USER ASSOCIATION AND ROUTING IN UAV-SUPPORTED HETNETS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

BERKE TEZERGIL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
COMPUTER ENGINEERING

JANUARY 2022

Approval of the thesis:

USER ASSOCIATION AND ROUTING IN UAV-SUPPORTED HETNETS

submitted by **BERKE TEZERGIL** in partial fulfillment of the requirements for the degree of **Master of Science in Computer Engineering Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. Halit Oğuztüzün
Head of Department, **Computer Engineering** _____

Prof. Dr. Ertan Onur
Supervisor, **Computer Engineering, METU** _____

Examining Committee Members:

Prof. Dr. İbrahim Körpeoğlu
Computer Engineering, Bilkent University _____

Prof. Dr. Ertan Onur
Computer Engineering, METU _____

Assist. Prof. Dr. Pelin Angın
Computer Engineering, METU _____

Date:24.01.2022

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Berke Tezergil

Signature :

ABSTRACT

USER ASSOCIATION AND ROUTING IN UAV-SUPPORTED HETNETS

Tezergil, Berke
M.S., Department of Computer Engineering
Supervisor: Prof. Dr. Ertan Onur

January 2022, 116 pages

With the introduction of millimeter waves in 5G, using wireless backhaul has become feasible with higher performance comparable to that of fiber cables. Using unmanned aerial vehicles as small cells enabled many use-cases by reducing the average link length and increasing the line-of-sight probability. In this work, a heterogeneous network with users, flying small cells on unmanned aerial vehicles, and macro base stations are considered. We introduce two main problems: establishing backhaul routes for small cells to maximize data capacity for users, and associating every user with a base station.

The problem is named as the UAV-UAR problem. Initially, a mixed-integer linear programming formulation is given, which is optimal, but requires considerable time to find a solution. Using this formulation, a flow network definition is given for the heterogeneous network, and used to formulate relabel-to-front algorithm-based heuristics. While these heuristic methods do not guarantee optimality, they are significantly faster than the exact solution.

The first developed heuristic, Relabel-to-Front-Eliminate, eliminates all edges that users allocate capacity except one. The second heuristic, Relabel-to-Front-Heuristic,

uses a heuristic preflow initialization to associate users before execution. The final heuristic, Relabel-to-Front-Iterative, uses the second method, but iteratively changes user association until the result no longer improves. Monte Carlo Simulation results show that relabel-to-front-based heuristics have comparable, and usually the same throughput performance to that of linear programming optimization, but with a sliver of the execution time, outperforming by 20 to 1000 times depending on other parameters.

Keywords: 5G and Beyond Mobile Networks, Heterogeneous Networks, User Association, Routing, mmWave, Flow Networks

ÖZ

İHA DESTEKLİ HETEROJEN AĞLARDA KULLANICI İLİŞKİLENDİRMESİ VE YÖNLENDİRME

Tezergil, Berke
Yüksek Lisans, Bilgisayar Mühendisliği Bölümü
Tez Yöneticisi: Prof. Dr. Ertan Onur

Ocak 2022 , 116 sayfa

5G ile birlikte milimetre dalga frekanslarının kullanıma alınmasıyla, kablosuz geri taşıma kullanımı fiber kablolara yakın performans verebilmeye başlamıştır. Böylece insansız hava araçlarını küçük hücreler olarak kullanmanın önü açılmıştır. Küçük hücreler, bağlantıların uzunluğunu kısaltarak ve görüş olasılığını artırarak ağ performansını artırırlar. Bu tezde, kullanıcıların, insansız hava aracı üstünde küçük hücrelerin, ve makro baz istasyonlarının oluşturduğu bir heterojen ağı ele alıyoruz. Bu ağın iki ana problemi İHA küçük hücreler için geri taşıma yolları oluşturup kullanıcılara ayrılan ağ kapasitesini artırmak, ve kullanıcıları baz istasyonlarıyla eşleştirmektir.

