].

Due to the above problems, associating high speed user to LPNs which has a small footprint consequently results in overhead of frequent cell re-association. So, categorized UA based on the relative speed of UE is necessary. In this scheme, fast moving UEs chose to associate to MC while slow moving UEs associate to LPN [145, 115]. I.e., high mobility users should associate with MC so as to incur less beam switching overhead, whereas low

mobility ones should be associated with LPNs.

In mmWave communication, managing the co-channel interference due to network density is very crucial [146]. Using well designed clustering, not only reducing co-channel interference, but also the performance gains from the MA can be enhanced. In 5G networks, the move from cell-centric to user-centric clustering is getting attention. Similar to the defintion in [147], we define user-centric UDN as network where each UE is individually served by its nearby BSs. Such networks form a dynamic BS group or cell-cluster to enable MA, and resources are allocated to each user in a flexible and seamless fashion [148]. The challenge here is the clustering approach and the choice of parameters which defines the cluster elements.

5.2 Contribution

In this chapter, we develop a MA system with mobility support. A clustering scheme is devised to cluster cells that can potentially serve a given user to enhance the aggregate throughput. In addition to serving high data rate UEs, the clustering-based MA with mobility support creates a network, which is resilient to UE mobility and small footprint issues of the UDN. The contributions of this chapter can be summarized as follows.

- We develop a clustering approach, which produce virtual cells or group of cells with which UEs will be associated. Combining of MA with clustering improves collaboration between most appropriate cells to serve a given UE. Our clustering consists of two stages: a master cluster and a user-centric cluster. A master cluster is used for the mobility management while a user-centric cluster is used for serving high data rate UEs.

- We use mobility management approach in LTE-A/5G combining with MA. I.e., high mobility users associate with MC, whereas low mobility ones associate with LPNs. To enable this, we develop a scheme which attempts to separately treat UEs based on their speed by setting some predefined thresholds.

- Results are taken from synthetic and realistic network simulations which drives important conclusions. In both cases, the Brownian Motion (BM) is used to simulate a

given UE's mobility, and the negative exponential distribution is used to model each UE's velocity. Key performance metrics taken into account include spatial changes in SINR, user throughputs, and frequency of re-association.

5.3 Related Work

To the best of authors' knowledge, the work by Kamel et al. in [27] was the first work on UA in UDN, which considers MA. In a MA scenario, the UE distributes its traffic amongst many cells breaking the backhaul limitation of individual cells and aggregating higher data rates. The idle mode probability is shown to be less in large Multi-Cell sizes. This is due activation of more cells in the user neighborhood. The results also show that the ASE in case of MA improves significantly with larger Multi-Cell sizes, and with higher user densities. The analysis was limited to DL user association and mobility was not considered.

Kamel et al. [28] analyzed the average downlink rate of a multiple association in UDN environment. Mathematical formulation and simulation has shown that there is a significant increase in the average DL rate with the increasing LPNs density for different pathloss exponents. Also, turning off the idle BSs is found to be critical for the realization of the gains from network densification.

When assisted by C/U-plane splitting, traffic load of a user can be shared by multiple BSs (all potentially serving cells at each tiers in the network) in the vicinity of the user using MA scheme. A weighted sum-rate maximization for the DL transmission in Multi-cell MA in OFDMA-based HetNets was proposed in [102] by formulating a joint user association, sub-channel allocation, and power allocation optimization problem. Their work do not particularize the use of the DUDe and C/U-plane splitting options. The DUDe based association in [95] proposed an adaptive decoupling and MA technique that allows users to be connected to multiple BSs to maximize the spatial frequency reuse by allowing users access to more resources. They developed algorithm for choosing between alternatives whether the users in the UL session should be coupled/decoupled from the DL in addition to introducing a mechanism to decide whether users should be associated to a single BS or multiple BSs.

The control mechanism impacts the performance of a UA, classified as centralized or distributed control mechanisms. For instance, in LTE, mobility is managed by the mobility management entity with the serving gateway being the mobility anchor, and hence, it is a centralized solution. However, distributed control aids in decentralization of the conventional mobility management mechanisms, wherein the anchors are now dispersed rather than having a single anchor for all the data services of an UE. [149]. For both mechanisms, cell clustering play important role in MA. A cluster of LPNs form a virtual cell to serve a single UE. This enhanced solution is known as the soft-association control, in which each UE with MA capability is associated with a virtual cell [111]. The effort on soft-association can be enabled by C-RAN as proposed in [112] using joint UA and resource allocation supported by software-defined radio. However, further investigation is needed to take into account a truly UDN environment both in the C-RAN and dense LPN topologies [147]. This includes questions among others: how cells are to be clustered?, how is highly mobile UEs get served? The solutions to these problems should include dynamic clustering and adaptive mobility management among others.

In [150], dynamic clustering is proposed as the core function of user-centric UDN. Such network will create a dynamic set of active cells or group of cells to serve each UE. The choice of cluster center is also an important issue; especially, in implementation of the scheme in existing networks. Zhang and Huang [138], selects nearest cells in UE's vicinity to serve the user in the U-plane, while the best one among them serves the C-plane, which reduces re-association probability and network overheads without a loss in the throughput performance gain. In our work, we consider the MC as cluster center and believe that this further ease the mobility management problem. All the C-plane associations are made to the MC-tier while the U-plane service the data transmission and reception. We achieve this by developing a master and user-centric clustering approach.

The study in [151] proposes a novel adaptive cell selection scheme. The scheme adapts to various characteristics of ultra-dense HetNets and UE movements. It was shown that for low- and medium-speed vehicles, the scheme outperforms the traditional protocol in terms of the average number of re-associations by 42.39%. This shows that a categorized treatment of mobile users based on UE speed is important in dense HetNets and mobile environment. Speed-based optimization algorithm was utilized to adjust the handover

control parameters in 4G/5G networks [152]. Hence, improve mobility management. UE's received power and speed are used to adjust the handover margin and the time to trigger. The gains from multi-connectivity or MA and C/U-plane splitting must also be studied to harness the expected system performances.

Kela et al. [46] proposed a continuous UDN to provide 5G services to mobile UE which is found to outperform the widely accepted solution based on macro cells and massive MIMO systems. They designed a frame structure which carries UL pilot resources constituting the basis of mobility and user tracking. In their work, a continuous UDN is a network composed of a high number of LPNs providing continuous small cell coverage. Here too, there is profound motivation for intelligent clustering-based UA for improving the system performance.

In this work, we devise a scheme which attempts to separately treat UEs based on their speed by setting some predefined thresholds. In ultra-dense HetNets, MCs are still required for high-speed mobile users. Low and medium speed users can associate with LPNs to benefit from ample resources due to MA. In addition, we combine clustering with MA in order to enhance cooperation between most appropriate cells to serve a given UE. The master and user-centric clustering approach in our work solves the problem of small footprint LPNs and mobility management requirements.

5.4 System model and topology

5.4.1 Network topology and model

The network model considers a two-tier network in which the MCs and LPNs are identically and independently distributed according to the PPP in two dimension denoted by Φ_M and Φ_μ, with density λ_m and λ_l, respectively. Let a set denoted as $\mathcal{N}$ represent the totality of network nodes (both MCs and LPNs). Also, lets denote the sets of MCs and LPNs by $\mathcal{N}_m$ and $\mathcal{N}_l$, respectively. Here, a BS $B_j^k \in \mathcal{N}$ is defined as j^{th} BS in the k^{th}-tier and X_j^k is a location of B_j^k as (x, y, h).

The network diagram in Fig.5.1 shows the voronoi diagram of the BS coverages considering the shortest distance-based association for both tiers independently and the BM for the users, which is to be discussed in the next subsection. As can be seen from the

diagram, the overlapping coverages are due to the assumption that both tiers operate at different spectrum which will be refined later in connection with UA policies.

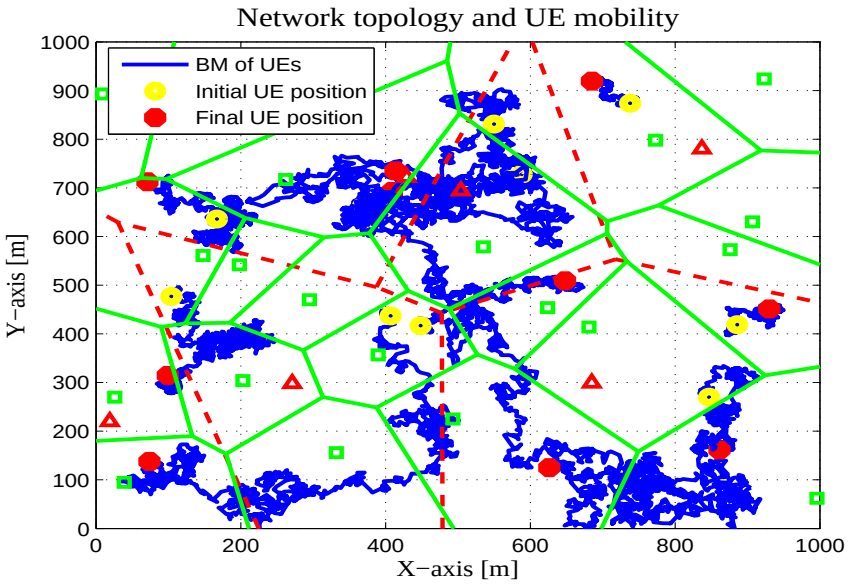


Figure 5.1: Network Topology with Voronoi tessellation: the coverage assumes shortest distance-based association, green lines for LPNs and red dotted lines for MCs coverage, the blue for the BM of sample users

The air interface technology under consideration is assumed to be the same for both tiers. However, the spectrum range used is below 6GHz for MCs or the conventional cellular network and the LPNs uses mmWave spectrum. OFDMA is considered and adjacent band interference is neglected. Furthermore, the effect of UE mobility on the propagation is assumed to be negligible. The effort ia mainly on reducing the ping-pong association that may result due to high speed of UEs. Also, the MCs are equipped with Massive MIMO capability with spatial multiplexing gain or antenna beamforming layers of m. We assume a Rayleigh fading channel, h with which the path loss PL is expressed as:

$$PL(d, h, \gamma) = PL(d_0)hd^{\gamma}, \tag{5.1}$$

where, d is the T-R separation distance, d_0 is the reference distance and γ is the PLE. Here, the PLE will be different for different tiers as the assumption above considers different spectrum.

In agreement with the above assumptions, the inter-tier interference is negligible and we consider intra-tier or ICI separately for each tier. For a transmit power of P_j and an

additive white Gaussian noise power density σ^2, we can express the SINR, ψ_j as:

$$\psi_j = \frac{P_j h_j d_j^{-\gamma}}{\sum_{b\in\{\Phi_M,\ \Phi_\mu\}} P_b h_b d_b^{-\gamma} + \sigma^2} \tag{5.2}$$

The operation $b \in \{\Phi_M,\ \Phi_\mu\}$ indicates either of the two tier will be considered at a time constraining the association to a single BS which is to be redefined later in this work to incorporate MA.

5.4.2 User Mobility Model

For simplicity of analysis using the tools from stochastic geometry, it is assumed that the initial position of the UE is at the center of the area of interest. Note that the homogeneous point process is also isotropic, i.e. the process is the same with respect to rotation around an arbitrarily chosen origin [37]. Let $\mathcal{U}$ represent a set of UEs in the area with numbers distributed according to Poisson Process and $u_{i,j} \in \mathcal{U}$ is the i^{th} UE associated to the j^{th} BS. Then for a UE at the center, r_i is the distance between $u_{i,j}$ and B_j^k. The distance (r_i) of a UE from the n^{th} closest BS can be expressed as a probability $f_n(r)$ as:

$$f_n(r) = \frac{2}{\Gamma(n)}(\lambda_b \pi)^n r^{2n-1} e^{(-\lambda_s \pi r^2)} \tag{5.3}$$

Where λ_b is the average number of BSs and $\Gamma(n)$ is the gamma function. The distance based UA for both tier is governed by (5.3) with the highest probability for 1^{st} closest BS which can be shown using Voronoi cells.

The mobility model uses the Brownian Motion (BM) or a random-walk with random step size and Fig.5.1 depicts BM of sample users implemented on Matalb based simulator. To conform to the mobility, let's relax the initial UE location from the center to that can be obtained from Poisson random distribution. Therefore, the spatial dimension of the initial UE locations can be modeled with the homogeneous PPP as follows, which gives the probability that u UEs exist in two dimensional area, $\mathcal{A}$ as,

$$\mathbf{P}(\mathcal{U} = u) = \frac{e^{-\lambda_u \cdot |\mathcal{A}|}}{u!}(\lambda_u |\mathcal{A}|)^u, u \epsilon Z^+, \tag{5.4}$$

where λ_u is the user density per KM^2.

For a UE at an initial location $l(t) = \{x(t), y(t)\}$ at time $t = 0$, if the user moves according to a BM, then at any time $t^{'} > 0$, the stochastic UE location $l(t^{'})$ is given by

[153]:

$$l(t^{'}) = l(t) + dl(t) = l(t) + \frac{\upsilon^2}{\pi} d\mathcal{W}(t) \tag{5.5}$$

Where $\mathcal{W}(t) = \{\mathcal{W}_x(t), \mathcal{W}_y(t)$ is a standard Wiener Process $(\mathcal{W}_x(t), \mathcal{W}_y(t) \sim \mathcal{N}(0,t))$ and υ is the drift speed.

The individual UE speed is modeled by a negative exponential distribution. I.e., there are a few number of UEs with high speed (υ greater than average speed).

5.5 Clustering and multiple association

5.5.1 Clustering

In Chapter 3, clustering was used for the purpose of finding appropriate LPN to offload a UE. Here, the purpose of using cell clustering is different. The clustered cells are used to serve a given UE with ample resource in order to meet the huge data demand.

Our clustering involves two stages: creating a master cluster and a user-centric cluster as shown in Fig. 5.2. The processes of developing both cluster stages and related requirements are presented next.

Master cluster

- Clustering should allow the MA system to support mobility. Here, UEs are categorized into high speed and low speed users. We denote the threshold velocity as v_{th} and individual UE speed as v_i. The cell clusters are to serve the low speed user while the MC-tier serve the high speed users.

- With the same procedure as in Subsection 3.5, cells that are highly interfering are grouped together. Let the neighbor list of the j^{th} cell cluster be contained in CC_j with cluster center MC_j. Each LPN $(LPN_k, \forall k \in \{1, 2, \cdots, N_l\})$ measures the interference in terms of the power received from the neighboring MCs at its own location $R^{rx}_{k,j}$. For an interference threshold I_{th}, select all MCs with an interference greater than the threshold by a certain amount β. In other words, $LPN_k \in CC_j$, $\forall j$, where $R^{rx}_{k,j} - \beta \geq I_{th}$. For a matrix M with size $N_m \times N_l$, the rows indicate MCs and the columns indicate the cluster elements or LPNs. If $LPN_k \in \{CC_j\}$, then the corresponding $M(j,k) = 1$; otherwise, $M(j,k) = 0$.

- The MC, or cluster center, is in charge of coordinating the communication between cluster elements.

User-centric cluster – After setting-up of the master clusters, a user-centric clusters are created as follows:

- A UE measures the signal strength of neighboring LPNs in the same cluster, $LPN_k \in \{CC_j\}$.
- According to the predefined threshold and/or possible number of serving cells, the UE selects LPNs and reports to the cluster coordinator.
- The cluster coordinator creates the user-centric cluster CC_i and assign resources.

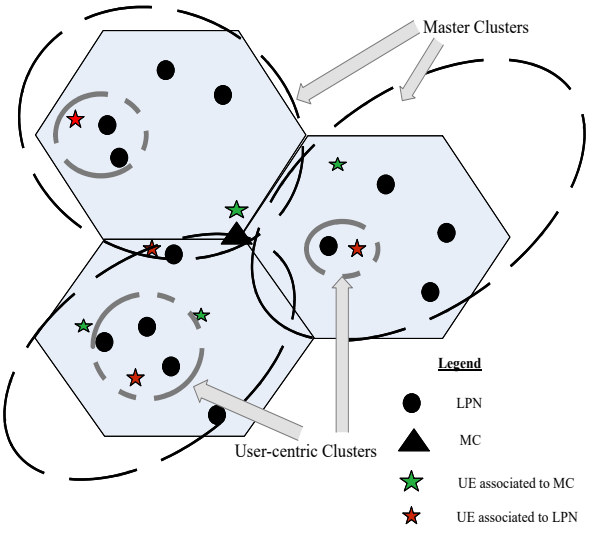


Figure 5.2: Master and user-centric clusters for Multiple Association

5.5.2 Algorithm for mobility, clustering and multiple association

A brief listing of procedures for the joint clustering, multiple association and mobility support is given in Algorithm 2. A detailed listing depends on the implementation options. The general routines involve initialization, master clustering, user velocity categorization

Algorithm 2 Procedure for mobility, clustering and multiple association

1: Initializations$\{\mathcal{N}, \mathcal{U}, P_t, P_{th}, v_{th}, v\}$
2: **for all** j in $\mathcal{N}$ **do**
3: $CC_j+ \leftarrow LPN_k$, if $R_{k,j} - \beta \geq I_{th}$ {For the master cluster, cluster the LPNs using interference relations between MCs and LPNs}
4: **end for**
5: **for all** UE_i in $\mathcal{U}$ **do**
6: **if** $v_i \geq v_{th}$ **then**
7: Associate UE_i to MC {use max-RSS for UA}
8: **else**
9: $CC_i+ \leftarrow LPN_k$, if $LPN_k \in CC_j$ and $R_{i,j} \geq R_{th}$ or $|CC_i| \leq K$ {Create user-centric cluster}
10: Associate UE_i to CC_i
11: **end if**
12: **end for**
13: **End**

and association, and user-centric clustering subroutines. The clustering subroutines involves measuring signals, building inter-tier interference relations, and populating cluster elements.

The master clustering subroutine is performed periodically or can be triggered owing to changes in the network deployment, which happens to be very rarely. Therefore, as clustering divides the problem space, the problem size is reduced to the size of each master cluster (N_c). Hence, we achieve the time complexity of $\Theta(N_c \cdot |\mathcal{U}|)$ for the user-centric clustering and UE association. The communication overheads, including UE measurement reports, are the same with other reference UA except for the clustering subroutines.

5.5.3 Performance of Multiple Association

The number of cells in a user-centric cluster and their average signal strength at the center or UE location determines the performance. For this, first, we need to obtain the expected number of cells in the user-centric cluster or cluster size, K.

The user neighborhood for a given UE is the set of LPNs where the average received

signal strength from a UE is above a given threshold, denoted by P_{th}. In other words, the user neighborhood is the set of LPNs which are in close proximity to this UE. Since the positions of the LPNs follows a PPP, the probability of finding k nodes in a disc of radius R and centered at the origin (where the typical user resides) is given by the Poisson distribution

$$\mathbf{P}(k) = \frac{e^{-\lambda_l \cdot \pi R^2}}{k!}(\lambda_l \pi R^2)^k. \tag{5.6}$$

This gives the probability that a UE associates with k cells. Moreover, the expectation of the neighborhood size or user-centric cluster size K is given by [27]

$$K = \lambda_l \pi \cdot R^2 = \lambda_l \pi \cdot (\frac{P_{th}}{P_t})^{-2/\gamma}, \tag{5.7}$$

where, P_t is transmit power of the LPNs and γ is the PLE.

A user-centric cluster can be of variable size based on the signal and interference reports of the UE. However, for the ease of implementation, we choose to work with fixed size clusters of K-cells. On the X2-interface, the cluster of K-cells coordinate transmission and reception. Therefore, the SINR in (5.2) is modified to account for this and is given by

$$\psi_{i\in\{\mathcal{U}\}} = \frac{\sum_{j=1}^{K} P_{i,j}^{rx}(b)}{\sum_{j\in\{\mathcal{N}_m,\, \mathcal{N}_l\}} P_{i,j}^{rx}(b) - \sum_{j=1}^{K} P_{i,j}^{rx}(b) + P_n(b)}. \tag{5.8}$$

The rate can then be estimated using modified Shannon equation, and for UE $i \in \mathcal{U}$, it is given as:

$$r_{i\in\{\mathcal{U}\}} = b \cdot m \cdot A \log_2(1 + B\psi_{i\in\{\mathcal{U}\}}) \tag{5.9}$$

As MA should guarantee the requirement that revenue obtained from rate gain is greater than the costs, we define the rate gain as $\triangle R = r_i^* - r_i$. In this scheme, as the mobility of UEs increase, more and more UEs associates with the MC-tier. The speed categorized association, at a cost of reducing unnecessary association and re-association (ping-pong), of course, the system performance is expected to decrease.

5.6 System Level Simulation

Our analysis set-up consists of two approaches. The first one is a system level simulation, where all the input parameters are synthetically generated. The second approach involves

Table 5.1: Simulation Parameters and Values in Chapter 5

Parameters/descriptions	**Values**
Downlink transmit power	MC: 43dBm, LPN: 30dBm
Spectrum for MC-tier	2.6 GHz
Spectrum for the LPN-tier	28 GHz
Sub-carrier spacing	LPN: 2x15 KHz, PRB = 360 KHz MC: 1x15 KHz, PRB = 180 KHz
Antenna beamforming layers	MC: m = 2, and LPN: m = 1
Deployment area	1000m by 1000m
Number of Snapshots	300
Number of UEs	50
Threshold velocity	40 Km/h
Spatial UE distribution models	-Initial UE location follows homogeneous PPP -UE mobility modeled as Brownian motion -Individual UE speed follows negative exponential distribution
Path loss model	-DPM (from WinProp) for realistic deployment -MC: 3GPP UMa, LPN: 3GPP UMi [154] for system-level simulator
Deployment Scenario	4 Macro-sites, 38 LPNs

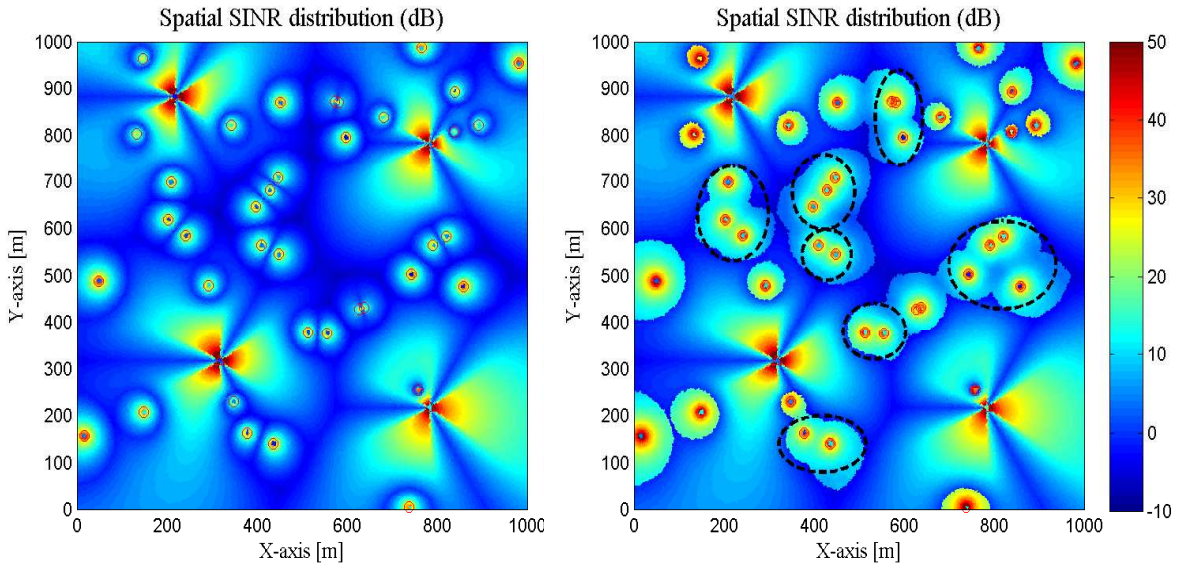
a test case in realistic scenario, which will be discussed in the next section. The former will be discussed here.