Bu probleme UAV-UAR problemi adını veriyoruz. İlk olarak, karışık tamsayılı doğrusal programlama kullanarak yapılan formülasyon optimal çözümü bulsa da, çalışma zamanı özellikle gerçek zamanlı kullanım için oldukca uzun kalmaktadır. Bu sorunu çözmek için, heterojen ağ bir akış çizgesine çevrilmiştir. Bu formülasyon sayesinde Relabel-to-Front algoritması tabanlı buluşsal algoritmalar geliştirerek en iyi çözümü bulamasak da çözüm hızını ciddi biçimde artırıyoruz.

İlk buluşsal metod Relabel-to-Front-Eleme, kullanıcıya atanan kenarları seçilen bir tanesi harici eleyerek doğru sonucu bulmaktadır. İkinci metod Relabel-to-Front-Buluşsal Önakış, relabel-to-front algoritmasının başında kullanılan önakış atamasını buluşsal bir şekilde yaparak kullanıcı atamasını baştan yapmaktadır. Son metod, Relabel-to-Front-Yineleme ise ikinci metodu kullanmakta ama sonuç iyileşmeyene kadar alternatifleri denemektedir. Monte Carlo simülasyonlarının sonucu bize relabel-to-front-tabanlı buluşsal metodların çalışma zamanını 20 ile 1000 kat arası hızlandırarak, doğrusal programlama optimizasyonuna yakın ağ verimi sağladığını göstermiştir.

Anahtar Kelimeler: 5G ve Ötesi Mobil Ağlar, Heterojen Ağlar, Kullanıcı İlişkilendirme, Yönlendirme, Milimetre Dalga, Akış Ağları

To my family and friends

ACKNOWLEDGMENTS

First and foremost, I'd like to thank my advisor Prof. Dr. Ertan Onur for his helpful comments and feedbacks throughout the writing process. During the publication process of our paper as well as this thesis, his guidance has been invaluable to me. He always managed to point me to the right direction, significantly improving the quality of work I produced.

I'd like to thank my father Mustafa and my mother Filiz for believing in me and helping me through the graduate study process. When the research processes were arduous, they have been there to support me and encourage me to move forward. I cannot ask for a more supportive pair of parents, and this thesis would not have been possible without them believing in me and guiding me in my 21 years of endless studying.

I'd also like to mention Prof. Dr. Hakkı Toroslu, since his help has been paramount to the success of the heuristic methods that I developed in this thesis. He has been always available to talk, with concise feedback and guidance. He has been nothing but encouraging and supportive to me.

I would like to thank Technology and Communication Authority and Turkcell for their support throughout my Master's studies. As a holder of 5G and Beyond Scholarship, I feel privileged to have worked with them in this area and this work would not be possible without their support.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

2D	2 Dimensional
2G	Second Generation
3D	3 Dimensional
3G	Third Generation
3GPP	Third Generation Partnership Project
5G	Fifth Generation
6G	Sixth Generation
AP	Access Point
ATM	Asynchronous Transfer Mode
AWGN	Additive White Gaussian Noise
BER	Bit-Error Rate
BS	Base Station
BSC	Base Station Controller
BTS	Base Transceiver Station
CapEx	Capital Expenditure
CoMP	Coordinated Multi-Point
CP	Content Provider
CSI	Channel State Information
CU	Centralized Unit
D2D	Device to Device
DSL	Digital Subscriber Line
DU	Distributed Unit
eMBB	Enhanced Mobile Broadband
eNB	Evolved Node-B

FDD	Frequency Division Duplexing
FSO	Free Space Optics
GEO	Geostationary Orbit
gNB	Next Generation Node-B
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
GSMA	GSM Association
HAPS	High Altitude Platform Station
HetNet	Heterogeneous Network
H-CRAN	Heterogeneous Cloud Radio Access Network
IAB	Integrated Access & Backhaul
IBFD	In-band Full-Duplex
IRS	Intelligent Reflecting Surfaces
KPI	Key Performance Indicator
LEO	Low Earth Orbit
LoS	Line of Sight
LTE	Long Term Evolution
MBS	Macro Base Station
MEC	Mobile Edge Computing
MEO	Medium Earth Orbit
MIDO	Multiple-Input Distributed-Output
MIMO	Multiple-Input Multiple-Output
MINLP	Mixed-Integer Nonlinear Programming
mMIMO	Massive Multiple-Input Multiple-Output
mMTC	Massive Machine Type Communication
mmWave	Millimeter Wave
MNO	Mobile Network Operator

MSC	Mobile Switching Center
NFV	Network Function Virtualization
NLoS	Non-Line of Sight
NOMA	Non-Orthogonal Multiple Access
NR	New Radio
OAB	Orthogonal Access & Backhaul
OBFD	Out-of-band Full-Duplex
OFDM	Orthogonal Frequency Division Multiplexing
OpEx	Operating Expenses
PON	Passive Optical Networks
PtMP	Point-to-Multipoint
PtP	Point-to-Point
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RA	Resource Allocation
RMa	Rural Macro
ROI	Return of Investments
RRM	Radio Resource Management
SAGIN	Space-Air-Ground Integrated Network
SBS	Small Base Station
SDN	Software Defined Network
SINR	Signal-to-Interference+Noise Ratio
TDD	Time Division Duplexing
TDM	Time Division Multiplexing
TDMA	Time Division Multiple Access
UA	User Association
UAV	Unmanned Aerial Vehicle

UE	User Equipment
UMa	Urban Macro
UMTS	Universal Mobile Telecommunications System
URLLC	Ultra Reliable Low Latency Communication
VNF	Virtual Network Function
VSAT	Very Small Aperture Terminal

CHAPTER 1

INTRODUCTION

With the introduction of 5G, mobile networks are evolving in a rapid pace. The introduction of new technologies such as millimeter wave (mmWave) frequencies and Integrated Access & Backhaul (IAB) paves the way for further developments while also making a case for the concepts that did not see wide use in the previous generations. Wireless backhauling and network densification are examples of such concepts. Improvements in wireless backhauling gave rise to the small cell architectures, which significantly aided in increasing user density in networks.

Small cells can be used in a myriad of use-cases, such as by increasing cell-edge performance, coverage, or acting as a base station (BS) without lengthy and expensive deployment procedures. These can also be deployed to various locations such as lamp posts, vehicles, or on top of aerial platforms. While the concepts of using aerial platforms and satellites has been around for a long time, the performance improvements in wireless backhauling make it feasible to deploy unmanned aerial vehicles (UAVs) or high altitude platform stations (HAPSs) with similar performance to that of fiber-backhauled macro base stations (MBS).

Using UAVs and HAPSs in cellular networks as relays or base stations significantly improve network flexibility and allow the network a greater degree of adaptability, albeit with its own problems. For instance, managing UAV trajectories with regard to energy consumption, recharging, and service performance is a prominent problem of UAVs as small cells. Furthermore, while using mobile UAVs as small cells improve the degree of freedom for deployment, the deployment problem with UAVs is three dimensional whereas conventional networks mostly do not consider the third dimension, the height of network elements.

Provided that UAVs are used as network elements and their logistical problems are solved, comes the network operation-related parts. After deploying the UAVs to their respective locations, the traffic has to be routed to the core network, preferably via MBSs with fiber connectivity. Establishing the links are often not the scope of the available works. Configuring the links after deployment in such a way to maximize the flow of traffic going to the core network can significantly increase the network performance. Moreover, the carrier frequency used for the communication, or the availability of multi-hop connections change the performance ceiling. Most of the available works perform on sub-6 GHz frequencies that have higher range and robustness at the cost of performance.

Combining small cell, UAV, mmWave, and IAB concepts results in a high-performance network that mitigates the shortcomings of individual concepts. For example, IAB network often suffer from interference as both access and backhaul traffic is sent using the same channel. To mitigate this, a network can employ mmWave links that are directive and noise-limited. While interference ceases to be a problem, network is now noise limited and has shorter range since mmWave inherently has high path loss. To solve this, small cell concepts are employed. Having a denser network consisting of small cells and MBSs significantly shortens the link distance, which somewhat reduces the negative effects of mmWave in terms of range and noise-limitedness. Using small cells require careful planning and deployment as there are few candidate locations that can support small cells with high performance. Employing UAVs for small cells increases the freedom of the network operator as they have mobility and do not require a fixed location, notwithstanding their own problems. Finally, as UAVs do not have wired backhaul, employing wireless backhaul is required. IAB comes to the rescue in that aspect as well.