5.6.1 Simulation Scenario

The initial UE locations are randomly generated from PPP. Then, the BM model is used to generate mobility pattern for each users. The individual UE speed is generated form negative exponential distribution with different average speeds. There are four 3-sectored micro-sites in a deployment area of $1Km^2$. There are 38 LPNs overlying the tier-1 network, whose locations are randomly generated. This deployment scenario considers a total of $42BSs/Km^2$ which is a typical 5G UDN as stated in [10]. Certain parameters and values, if not explicitly changed or mentioned, are also used for the realistic test case simulation. Table 5.1 presents important simulation parameters and their corresponding values.

5.6.2 Result Presentation

The aims of the analysis are to show the effectiveness of our clustering approach and understand the gains from MA. In addition, the effect of mobility on the performance will br presented in the next section.



(a) No clustering (b) User-centric clustering with K=4

Figure 5.3: Spatial SINR distribution: No clustering (a) and user-centric clustering with K=4 (b)

Fig. 5.3a and 5.3b compares the spatial SINR distribution between the no-clustering and clustering with $K = 4$. User-centric clusters are shown by dotted ellipses. Here, in Fig. 5.3b, maximum of 4 elements are observed as $K = 4$. It is shown that with clustering, the SINR improves. It can be seen that cell-edge between non-clustered cells are avoided when K-cell cluster is used. This is because as UEs associated to the cluster almost receive the same performance.

Average UE throughput is shown in Fig. 5.4 for pedestrian users and different cluster sizes. For stationary and pedestrian users, similar throughput enhancements are observed at cell-edge, cell-average and cell-center as cluster size increases. As can be seen, the marginal rate gains decreases as cluster size increases. For this particular simulation, the marginal gains between $K = 1$ and $K = 2$, $K = 2$ and $K = 3$, and $K = 3$ and $K = 4$ are $188Kbps$, $105Kbps$ and $74Kbps$, respectively. This amounts to percentage gains of 15.5%,

7.5%, and 4.9%, respectively.

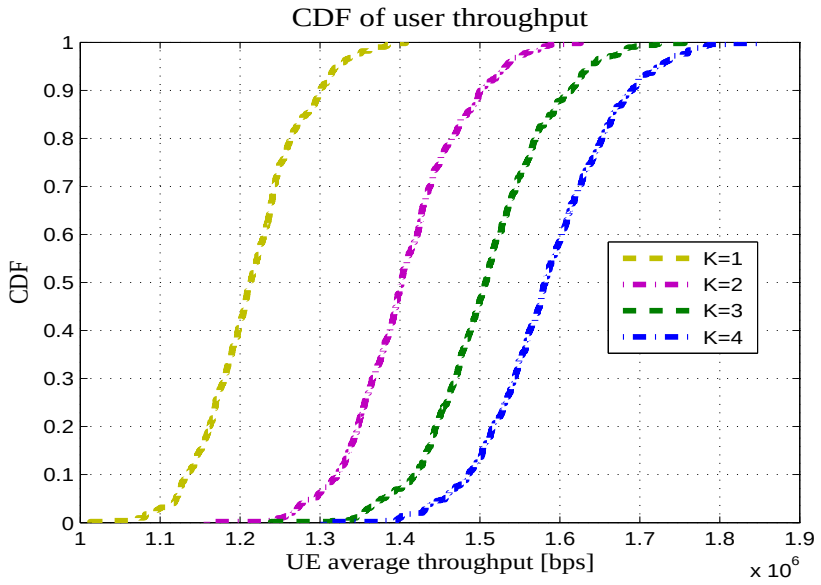


Figure 5.4: CDF of user throughputs for different cluster sizes (K)

5.7 Test Case in realistic Scenario

The test case realistic deployment involves the map and transmitter descriptions, used tools and result presentation.

5.7.1 Deployment scenario and evaluation setting

The digital map for Arat-kilo and Amist-kilo area is used covering an area of around 1000m by 1000m. The terrain elevation of the area varies from 2430m - 2490m and the building heights vary from 4m-19m. As in the case of synthetic simulation, we consider four 3-sectored micro-sites. There are again 38 LPNs overlying the MC-tier network. The deployment using the radio propagation and network planning tool called WinProp is depicted in Fig. 5.5 . The LPNs are deployed on lamp-posts, hot-spots and low coverage areas.

The received power and path losses are predicted at a receiver height of 1:5m from the ground for all transmitters. From the radio propagation and network planning tool WinProp, we used the Dominant Path Model (DPM) to generate the signal map. The path loss at each pixel in the considered computation area, generated from DPM is further

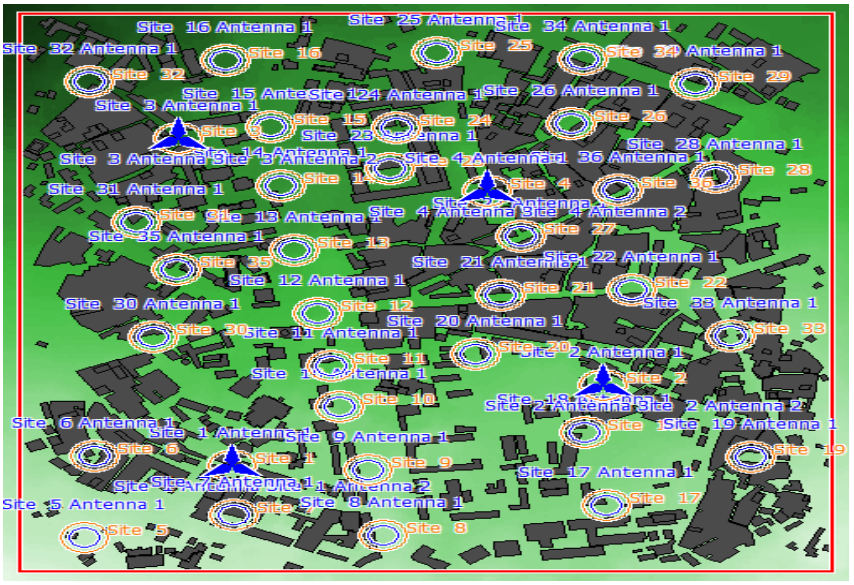


Figure 5.5: Deployment in the test case area

processed using Matlab-based simulator, which is also used to implement other empirical models discussed in the previous sections.

5.7.2 Result and Analysis

In addition to the clustering and MA effects on the performance, the influence of UE speed is presented in this section. It is assumed that pedestrian and low speed users are with speed $v \leq 10Km/h$. Other average UE speeds considered are $30Km/h$, $60Km/h$, $90Km/h$, and $120Km/h$. With the believe that the LPN can only serve low speed vehicles and pedestrian users, we consider threshold velocity of $40Km/h$ in our simulation. More dense network with smaller cell-coverage footprints compared to the current deploymet may require smaller velocity thresholds.

The re-association frequency per UE was one key performance indicator measured without mobility support and with designed mobility support. This was performed for both U-plane and C-plane associations. For each UE average speeds, we measure the average number of re-association for similar snapshots and in the same time period.

As can be seen from Fig.5.6 and 5.7, the U-plane and C-plane re-association increases as the UE mobility increases. In fact, the higher frequency of re-association is observed in the U-plane compared to the C-plane. This is because that the MA increases the probability that a UE changes from one user-centric cluster to the other as its speed gets

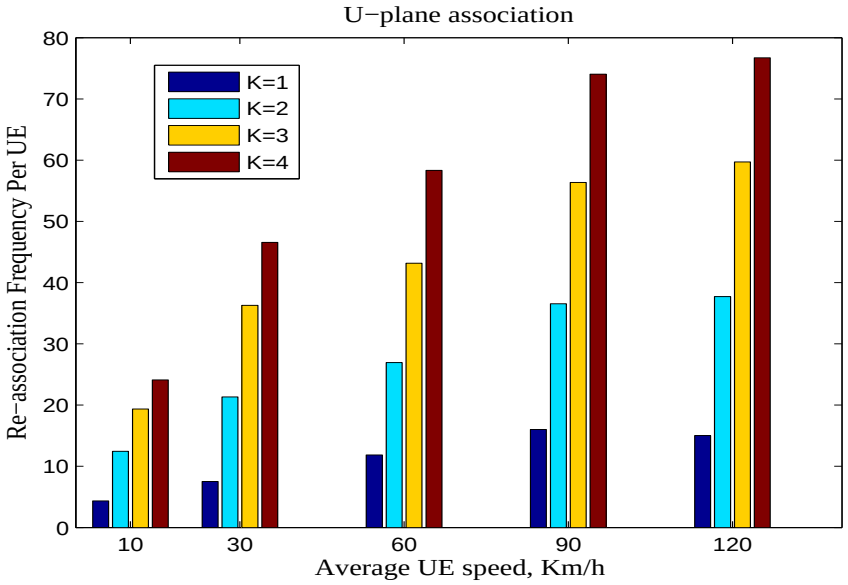


Figure 5.6: Re-association rate on U-plane without mobility support for different cluster size K and speed

increased. In the case of C-plane association, as MA only considered for the LPNs, there is change of re-association frequency with the change in average UE speed.

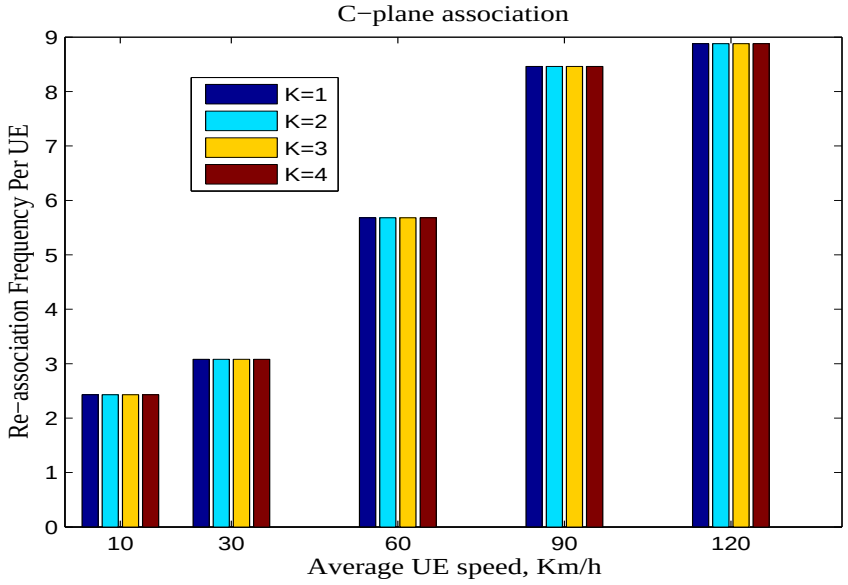


Figure 5.7: Re-association rate on C-plane without mobility support for different cluster size K and speed

From Fig.5.8 and 5.9, we observe that the U-plane and C-plane re-association increases as the UE mobility increases. The difference is that the frequency of re-association in the U-plane is now smaller compared to the case without mobility support. Typical re-

association frequency decrease when the speed is $90Km/h$ and $K = 4$ is 65.7%. This is a significant re-association frequency reduction as expected from the proposed mobility support design with MA and clustering. Again, in the case of C-plane association, as MA only considered for the LPNs, there is change of re-association frequency with the change in cluster sizes. The fact that Fig. 5.7 and Fig.5.9 are the same shows there is no effective change in C-plane association for the UE speed range under consideration. Therefore, the proposed clustering and MA with mobility support can be used to decrease the undesired U-plane re-association frequency without affecting the C-plane re-association frequency.