Employing all these concepts together, a quite effective network can be designed. This network can operate on multiple use-cases thanks to the degree of freedom that small cell architecture generally, and UAVs specifically, support. On top of this network scheme, we define the problem of user association and routing in the next section.

1.1 Motivation and Problem Definition

Consider the scenario in Figure 1.1. In an urban setting, users have multiple alternatives available to choose a base station that serves them. However, there is no single parameter that can govern this choice. For example, if links are chosen by proximity, UE1 has to be served by UAV small cell (UAV-SC), but this link is non-line-of-sight (NLoS), which significantly degrades performance. Users can prioritize connecting to MBSs as they have fiber connectivity to the core network, but if UE2 in Figure 1.1 connects to the MBS, as the link is NLoS, the performance will not be as high as what can be attained from the UAV-SC.

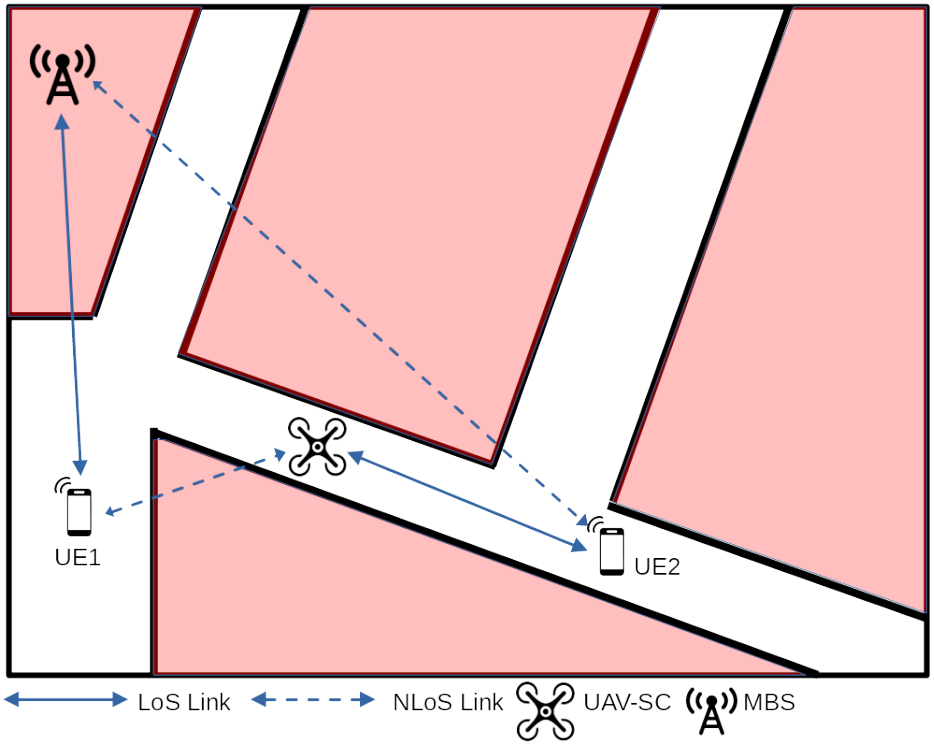


Figure 1.1: An example urban scenario consisting of two users, a UAV-SC, and a MBS.

User association is only one part of the problem. Suppose UE2 chooses the UAV-SC as its base station. Since UAV-SC does not have any wired connection like MBS, it has to employ wireless backhaul links to send any data to the core network. For this purpose, UAV-SCs can either directly use MBSs, or they can employ other UAV-SCs

for backhauling their traffic, using a multi-hop structure for backhauling. Routing of backhaul links is another problem that has to be considered for this network.

In this thesis, we consider an IAB Heterogeneous Network (HetNet) with UAVs as small cells. In this network, we consider the problem of user association and routing, which we name the UAV-UAR problem. To the best of our knowledge, user association and routing problem for UAV small cells is not jointly studied in the literature. Furthermore, available studies mostly use carrier frequencies of sub-6 GHz, which does not have the capability to satisfy the 5G data rate requirements. For this purpose, we employ 28 GHz that lies in FR2 frequency band, and 60 GHz V-band, which is a mmWave band that does not require licensing. To the best of our knowledge, there are no works on UAV small cells that operate on V-band frequencies.

We use 28-60 GHz mmWave frequencies due to them being able to deliver 5G data rates. While the mmWave frequencies have less range compared to sub-6 GHz frequencies due to higher free-space pathloss and susceptibility to atmospheric attenuation, they can also achieve much better throughput. Since we employ UAV-SCs to shorten the link distances, the effect of shorter range is somewhat mitigated. Furthermore, interference is not as destructive for mmWave frequencies as sub-6 GHz as they often employ directive links and beamforming. Directivity also aids in the network security as it is much harder to intercept the links without alerting the network. Finally, IAB networks are often interference-limited but due to the directivity of mmWave links, the network becomes noise-limited. Furthermore, the effect of noise on the network performance is also reduced as UAV-SCs are employed to shorten the links and improve signal strength.

Usually, the go-to solution method for optimization problems is integer programming methods. While these methods find the optimal solution, they take quite a long time to find the optimal solution and any system that uses such algorithms cannot be real-time responsive. For this purpose, the proposed algorithm makes use of a flow network-based modeling of the HetNet to provide real-time solution in seconds. With this speed, the proposed algorithm can be used in cellular networks for dynamic configurations.

Finally, since UAV small cells and their characteristics are handled during network

modeling, the proposed algorithm supports any type of user equipments (UEs), SBSs, or MBSs, or even multiple types of a network element in the same network. While this thesis finds the solution of the problem in IAB HetNets with UAV small cells, the proposed algorithm can also be employed for other networks that use different type of small cells, provided that the modeling is made accordingly.

1.2 Contributions

Our contributions are as follows:

- A survey on wireless backhauling in 5G is conducted prior to experiments. This survey focuses on the literature employing wireless backhauling on five main topics, namely resource allocation, deployment, scheduling, performance evaluation, coverage, and security. Furthermore, we also studied opportunities that wireless backhaul brought about, which are UAV/HAPS usage, mobile edge computing (MEC), rural connectivity, and satellite backhauling. Finally, we also briefly mention the technologies that might aid wireless backhauling beyond 5G. Sections directly related to this work can be found in Chapter 2.5

- An IAB HetNet model that employs UAVs as small cells is proposed. UEs connect to either the MBSs or the UAVs, depending on the capacity of the available links. UAVs employ multi-hop connections to bring the user data to any MBS, which have fiber connectivity to the core. The network uses V-band frequencies for communication, which is supported for 5G. This model is presented in Chapter 3.1

- The problem of user association and routing in HetNets with UAV-SC's (UAV-UAR) are formulated using the aforementioned network model. For this problem, a mathematical formulation is presented using constraints and the objective of maximizing the network throughput. Using this formulation as the basis, a mixed integer linear program (MILP) developed to solve the problem in Chapter 3.2.

- While the aforementioned MILP finds the optimal solution, its execution time

is not acceptable to be used in a dynamic system. Because of this shortcoming, an algorithm is proposed that has $O(V^3)$ complexity with V being the number of nodes in network, which is suitable for real-time usage in dynamic networks. Since the proposed HetNet model can be modeled as a flow network, an algorithm based on the relabel-to-front algorithm for flow maximization is proposed for UAV-UAR problem. The initial algorithm uses the relabel-to-front algorithm, and after execution eliminates all outgoing edges from UEs but one of them, conforming to the user association constraint. We call this algorithm Relabel-to-Front-and-Eliminate (RTF-E) and introduce this heuristic in Chapter 4.1.