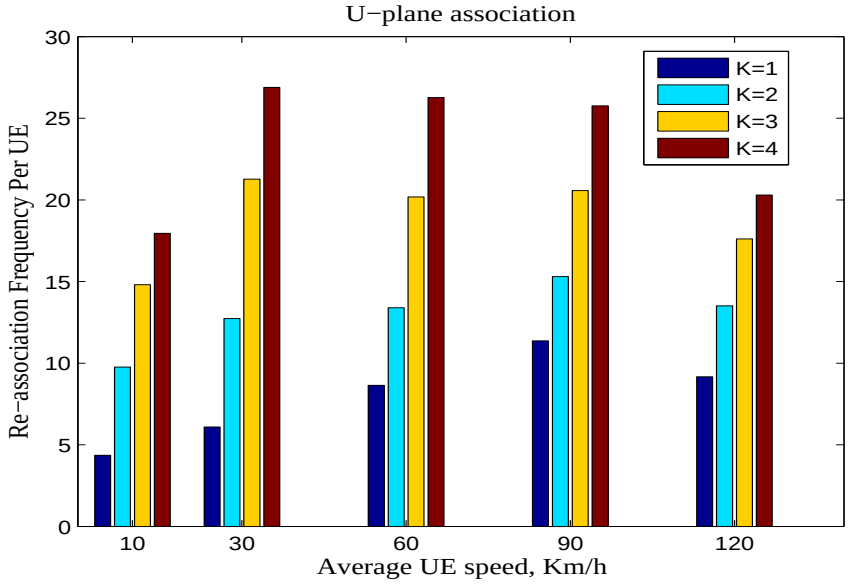


Figure 5.8: Re-association rate on U-plane with mobility support for different cluster size K and speed

Results in Fig. 5.10 and 5.11 shows UE throughput for different UE speeds and two cluster sizes: $K = 2$ and $K = 4$, respectively. In the realistic simulation, the gain from MA is only obsereved when UE speed is low and almost no gain is obtained when UE speed is beyond threshold value. This is because that more users begin to associate with MC-tier when UE speed gets increased. The 50^{th} percentile UE throughput when $k = 4$ is about $52.2Kbps$ above the case when $K = 2$ for pedestrian and low speed vehicle users, i.e., $\leq 30Km/h$. This amounts to 5.5% increase in user throughput. Another important observation is that in both cases, as UE speed increases the cell-center UE throughput decreases accompanied by cell-edge UE throughput increase. This may be attributed to the decrease in the interference level in the LPN-tier as more UEs associate with the

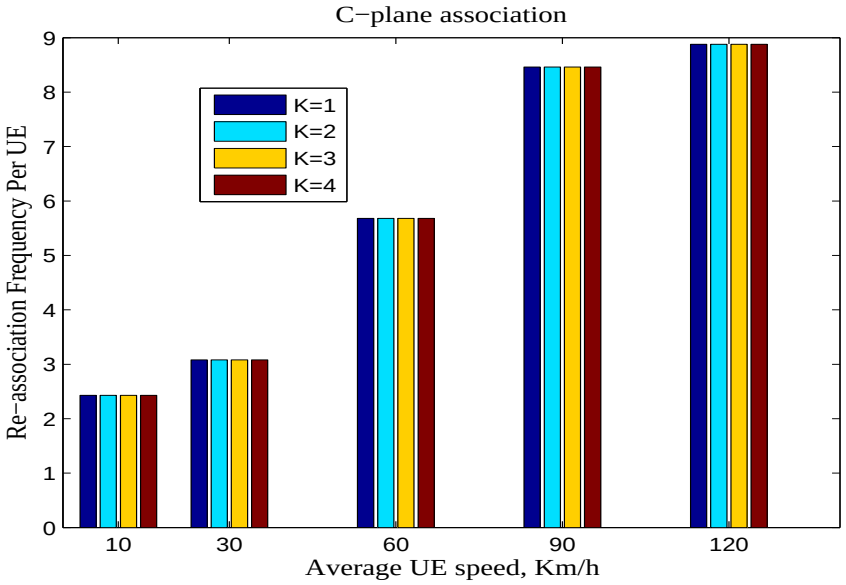


Figure 5.9: Re-association rate on C-plane with mobility support for different cluster size K and speed

MC-tier.

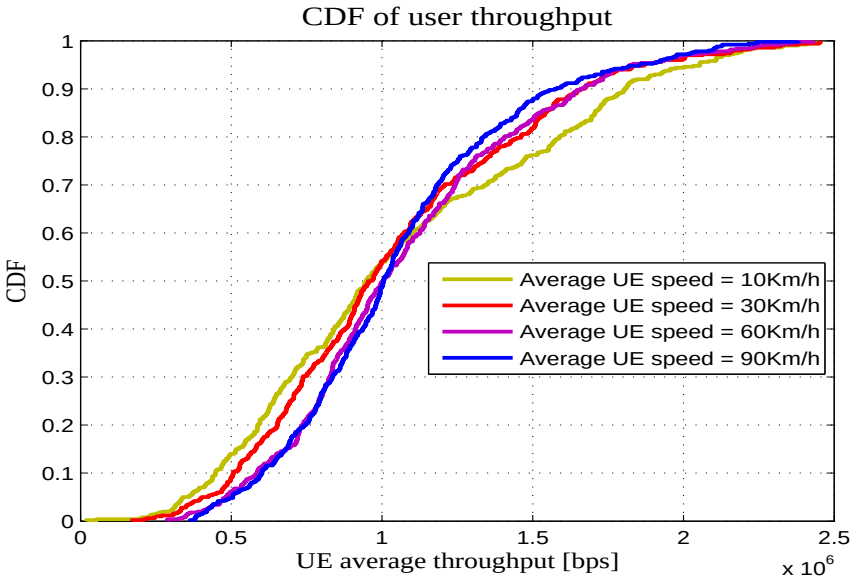


Figure 5.10: CDF of user throughput for different UE speed with K = 2

5.8 Conclusions

In this chapter, we introduced the MA and clustering to improve re-association frequency and throughput performance. To achieve this, we have created a technique for mobility

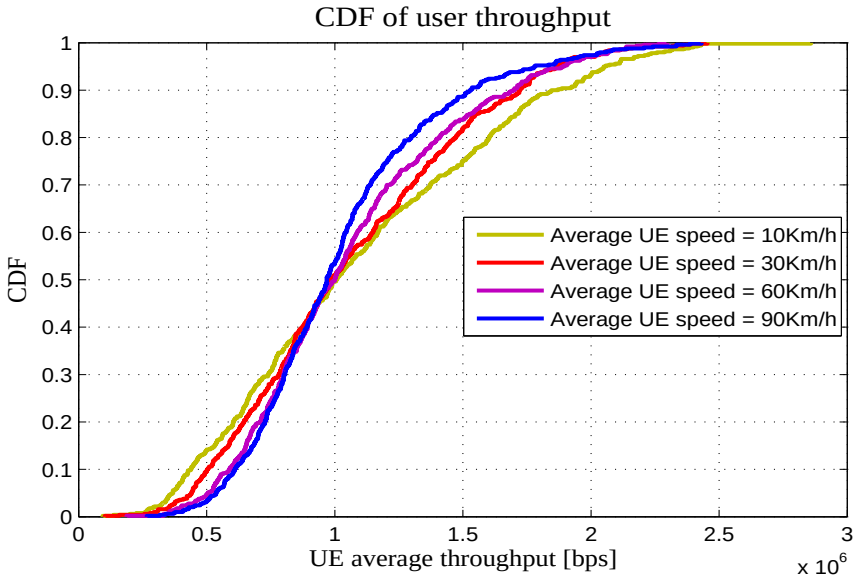


Figure 5.11: CDF of user throughput for different UE speed with K = 4

support in an UDN environment. The emphasis was on solutions for UEs with high data rate demands, as well as mobility management. C/U-plane split architecture was assumed for the development of mobility support. This goes with the approaches to mobility management in LTE-A/5G.

The developed clustering approach is divided into two stages: master and user-centric clustering. The master cluster is used to support the mobility of users and avoid frequent cell re-association, whereas user-centric clustering is used to serve the U-plane of UEs. Based on their speed category, UEs are assigned to either the MC or the clustered LPNs. The results were examined using synthetic and realistic test case simulations. The results show that MA with clustering improves performance while decreasing marginal gains with increasing cluster size. It is also demonstrated that the proposed scheme negatively impacts cell-center performance while improving the cell-edge performance. Using the mobility support, the re-association frequency reduced significantly for the U-plane while the C-plane remains unchanged. Therefore, the proposed clustering and MA with mobility support can be used to decrease the undesired U-plane re-association frequency without affecting the C-plane re-association frequency.

Chapter 6

Summary and Recommendations

6.1 Summary

The thesis is concerned with the concept of UA in dense HetNets. Chapter One focused on the motivation for the task, the problem definition, the objectives, and the methodologies. There was also a definition and discussion of basic terms. The background and state-of-the-art approaches in UA were covered in Chapter Two. Future challenges were also discussed. The author's main research contributions are chapters three, four, and five.

This chapter contains a summary of the methodologies, as well as the findings, conclusions, and recommendations. The presentation is sequenced in accordance with the work's objectives.

The flaring data service demand is driving the multitudes of technological advances in wireless communications. Among these is cell densification, where regional variation is observed due to the digital divide across the globe.

One technique in radio resource management is the UA. This work takes on different requirements for resource management based on the regional variation in cell densification. The approach was step-by-step treatment of the UA problems and corresponding challenges.

For less dense HetNets, where coupled association is still a viable solution, the focus is on the offloading and interference coordination. With this, cell load balancing is crucial way of performance enhancement and effective resource sharing. The first objective of this work was to come up with UA scheme, which is load-aware and adaptive resource

sharing. It employs clustering approach, which is effective in load sharing and easy for implementation in existing networks.

We introduce clustering-based LA-OLPS in dense HetNets. An iterative load-aware offsetting and an enumeration-based LPS configuration are formulated and evaluated using system-level simulations. The proposed scheme balances load among cells and converges after some iterations. Using clustering-based LA-OLPS, the LPN cell loads are more uniformly distributed than those in the case of the static offsetting and LPS configuration. The average user throughput is enhanced, particularly for cell-edge and cell-average users, without reducing the cell-center users' performance compared with LPS with static offsetting. The users require only the received signal strength and cell-specific offset values to initiate UA. Moreover, all optimization computations are performed on the network. This enables the easy implementation of the clustering-based LA-OLPS method in existing systems with neighbor-list management schemes. Our system modeling approach can also be used in further studies that aim at reducing the gap between cell-edge and cell-center users' performances and achieving a fair distribution of service provision. In addition, the results demonstrate performance changes in the four service regions.

Another aspect of the study of UA is whether to use coupled or decoupled association. For which densification level an operator choose coupled association over decoupled association or vise versa? The Fourth Chapter dealt with the answer for this critical problem in decision making in radio resource management. A critical densification level was identified using stochastic geometry tools. Intensive numerical evaluation, system level simulation and realistic test case simulations were carried out.

For full-load condition, the critical level of densification depends on the DL transmit powers, the PLE and intensity ratio between different tiers. Network operator can chose its plan based on its network's relative densification level. The choice between coupled and decoupled association is then a function of transmit powers, intensity ratio and the PLE.

In ultra-dense heterogeneous wireless networks, the load distribution among different tiers changes at a certain intensity ratio where different load balancing and interference coordination is required. Also, the user association choices impact the network performance differently for different relative intensities of nodes at each tiers. In this paper, we

presented the different critical densification levels at which user association choices can be optimized. For smaller relative intensity (or $\alpha \leq \frac{\overline{P}^{\frac{2}{\gamma_l}}}{\overline{P}^{\frac{2}{\gamma_l}}-2}$), coupled association to the tier-1 is preferred by users. In this sub-optimal association, offloading with appropriate interference coordination can be used to enhance capacity. For a medium relative intensity level ($\frac{\overline{P}^{\frac{2}{\gamma_l}}}{\overline{P}^{\frac{2}{\gamma_l}}-2} < \alpha \leq \overline{P}^{\frac{2}{\gamma_l}} - 2$), users prefer the decoupled association. At higher relative intensity ($\alpha > \overline{P}^{\frac{2}{\gamma_l}} - 2$), users choose to associate to lower tier both in UL and DL. In this case, other capacity enhancement and mobility support approaches are required.

The study of the non-full-load condition involves a number of complex and interplaying parameters. The first is the high data rate demand that can not be satisfied by a single serving cell. The other is, due to the high node intensity and hence small footprints, mobility management becomes crucial. Related to this, the control mechanism is important: centralized or decentralized. Whether to associate UEs to MCs or LPNs should consider UE speed and coordination mechanism should be devised. In this regard, Chapter Five focused on MA, clustering and association based on speed.

The developed clustering approach consists of two stages, namely: master and user-centric clustering. UEs are associated to either the MC or to the clustered LPNs based on their speed category. The performances area analyzed using synthetic and realistic test case simulations. Results show that MA with clustering enhances performance with decreasing marginal gains with respect to increasing cluster sizes. It is also shown that UE speed affects the cell-center performance.

One contribution of this work is the use of ray-trancing based real network simulation. Accurate digital maps were used with DPM of the WinProp tool to generate accurate signal map at the deployment area. In addition, this has enabled us study the implementation issues in the existing networks.

6.2 Recommendations

The author recommends the following as future research work:

- For the test cases in realistic network deployment, inclusion of ray-tracing based indoor propagation and its effect should be considered. This can capture the performance study of the vertical densification, as it is mainly indoor scenario.

- The densification with massive MIMO mounted on the MC towers has also gained momentum in the response to the huge data traffic demand. In addition, the high bandwidth provided by mmWave spectrum range added with massive MIMO brings a new paradigm in UA. In regards to this, UA to beams and related algorithms and interference management will be studied.

- Massive connection of devices with stand-alone batteries or connected renewable energy supply. A future work will also embark on UA as energy optimization mechanism.

- In cases where MA are possible in an UDN, and efficient and effective UA should consider the optimization options for instance between fiber or mmWave links. Therefore, intelligible UA algorithms which works in ultra-dense SC, massive MIMO, mmWave link and C-RAN network could potentially lead to the betterment of system performance. The problem of how different backhaul options impact the UA decision is not well studied and needs further research efforts.

Bibliography

[1] ITU, "ITU-R M.2370-0: IMT traffic estimates for the years 2020 to 2030," tech. rep., ITU, 2015.

[2] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2017-2022," white paper, Cisco, Feb 2019.

[3] 5GAmericas, "LTE to 5G: Cellular and Broadband Innovation," tech. rep., Rysavy Research/5G Americas, August 2017.

[4] Cisco, "Cisco Annual Internet Report (2018–2023)," white paper, Cisco, 2020.

[5] M. Olsson, C. Cavdar, P. Frenger, S. Tombaz, D. Sabella, and R. Jantti, "5GrEEn: Towards Green 5G Mobile Networks," in *IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Nov 2013.

[6] M. Agiwal, A. Roy, and N. Saxena, "Next Generation 5G Wireless Networks: A Comprehensive Survey," *IEEE Comm. Sur. & Tut.*, vol. 18, no. 3, pp. 1617–1655, 2016.

[7] E. Hossain and M. Hasan, "5G Cellular: Key Enabling Technologies and Research Challenges," *IEEE Instr. & Meas. Magazine*, vol. 18, pp. 11 – 21, Jun 2015.

[8] C. G. Guy and M. R. Tabany, "LTE and LTE-A Interworking and Interoperability with 3GPP and non-3GPP Wireless Networks," *Journal of Emerging Trends in Computing and Information Sciences*, vol. 4, Aug 2013.

[9] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. Teja, Sukhavasi, C. Patel, and S. Geirhofer, "Network Densification: The Dominant

Theme for Wireless Evolution into 5G," *IEEE Comm. Magazine*, vol. 52, pp. 82–89, Feb 2014.

[10] X. Ge, S. Tu, G. Mao, C.-X. Wang, and T. han, "5G Ultra-Dense Cellular Networks," *IEEE Wireless Comm.*, vol. 23, pp. 72–79, February 2016.

[11] Nokia, "Ultra Dense Network (UDN) White Paper," tech. rep., Nokia, 2016.

[12] M. Ding, D. López-Pérez, G. Mao, P. Wang, and Z. Lin, "Will the Area Spectral Efficiency Monotonically Grow as Small Cells Go Dense?," *2015 IEEE Global Communications Conference (GLOBECOM)*, 2015. [online] available at: http://arxiv.org/abs/1505.01920.

[13] W. Yu, H. Xu, H. Zhang, D. Griffith, and N. Golmie, "Ultra-Dense Networks: Survey of State of the Art and Future Directions," in *2016 25th Internation Conference on Computer Communication and Networks (ICCCN)*, pp. 1–10, 2016.

[14] D. López-Pérez, M. Ding, H. Claussen, and A. Jafari, "Towards 1 Gbps/UE in cellular systems: Understanding ultra-dense small cell deployments," *IEEE Comm. Sur & Tut.*, vol. 17, no. 4, p. 2078 – 2101, 2015.

[15] SCF, "Network densification in the 5G era," tech. rep., Small Cell Forum, Reprot Version 201.10.01, 2017.

[16] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-Dense Networks: A Survey," *IEEE Comm. Sur. & Tut.*, pp. 1–24, 2016.

[17] P. Cerwall, A. Lundvall, P. Jonsson, and S. C. et. al., "November 2020 Ericsson Mobility Report," tech. rep., Ericsson, November 2020.

[18] E. Mutafungwa, L. Baracchi, O. Renda, and G. Micheletti, "Study on Small Cells and Dense Cellular Networks Regulatory Issues," white paper, Global vision, standardisation & stakeholder engagement in 5G, Global5G.org, Jan 2018.

[19] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Drone Small Cells in the Clouds: Design, Deployment and Performance Analysis," in *IEEE Global Comm. Conf. (GLOBECOM), San Diego, CA*, pp. 1–6, 2015.

[20] D. Amzallag, R. Bar-Yehuda, D. Raz, and G. Scalosub, "Cell Selection in 4G Cellular Networks," *IEEE Trans. on Mob. Comp.*, vol. 12, pp. 1443–1455, Jul 2013.

[21] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective," *IEEE Wireless Comm.*, vol. 21, pp. 118–127, Jun 2014.

[22] J. Sangiamwong, Y. Saito, N. Miki, T. Abe, S. Nagata, and Y. Okumura, "Investigation on Cell Selection Methods Associated with Inter-cell Interference Coordination in Heterogeneous Networks for LTE-Advanced Downlink," in *Wireless Conference 2011 - Sustainable Wireless Technologies (European Wireless), Vienna, Austria*, pp. 1–6, 2011.

[23] S. Corroy, L. Falconetti, and R. Mathar, "Cell association in small heterogeneous networks: Downlink sum rate and min rate maximization," in *2012 IEEE Wireless Comm. and Net. Conf. (WCNC), Shanghai*, pp. 888–892, 2012.

[24] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "User Association for Load Balancing in Heterogeneous Cellular Networks," *IEEE Trans. on Wireless Comm.*, vol. 12, pp. 2706–2716, Jun 2013.

[25] J. G. Andrews, S. Singh, Q. Ye, X. Lin, and H. S. Dhillon, "An Overview of Load Balancing in HETNETs: Old Myths and Open Problems," *IEEE Wireless Comm.*, vol. 21, pp. 18–25, Apr 2014.

[26] H. Hu, J. Weng, and J. Zhang, "Coverage Performance Analysis of FeICIC Low Power Subframes," *IEEE Trans. on Wireless Comm.*, pp. 1–12, May 2016.

[27] M. I. Kamel, W. Hamouda, and A. M. Youssef, "Multiple Association in Ultra-Dense Networks," in *IEEE ICC 2016 - Wireless Comm. Symposium*, pp. 1–6, 2016.

[28] M. Kamel, W. Hamouda, and A. Youssef, "Performance Analysis of Multiple Association in Ultra-Dense Networks," *IEEE Trans. on Comm.*, vol. 65, no. 9, pp. 3818–3831, 2017.

[29] V. Poirot, "Energy Efficient Multi-Connectivity for Ultra-Dense Networks," thesis, Department of Computer Science, Electrical and Space Engineering, Lulea University of Technology, 2017.

[30] UNCTAD, "Digital Economy Report2021: Cross-border data flows and development: For whom the data flow," tech. rep., United Nations Conference on Trade and Development (UNCTAD), 2021.

[31] K. Okeleke, D. George, and E. Obiodu, "5G in Sub-Saharan Africa: laying the foundations," tech. rep., GSMA Intelligence, 2019.

[32] V. Degli-Esposti, "Ray Tracing Propagation Modelling: Future Prospects," in *The 8th European Conference on Antennas and Propagation (EuCAP 2014)*, p. 2232, 2014.

[33] H. Zhang, H. Ji, and Y. Wang, "User Association Scheme In Heterogeneous Networks Considering Multiple Real-World Policies," in *National Doctoral Academic Forum on Information and Communications Technology*, pp. 1–7, 2013.

[34] H. Munir, S. A. Hassan, H. Pervaiz, Q. Ni, and L. Musavian, "User Association in 5G Heterogeneous Networks Exploiting Multi-Slope Path Loss Model," in *IEEE*, 2017.

[35] F. Baccelli and B. Blaszczyszyn, *Stochastic Geometry and Wireless Networks: Volume I - Theory*. NoW Publishers, 1, pp.150, 2009, Foundations and Trends in Networking Vol. 3, Nos. 3-4, 2009. DOI: 10.1561/1300000006.

[36] J. Kingman, *Poisson Processes*. Oxford Studies in Probability, Oxford University, 1993.

[37] D. Moltchanov, "Distance distributions in random networks," *Ad Hoc Networks*, vol. 10, no. 6, pp. 1146–1166, 2012.

[38] J.-S. Ferenc and Z. Néda, "On the size-distribution of Poisson Voronoi cells." arXiv:cond-mat/0406116v2 [cond-mat.soft] 20 Feb 2008.

[39] P. Morters and Y. Peres, *Brownian Motion*. May 2008. Draft version.

[40] AWE, "WinProp: Propagation Modeling and Network Planning Software Suite." *AWE Communications GmbH.* Internet: www.awe-com.com, accessed on: Jun 24, 2015.

[41] R. Wahl and G. Wölfle, "Combined Urban and Indoor Network Planning Using the Dominant Path Propagation Model," in *Proceeding of EuCAP 2006, Nice, France*, Nov 2006.

[42] T. Nakamura, A. Benjebbour, Y. Kishiyama, S. Suyama, and T. S. Imai, "5G Radio Access: Requirements, Concept and Experimental Trials," *IEICE Trans. Comm.*, vol. E98–B, pp. 1396–1406, Aug 2015.

[43] J. Madden, "Cloud RAN or Small Cells?," tech. rep., Mobile Experts, Apr 2013.

[44] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A Survey on 3GPP Heterogeneous Networks," *IEEE Wireless Comm.*, pp. 10–21, Jun 2011.

[45] A. G. Gotsis, S. Stefanatos, and A. Alexiou, "Optimal User Association for Massive MIMO Empowered Ultra-Dense Wireless Networks," in *International Conference on Communication Workshop (ICCW)*, pp. 2238–2244, 2015.

[46] P. Kela, X. Gelabert, J. Turkka, M. Costa, K. Heiska, K. Leppänen, and C. Qvarfordt, "Supporting Mobility in 5G: A Comparison Between Massive MIMO and Continuous Ultra Dense Networks," in *IEEE ICC 2016 - Mobile and Wireless Networking Symposium*, pp. 1–6, 2016.

[47] W. Liu, S. Han, C. Yang, and C. Sun, "Massive MIMO or Small Cell Network: Who is More Energy Efficient?," in *Wireless Comm. and Networking Conference Workshops (WCNCW), 2013 IEEE*, pp. 24–29, 2013.

[48] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!," *IEEE Access*, vol. 1, pp. 335–349, 2013.

[49] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-Wave Cellular Wireless Networks: Potentials and Challenges," in *Proc. of the IEEE*, vol. 102, pp. 366–385, Mar 2014.

[50] C. J. Hansen, "WIGIG: Multi-GIgabit Wireless Communications In the 60 Ghz Band," *IEEE Wireless Comm.*, pp. 6–7, Dec 2011.

[51] T. Huynh, K. Kuroda, and M. Hasegawa, "User Association for Massive MIMO Cellular Networks with Small Cell Wireless Backhaul," in *19th International Symposium on Wireless Personal Multimedia Comm. (WPMC)*, 2016.

[52] K. Boulos, M. E. Helou, and S. Lahoud, "RRH Clustering in Cloud Radio Access Networks," in *International Conference on Applied Research in Computer Science and Engineeing (ICAR)*, pp. 1–6, 2015.

[53] D. Liu, L. Wang, Y. Chen, M. Elkashlan, K.-K. Wong, R. Schober, and L. Hanzo, "User Association in 5G Networks: A Survey and an Outlook," *IEEE Comm. Surv.& Tut.*, vol. 18, no. 2, pp. 1018–1044, 2016.

[54] E. Rakotomanana and c. G. Fran "Optimum Biasing for Cell Load Balancing Under QoS and Interference Management in HetNets," *IEEE Access*, vol. 4, 2016.

[55] M. I. Kamel and K. M. F. Elsayed, "ABSF Offsetting and Optimal Resource Partitioning for eICIC in LTE-Advanced: Proposal and Analysis using a Nash Bargaining Approach," in *IEEE ICC 2013 - Wireless Networking Symposium*, pp. 6240–6244, 2013.

[56] M. S. ElBamby, M. Bennis, and M. Latva-aho, "UL/DL Decoupled User Association in Dynamic TDD Small Cell Networks," in *2015 International Symposium on Wireless Communication Systems (ISWCS)*, 2015.

[57] J. Choi, W.-H. Lee, Y. Kim, and S.-C. Kim, "Dynamic User Association and eICIC Management in Heterogeneous Cellular Networks," in *IEEE International Conference on Comm. (ICC), Kuala Lumpur*, pp. 1–6, 2016.

[58] C. Bottai, C. Cicconetti, A. Morelli, M. Rosellini, and C. Vitale, "Energy-Efficient User Association In Extremely Dense Small Cells," in *Networks and Communications (EuCNC), Bologna*, pp. 1–5, 2014.

[59] D. Liu, Y. Chen, K. K. Chai, T. Zhang, and M. Elkashlan, "Two-Dimensional Optimization on User Association and Green Energy Allocation for HetNets With Hybrid Energy Sources," *IEEE Trans. on Comm.*, vol. 63, pp. 4111–4124, Nov 2015.

[60] T. Han and N. Ansari, "Green-energy Aware and Latency Aware User Associations in Heterogeneous Cellular Networks," in *Globecom 2013 - Wireless Networking Symposium*, pp. 4946–4951, 2013.

[61] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: a survey," *IEEE Comm. Magazine*, vol. 46, pp. 59–67, Sep 2008.

[62] M. Mirahsan, R. Schoenen, and H. Yanikomeroglu, "HetHetNets: Heterogeneous Traffic Distribution in Heterogeneous Wireless Cellular Networks," *IEEE Journal on Selected Areas in Comm.*, vol. 33, pp. 2252–2265, Oct 2015.

[63] M. Li, Z. Zhao, Y. Zhou, X. Chen, and H. Zhang, "On the Dependence Between Base Stations Deployment and Traffic Spatial Distribution in Cellular Networks," in *23rd International Conference on Telecomm. (ICT), Thessaloniki*, pp. 1–5, May 2016.

[64] M. Laner, P. Svoboda, and M. Rupp, "Parsimonious Network Traffic Modeling By Transformed ARMA Models," *IEEE*, vol. 2, pp. 40–55, Jan 2014.

[65] D. Lee, S. Zhou, X. Zhong, Z. Niu, X. Zhou, and H. Zhang, "Spatial Modeling of the Traffic Density in Cellular Networks," *IEEE Wireless Comm.*, pp. 80–88, Feb 2014.

[66] D. G. González and J. Hämäläinen, "Topology and Irregularity in Cellular Networks," in *2015 IEEE Wireless Comm. and Networking Conference (WCNC), New Orleans, LA*, pp. 1719–1724, Mar 2015.

[67] H. Kim, G. de Veciana, X. Yang, and M. Venkatachalam, "Alpha-Optimal User Association and Cell Load Balancing in Wireless Networks," in *Proceedings IEEE INFOCOM, San Diego, CA*, pp. 1–5, 2010.

[68] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A Tractable Approach to Coverage and Rate in Cellular Networks," *IEEE Trans. on Comm.*, vol. 59, pp. 3122–3134, Nov 2011.

[69] P. V. Bahl, M. T. Hajiaghayi, K. Jain, S. V. Mirrokni, L. Qiu, , and A. Saberi, "Cell Breathing in Wireless LANs: Algorithms and Evaluation," *IEEE Trans. on Mobile Comp.*, vol. 6, Feb 2007.

[70] L. Du, J. Bigham, and L. Cuthhen, "A Bubble Oscillation Algorithm for Distributed Geographic Load Balancing in Mobile Networks," in *IEEE INFOCOM 2004*, 2004.

[71] D. A. Bulti, D. H. Woldegebreal, G. D. González, B. B. Haile, and J. Hämäläinen, "User Association and Load Balancing in Long Term Evolution Network in Addis Ababa, Ethiopia," in *AFRICON*, pp. 1–5, Sep 2015.

[72] M. Huang, S. Feng, and J. Chen, "A Practical Approach for Load Balancing in LTE Networks," *Journal of Communications, Engineering and Technology Publishing*, vol. 9, pp. 490–497, June 2014.

[73] K. Son, S. Chong, and G. de Veciana, "Dynamic Association for Load Balancing and Interference Avoidance in Multi-Cell Networks," *IEEE Trans. on Wireless Comm.*, vol. 8, pp. 3566–3576, Jul 2009.

[74] 3GPP, "TS 36.211, Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation," Tech. Rep. v8.4.0, 3GPP Technical Specification, Sep 2008.

[75] M. S. Ali, P. Coucheney, and M. Coupechoux, "Learning Annealing Schedule of Log-Linear Algorithms for Load Balancing in HetNets," in *European Wireless 2016; 22th European Wireless Conference*, pp. 1–6, 2016.

[76] Q. Kuang, "Joint User Association and Reuse Pattern Selection in Heterogeneous Networks," in *11th International Symposium on Wireless Comm. Systems (ISWCS), Barcelona*, pp. 401–405, 2014.

[77] 3GPP, "3GPP TS 36.300 version 10.10.0 Release 10: Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2," tech. rep., 3rd Generation Partnership Project (3GPP), Jul 2013.

[78] I. Siomina and D. Yuan, "Load Balancing in Heterogeneous LTE: Range Optimization via Cell Offset and Load-Coupling Characterization," in *IEEE ICC 2012 - Comm. QoS, Reliability and Modeling Symposium*, pp. 1357–1361, 2012.

[79] M. A. Gadam, M. A. Ahmed, C. K. Ng, N. K. Nordin, A. Sali, and F. Hashim, "Review of Adaptive Cell Selection Techniques in LTE-Advanced Heterogeneous Networks," *Journal of Computer Networks and Communications*, vol. 2016, May 2016.

[80] T. Kudo and T. Ohtsuki, "Cell range expansion using distributed Q-learning in heterogeneous networks," *EURASIP Journal on Wireless Comm. and Networking*, vol. 2013, no. 1, 2013.

[81] M. Fenton, D. Lynch, S. Kucera, H. Claussen, and M. O'Neill, "Load Balancing in Heterogeneous Networks using an Evolutionary Algorithm," *IEEE*, 2015.

[82] J. Oh and Y. Han, "Cell Selection for Range Expansion with Almost Blank Subframe in Heterogeneous Networks," in *IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications - (PIMRC), Sydney, NSW*, pp. 653–657, Sep 2012.

[83] I. Guvenc, "Capacity and Fairness Analysis of Heterogeneous Networks with Range Expansion and Interference Coordination," *IEEE Comm. Letters*, vol. 15, pp. 1084 – 1087, Oct 2011.

[84] S. Deb, P. Monogioudis, J. Miernik, and J. P. Seymour, "Algorithms for Enhanced

Inter-Cell Interference Coordination (eICIC) in LTE HetNets," *IEEE/ACM Trans. on Networking*, vol. 22, pp. 137–150, Feb 2014.

[85] M. Y. Umair, D. Xiao, Y. Dongkai, and F. Fanny, "Identification of Interferers in Het-Net in LTE-A Systems Based on FeICIC with Cell Range Expansion," in *Information and Communication Technology (ICoICT), Bandung*, pp. 198–201, Mar 2013.

[86] W. Jin, J. Huilin, P. Zhiwen, L. Nan, Y. Xiaohu, and D. Tianle, "Joint User Association and ABS Proportion Optimization for Load Balancing in HetNet," in *2015 International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 1–6, 2015.

[87] Z. H. Abbas, F. Muhammad, and L. Jiao, "Analysis of Load Balancing and Interference Management in Heterogeneous Cellular Networks," *IEEE Access*, 2017.

[88] M. Haddad, P. Więcek, H. Sidi, and E. Altman, "An Automated Dynamic Offset for Network Selection in Heterogeneous Networks," *IEEE Trans. on Mobile Comp.*, vol. 15, no. 9, pp. 2151 – 2164, 2015.

[89] J. B. Abderrazak, A. Zemzem, and H. Besbes, "A Distributed Muting Adaptation Solution for a QoS-Aware User Association and Load Balancing in HetNets," in *ICTC 2015*, pp. 419–424, 2015.

[90] Z. Feng, Z. Feng, W. Li, and W. Chen, "Downlink and Uplink Splitting User Association inTwo-tier Heterogeneous Cellular Networks," in *Globecom 2014 - Wireless Networking Symposium*, pp. 4659–4664, 2014.

[91] F. Boccardi, J. Andrews, H. Elshaer, M. Dohler, S. Parkvall, P. Popovski, and S. Singh, "Why to Decouple the Uplink and Downlink in Cellular Networks and How To Do It," *IEEE Comm. Mag.*, 2016.

[92] F. Muhammad, Z. H. Abbas, G. Abbas, and L. Jiao, "Decoupled Downlink-Uplink Coverage Analysis with Interference Management for Enriched Heterogeneous Cellular Networks," *IEEE Access*, vol. 4, pp. 6250–6260, Oct. 2016.

[93] M. A. Lema, EnricPardo, O. Galinina, S. Andreev, and M. Dohler, "Flexible Dual-Connectivity Spectrum Aggregation for Decoupled Uplink and Downlink Access in 5G Heterogeneous Systems," *IEEE Journal on Selected Areas in Comm.*, vol. 34, no. 11, pp. 2851 – 2865, 2016.

[94] H. Wang, M. Garcia-Lozano, E. Mutafungwa, X. Yin, and S. Ruiz, "Performance Study of Uplink and Downlink Splitting in Ultradense Highly Loaded Networks," *Wireless Communications and Mobile Computing, Hindawi*, vol. 2018, jul 2018.

[95] A. Ebrahim and E. Alsusa, "Adaptive De-Coupling and Multi-BS Association in Heterogeneous Networks," *IEEE Access*, vol. 5, pp. 18121 – 18131, aug 2017.

[96] M. Kim, S. Y. Jung, and S.-L. Kim, "Sum-Rate Maximizing Cell Association via Dual-Connectivity," in *2015 International Conference on Conputer, Information and Telecommunication Systems (CITS)*, pp. 1–5, 2015.

[97] 3GPP, "Study on Small Cell Enhancements for E–UTRA and E-UTRAN–Higher layer aspects," tech. rep., 3GPP TR 36.842., 2013.

[98] 3GPP, "Scenarios and Requirements for Small Cell Enhancements," tech. rep., 3GPP TR 36.932., 2013.

[99] 3GPP, "3GPP work items on Small Cells v0.0.4," tech. rep., 3GPP, 2014.

[100] Z. Lei and C. Gang, "QoS-Aware User Association for Load Balancing in Heterogeneous Cellular Network with Dual Connectivity," in *2nd IEEE International Conference on Computer and Comm.*, pp. 2878–2883, 2016.

[101] J. Liu, M. Sheng, L. Liu, and J. Li, "Network Densification in 5G: From the Short-Range Communications Perspective," *IEEE Comm. Magazine*, pp. 96–102, December 2017.

[102] F. Wang, W. Chen, H. Tang, and Q. Wu, "Joint Optimization of User Association, Subchannel Allocation, and Power Allocation in Multi-cell Multi-association OFDMA Heterogeneous Networks," *IEEE Trans. on Comm.*, vol. 65, no. 6, pp. 2672–2684, 2017.

[103] U. Paul, A. P. Subramanian, M. M. Buddhikot, and S. R.Das, "Understanding Traffic Dynamics in Cellular Data Networks," in *IEEE INFOCOM 2011*, pp. 882–890, 2011.

[104] B. Błaszczyszyn, M. Jovanovicy, and M. K. Karray, "How User Throughput Depends on the Traffic Demand in Large Cellular Networks," in *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2014 12th International Symposium on, Hammamet*, pp. 611–619, 2014.

[105] A. S. Hamza, S. S. Khalifa, H. S. Hamza, and K. Elsayed, "A Survey on Inter-Cell Interference Coordination Techniques in OFDMA-Based Cellular Networks," *IEEE Comm. Sur. & Tut.*, vol. 15, no. 4, pp. 1642–1670, 2013.

[106] A. Tall, Z. Altman, and E. Altman, "Self Organizing strategies for enhanced ICIC (eICIC)," *available at arXiv:1401.2369v1 [cs.NI] 10 Jan 2014*, 2014.

[107] K. I. Pedersen, Y. Wang, S. Strzyz, and F. Frederiksen, "Enhanced Inter-cell Interference Coordination in Co-channel Multi-layer LTE-Advanced Networks," *IEEE Wireless Comm.*, pp. 120–127, Jun 2013.

[108] L. Chen, L. Ma, Y. Xu, V. C. Leung, and X. Wang, "Cluster-based Joint Cell Association and Interference Coordination Control in Heterogeneous Networks," in *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*, 2016.

[109] W. Qi, B. Zhang, B. Chen, and J. Zhang, "A User-Based K-means Clustering Offloading Algorithm for Heterogeneous Network ," in *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, 2018.

[110] L. Liu, Y. Zhou, V. Garcia, L. Tian, and J. Shi, "Load Aware Joint CoMP Clustering and Inter-cell Resource Scheduling in Heterogeneous Ultra Dense Cellular Networks," *IEEE Trans. on Vehicular Technology*, vol. 67, pp. 2741–2755, Mar. 2018.

[111] G. Sun, H. Zhang, and G. Liu, "User Demand Aware Soft-Association Control in Ultra-Dense Small Cell Networks," in *2016 IEEE 41st Conf. on Local Computer Networks*, pp. 651–654, 2016.

[112] M. Awais, A. Ahmed, M. Naeem, M. Iqbal, W. Ejaz, A. Anplagan, and H. S. Kim, "Efficient Joint User Association and Resource Allocation for Cloud Radio Access Networks," *IEEE Access*, vol. 5, pp. 1439–1448, 2017.

[113] Y. Xu and S. Mao, "User Association in Massive MIMO HetNets," *IEEE Systems Journal*, vol. 11, pp. 7–19, Mar. 2017.

[114] 5G PPP, "View on 5G Architecture," White Paper Version 3.0, 5G PPP Architecture Working Group, Feb 2020. Accessed on May 11, 2022.

[115] P. Luong, T. M. Nguyen, L. B. Le, N.-D. Dáo, and E. Hossain, "Energy-Efficient WiFi Offloading and Network Management in Heterogeneous Wireless Networks," *IEEE Access*, vol. 4, pp. 10210–10227, Jan 2016.

[116] M. Jaber, M. A. Imran, R. Tafazolli, and A. Tukmanov, "5G Backhaul Challenges and Emerging Research Directions: A Survey," *IEEE Access*, vol. 4, pp. 1743–1766, 2016.

[117] T. Han and N. Ansari, "User Association in Backhaul Constrained Small Cell Networks," in *2015 IEEE Wireless Comm. and Networking Conf. (WCNC)*, pp. 1637–1642, 2015.

[118] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H.-P. Mayer, L. Thiele, and V. Jungnickel, "Coordinated Multipoint: Concepts, Performance, and Field Trial Results," *IEEE Comm. Magazine*, vol. 49, pp. 102–111, Feb 2011.

[119] L. Liu, Y. Zhou, A. V. Vasilakos, L. Tian, and J. Shi, "Time-domain ICIC and optimized designs for 5G and beyond: a survey," *Science China Information Sciences*, vol. 62, pp. 1–28, Feb. 2019.

[120] N. Trabelsi, C. Chen, R. El-Azouzi, L. Roullet, and E. Altman, "User Association and Resource Allocation Optimization in LTE Cellular Networks," *IEEE Tran. on Network and Service Management, IEEE*, vol. 14 (2), pp. 429–440, 2017.

[121] L. Liu, Y. Zhou, J. Yuan, W. Zhuang, and Y. Wang, "Economically Optimal MS Association for Multimedia Content Delivery in Cache-Enabled Heterogeneous

Cloud Radio Access Networks," *IEEE Journal on Selected Areas in Comm.*, vol. 37, pp. 1584–1593, Jul. 2019.

[122] X. Yang, "A Multi-level Soft Frequency Reuse Technique for Wireless Communication Systems," *IEEE Comm. Let.*, vol. 18, pp. 1983 – 1986, oct 2014.

[123] 3GPP TS 36.300 version 8.12.0 Release 8, "LTE; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2 ," Apr 2010.

[124] H. Holma and A. Toskala, *LTE for UMTS: OFDMA and SC-FDMA Based Radio Access*. John Wiley & Sons, Ltd., 2009.

[125] 3GPP, "3GPP TR 36.814 V9.2.0; TS Group Radio Access Network; E-UTRA; Further advancements for E-UTRA physical layer aspects," tech. rep., 3GPP, 2017.

[126] Y. Shi, E. Alsusa, and M. W. Baidas, "A survey on downlink – uplink decoupled access: Advances, challenges,and open problems," *Computer Networks*, vol. 213, Aug. 2022.

[127] K. Hou, Q. Xu, X. Zhang, Y. Huang, and L. Yang, "User Association and Power Allocation Based on Unsupervised Graph Model in Ultra-Dense Network," in *IEEE Wireless Comm. and Networking Conference (WCNC)*, 2021.

[128] K. Smiljkovikj, P. Popovski, and L. Gavrilovska, "Analysis of the Decoupled Access for Downlink and Uplink in Wireless Heterogeneous Networks," *IEEE Wireless Comm. Letters*, vol. 4, pp. 173 – 176, apr 2015.

[129] M. N. Sial and J. Ahmed, "A Realistic Uplink–Downlink Coupled and Decoupled User Association Technique forK-tier 5G HetNets," *Arabian Journal for Science and Engineering*, Jun. 2018.

[130] M. Saimler and S. Coleri, "Multi-Connectivity based Uplink/Downlink decoupled Energy Efficient User Association in 5G Heterogenous CRAN," *IEEE Comm. Letters*, 2020.

[131] Y. Shi, *5G and Beyond Wireless Networkd Optimization Through Uplink and Downlink Decoupled Access*. PhD thesis, School of Electrical and Electronic Engineering, University of Manchester, 2021.

[132] Z. Feng, Z. Feng, and T. A. Gulliver, "Joint user association and resource partition for downlink-uplink decoupling in multi-tier HetNets," *Frontiers of Information Technology & Electronic Engineering*, vol. 18, pp. 817–829, Jun 2017.

[133] E. Pardo, "Study of Decoupled Uplink and Downlink Access In 5G Heterogeus systems," Master's thesis, King's College London and Universitat politécnica de Catalunya (UPC), May 2016.

[134] M. N. Sial and J. Ahmed, "Analysis of K-tier 5G heterogeneous cellular network with dual-connectivity and uplink–downlink decoupled access," *Telecommun Syst*, Aug. 2017.

[135] O. W. Bhatti, H. Suhail, U. Akbar, S. A. Hassan, H. Pervaiz, L. Musavian, and Q. Ni, "Performance Analysis of Decoupled Cell Association in Multi-Tier Hybrid Networks using Real Blockage Environments," *arXiv:1705.04390v1 [cs.IT]*, may 2017.

[136] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic Geometry and Random Graphs for the Analysis and Design of Wireless Networks," *IEEE Journal on Selected Areas in Comm.*, vol. 27, pp. 1029 – 1046, Sept. 2009.

[137] T. S. Rappaport, J. George R. MacCartney, M. K. Samimi, and S. Sun, "Wideband Millimeter-Wave Propagation Measurements and Channel Models for Future Wireless Communication System Design," *IEEE Trans. on Comm.*, vol. 63, pp. 3029–3056, Sep 2015.

[138] H. Zhang and W. Huang, "Tractable Mobility Model for Multi Connectivity in 5G User-Centric Ultra-Dense Networks," *IEEE Access*, vol. 6, pp. 43100–43112, 2018.

[139] T. H. L. Dinh, M. Kaneko, K. Wakao, K. Kawamuran, T. Moriyama, H. Abeysekera, and Y. Takatori, "Distributed user-to-multiple access points association through deep learning for beyond 5G," *Computer Networks*, vol. 197, no. 108258, 2021.

[140] T. Sahin, M. Klügel, C. Zhou, and W. Kellerer, "Virtual Cells for 5G V2X Communications," *IEEE Comm. Standards Magazine*, vol. 2, no. 1, pp. 22–28, 2018.

[141] H. Wang, C. Rosa, and K. I. Pedersen, "Dual connectivity for LTE-advanced heterogeneous networks," *Wireless Networks*, vol. 22, pp. 1315–1328, 2016.

[142] "3GPP TR 36.932 V 12.1.0, Technical Specification Group Radio Access Network; Scenarios and requirements for small cell enhancements for E-UTRA and E-UTRAN, Release 12," Mar. 2013.

[143] R. P. Antonioli, G. C. Parente, C. F. M. e Silva, D. A. Sousa, E. B. Rodrigues, T. F. Maciel, and F. R. P. Cavalcanti, "Dual Connectivity for LTE-NR Cellular Networks: Challenges and Open Issues," *Journal of Communication and Information Systems*, vol. 33, pp. 282–294, aug 2018.

[144] A. Zakrzewska, D. López-Pérez, S. Kucera, and H. Claussen, "Dual Connectivity in LTE HetNets with Split Control- and User-Plane," in *Globecom 2013 Workshop - Broadband Wireless Access*, 2013.

[145] P. Popineau, S. Kalamkar, and F. Baccelli, "On Velocity-based Association Policies for Multi-tier 5G Wireless Networks," in *2021 IEEE GLOBECOM*, pp. 1–6, 2021.

[146] B. Soleimani and M. Sabbaghian, "Cluster-based Resource Allocation and User Association in mmWave Femtocell Networks," *IEEE Trans. on Comm.*, vol. 68, no. 3, pp. 1746–1759, 2018.

[147] C. Pan, M. Elkashlan, J. Wang, J. Yuan, and L. Hanzo, "User-centric C-RAN Architecture for Ultra-dense 5G Networks: Challenges and Methodologies," *IEEE Comm. Magazine*, vol. 56, pp. 14–20, June 2018.

[148] S. Choi, J.-G. Choi, and S. Bahk, "Mobility-Aware Analysis of Millimeter Wave Communication Systems With Blockages," *IEEE Trans. on Vehicular Technology*, vol. 69, no. 6, pp. 5901–5912, 2020.

[149] A. Jain, *Enhanced Mobility Management Mechanisms for 5G Networks*. phd thesis, Department of Network Engineering, Technical University of Catalunya, Barcelona, Spain, 2020.

[150] S. Chen, F. Qin, B. hu, X. Li, and zhonglin Chen, "User-centric ultra-dense networks for 5G: Challenges, methodologies, and directions," *IEEE Wireless Comm.*, vol. 23, pp. 78 – 85, Apr. 2016.

[151] I. A. Alablani and M. A. Arafah, "An Adaptive Cell Selection Scheme for 5G Heterogeneous Ultra-Dense Networks," *IEEE Access*, vol. 9, pp. 64224 – 64240, Apr. 2021.

[152] A. Alhammadi, M. Roslee, M. Y. Alias, I. Shayea, and A. Alquhali, "Velocity-Aware Handover Self-Optimization Management for Next Generation Networks," *MDPI, Applied Sciences*, vol. 10, pp. 1–14, Feb. 2020.

[153] J. Park, S. Y. Jung, S.-L. Kim, M. Bennis, and M. Debbah, "User-Centric Mobility Management in Ultra-Dense Cellular Networks under Spatio-Temporal Dynamics," in *2016 IEEE GLOBECOM*, pp. 1–6, 2016.

[154] 3GPP, "3GPP TR 38.901 version 16.1.0 Release 16: 5G;Study on channel model for frequencies from 0.5 to 100 GHz," tech. rep., 3GPP, 2020. Accessed on Nov. 15, 2022.

[155] D. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods*, vol. 1 of *A Series of the Applied Probability Trust*. Springer, second edition ed., 2003.

Appendix A

Appendix

The derivations for the ergodic rates are presented below.

A.1 Proof of Lemma 4.6.1 – UL Ergodic Rates

When a typical UE is associated to the MC in the UL, the ergodic rate is given by:

$$\begin{aligned} R_{UL}^{m} &= \frac{1}{\ln(2)}\mathbf{E}_{r,\psi}[\ln(1+\psi_{UL}^{m})] \\ &= \frac{1}{\ln(2)}\int_{0}^{\infty}\mathbf{E}_{\psi}[\ln(1+\frac{P_u h_m r^{-\gamma_m}}{I})]\cdot f(r,1)dr, \end{aligned} \tag{A.1}$$

where, $I = \sum_{k\in\Phi_u\setminus u} P_u g_k x^{-\gamma_k}$ is the interference from users except the typical user at the origin and $f(r,1)dr$ is the distance distribution of the serving node given in (4.1). The expectation of the spectral efficiency term in right-hand side (RHS) of (A.1) can be obtained as in [68].

$$\begin{aligned} R_{UL}^{m*} = \mathbf{E}_{\psi}[\ln(1+\frac{P_u h_m r^{-\gamma_m}}{I})] &= \int_{0}^{\infty}\mathbf{P}\{\ln(1+\frac{P_u h_m r^{-\gamma_m}}{I}) > y\}dy \\ &= \int_{0}^{\infty}\mathbf{P}\{h_m > I P_u^{-1} r^{\gamma_m}(e^{y}-1)\}dy \\ &\overset{a}{=} \int_{0}^{\infty}\exp\{-\mu I P_u^{-1} r^{\gamma_m}(e^{y}-1)\}dy \\ &= \int_{0}^{\infty}\mathcal{L}_I\{\mu P_u^{-1} r^{\gamma_m}(e^{y}-1)\}dy \end{aligned} \tag{A.2}$$

Where (a) follows from the exponentially distributed h_m with mean $1/\mu$ and the Laplace Transform (LT) of the interference can be expressed as:

$$\begin{aligned}\mathcal{L}_I(s) = \mathbf{E}_{\Phi_u,g_k}[e^{-sI}] &= \mathbf{E}_{\Phi_u,g_k}[\exp\{-s\sum_{k\in\Phi_u\setminus u} P_u g_k x^{-\gamma_k}\}] \\ &= \mathbf{E}_{\Phi_u,g_k}[\prod_{k\in\Phi_u\setminus u} \exp\{-sP_u g_k x^{-\gamma_k}\}] \\ &\overset{a}{=} \mathbf{E}_{\Phi_u}[\prod_{k\in\Phi_u\setminus u} \mathbf{E}_{g_k}[\exp\{-sP_u h_k x^{-\gamma_k}\}]]\end{aligned} \tag{A.3}$$

(a) follows from the independence between Φ_u and g_k. With help of Probability Generating Functional (PGFL) [36] and [155] of the PPP, which states for some function $f(x)$ that $\mathbf{E}[\prod_{x\in\Phi} f(x)] = \exp\{-\lambda\int_{R^2}(1-f(x))dx\}$, the equation in (A.3) becomes:

$$\begin{aligned}\mathcal{L}_I(s) = \mathbf{E}_{\Phi_u,g_k}[e^{-sI}] &= \exp\{-2\pi\lambda_u\int_r^\infty (1-\mathbf{E}_{g_k}[\exp\{-sP_u g_k x^{-\gamma_k}\}])xdx\} \\ &\overset{a}{=} \exp\{-2\pi\lambda_u^*\int_r^\infty (1-\frac{\mu}{sP_u x^{-\gamma_k}+\mu})xdx\}\end{aligned} \tag{A.4}$$

Where (a) follows from exponential distribution of g_k. Substituting $s = \mu P_u^{-1} r^{\gamma_m}(e^y - 1)$ and putting (A.4) in (A.2) and (A.1) with simplification gives the result.

The same procedure can be followed to obtain the ergodic user rate when a typical UE is associated to the LPN in the UL (the ergodic rate is given in (4.19)).

A.2 Proof of Lemma 4.6.2 – DL Ergodic Rates

When a typical UE is associated to the MC in the DL, the ergodic rate is given by:

$$\begin{aligned}R_{DL}^m &= \frac{1}{\ln(2)}\mathbf{E}_{r,\psi}[\ln(1+\psi_{DL}^m)] \\ &= \frac{1}{\ln(2)}\int_0^\infty \mathbf{E}_\psi[\ln(1+\frac{P_m h_m r^{-\gamma_m}}{I})]\cdot f(r,1)dr,\end{aligned} \tag{A.5}$$

where, $I = \sum_{k\in\Phi_m\setminus m} P_m g_k r^{-\gamma_k} + \sum_{k\in\Phi_l} P_l g_k r^{-\gamma_k}$ is the interference from MCs and LPNs to a typical user at the origin which being served by MC m and $f(r,1)dr$ is the distance distribution of the serving node. The expectation of the spectral efficiency term in RHS

of (A.5) can be obtained as follows.

$$\begin{aligned} R_{DL}^{m*} = \mathbf{E}_{\psi}[\ln(1+\frac{P_m h_m r^{-\gamma_m}}{I})] &= \int_0^{\infty} \mathbf{P}\{\ln(1+\frac{P_m h_m r^{-\gamma_m}}{I}) > y\}dy \\ &= \int_0^{\infty} \mathbf{P}\{h_m > I P_m^{-1} r^{\gamma_m}(e^y-1)\}dy \\ &\overset{a}{=} \int_0^{\infty} \exp\{-\mu I P_m^{-1} r^{\gamma_m}(e^y-1)\}dy \\ &= \int_0^{\infty} \mathcal{L}_I\{\mu P_m^{-1} r^{\gamma_m}(e^y-1)\}dy, \end{aligned} \tag{A.6}$$

where (a) follows from the exponentially distributed h_m with mean $1/\mu$. The LT of the interference can be expressed as:

$$\begin{aligned} \mathcal{L}_I(s) = \mathbf{E}_{\Phi_m,\Phi_l,g_k}[e^{-sI}] &= \mathbf{E}_{\Phi_m,\Phi_l,g_k}[\exp\{-s(\sum_{k\in\Phi_m\setminus m} P_m g_k r^{-\gamma_k} + \sum_{k\in\Phi_l} P_l g_k r^{-\gamma_k})\}] \\ &= \mathbf{E}_{\Phi_m,\Phi_l,g_k}[\prod_{k\in\Phi_m\setminus m} \exp\{-sP_m g_k r^{-\gamma_k}\} \times \prod_{k\in\Phi_l} \exp\{-sP_l g_k r^{-\gamma_k}\}] \\ &\overset{a}{=} \mathbf{E}_{\Phi_m}[\prod_{k\in\Phi_m\setminus m} \mathbf{E}_{g_k}[\exp\{-sP_m g_k r^{-\gamma_k}\}]] \times \mathbf{E}_{\Phi_l}[\prod_{k\in\Phi_l} \mathbf{E}_{g_k}[\exp\{-sP_l g_k r^{-\gamma_k}\}]] \end{aligned} \tag{A.7}$$

(a) follows from the independence between Φ_u, Φ_l and h_k. With help of PGFL [36] and [155] of the PPP, which states for some function $f(x)$ that $\mathbf{E}[\prod_{x\in\Phi} f(x)] = \exp\{-\lambda \int_{R^2}(1-f(x))dx\}$, and considering exponential distribution of g_k, equation in (A.7) becomes:

$$\begin{aligned} \mathcal{L}_I(s) = \mathbf{E}_{\Phi_m,\Phi_l,h_k}[e^{-sI}] &= \exp\{-2\pi\lambda_m \int_r^{\infty} (1-\frac{\mu}{sP_m x^{-\gamma_k}+\mu})xdx\} \\ &\times \exp\{-2\pi\lambda_l \int_r^{\infty} (1-\frac{\mu}{sP_l z^{-\gamma_k}+\mu})zdz\} \end{aligned} \tag{A.8}$$

Substituting $s = \mu P_m^{-1} r^{\gamma_m}(e^y-1)$ and putting (A.8) in (A.6) and (A.5) with simplification gives the result.

Similarly, the same procedure can be followed to obtain the ergodic user rate when a typical UE is associated to the LPN in the DL (the ergodic rate is given in (4.20)).