- While the RTF-E algorithm produces a valid result, performing initial user association assignment instead of eliminating edges is a better approach in terms of choosing the serving BS. For this purpose, we add a heuristic preflow initialization, and use it with the default relabel-to-front method. We call this algorithm Relabel-to-Front-Heuristic (RTF-H) and define it in Chapter 4.2. On top of RTF-H, we also define another heuristic named Relabel-to-Front-Iterative (RTF-I) in Chapter 4.3, which iteratively tries alternative BSs that have higher upper bounds than the initial result.

1.3 The Outline of the Thesis

The thesis is organized as follows:

Chapter 2 gives background information regarding the cellular networks and related works to this topic. In the beginning, an overview of pre-5G networks are given, with emphasis on wireless backhauling. After this, concepts such as network densification and aerial network elements are mentioned. These topics are important to explain the rationale behind the choice of network model in this thesis. Furthermore, the concepts of flow networks and maximum flow problem are explained in detail as these topics provide the basis to proposed heuristics. Finally, a comprehensive list of related work is given in the end of this chapter.

Chapter 3 introduces the problem and its solution with an MILP. On top of the back-

ground given in the previous chapter, the UAV-UAR problem is introduced in detail, albeit not from a purely mathematical perspective, initially. The network structure and the rationale behind choices such as using wireless backhauling and UAV's as small cells are also explained. Moving on, a mathematical formulation is made for the UAV-UAR problem, which provides a MILP optimization to perform the basis for our comparisons.

Chapter 4 presents the novel algorithm solution for the same problem. In this chapter, the flow network formulation is introduced. Following this formulation, three heuristics are introduced. These heuristics use the relabel-to-front algorithm as their basis, but they implement further operations to satisfy the user association constraint as maximum flow problem allocates all available flows from source nodes, which are users in our network model.

In Chapter 5, experiment setup is detailed, and the results are discussed. Initially, the simulation parameters and scenarios are explained, alongside the method of simulation. Following this, topology generation process is explained in detail. After the simulation related information, we present the results in five categories, namely the comparison between RTF-based and MILP methods, effect of user demand, transmit power, UAV altitude, and user association. The chapter is concluded with a discussion regarding simulation results.

Chapter 6 concludes the thesis. Final remarks and future work are briefly mentioned.

CHAPTER 2

BACKGROUND AND RELATED WORK

In this section, we will start by giving an overview of pre-5G cellular networks and then introduce the concept of network densification. We will mention UAVs and HAPSs as network elements and how they change the traditional network structure. Furthermore, flow networks and relabel-to-front algorithm, which forms up the basis of the novel algorithm proposed in this thesis, will also be briefly introduced in later sections.

2.1 An Overview of Pre-5G Cellular Networks and Wireless Backhaul

In the first 2G cellular networks, the mobile devices connected to the base transceiver stations (BTS's) wirelessly. These BTS's were then connected to the base station controllers (BSC), which had control plane functionality for the air interface, and a connection to the core network via mobile switching centers (MSC). All of this connection except the one between the mobile devices and BTS's were wired E-1 based or IP-based connections. Initially, GSM networks had the task of carrying voice but with the introduction of GPRS, GSM networks also started carrying data. The voice used circuit-switching, which meant that required resources for the call were allocated on the call setup and then released after the call. On the contrary, data was transported on the Internet using packet switching, meaning that 2G networks had to implement and use both stacks on their network elements.

In 3G, base stations of UMTS (which are called Node-B's) were connected with high-speed DSL connections, microwave radio, fiber and microwave Ethernet lines. Moreover, the transition was to the IP-based protocols instead of legacy SS7. Like its

predecessor, 3G networks also carried voice and data using the same infrastructure. This duality imposed more and more challenges in network design and operation. For example, any network element had to implement one stack for voice traffic with its own protocols and another one for data traffic. Moreover, scheduling and handling two different stacks added complexity to network management and operation. Because of these reasons, especially during 3G, the transition from circuit-switching to packet-switching became very common due to the increasing effect of data services over the whole network. With a packet-switched network, all services can be implemented on a unified core, resulting in better resource utilization and improved cost efficiency [1].

Despite the difference in voice and data service, networks usually used the same backhaul lines to the controllers and to the core network. The ATM-based transports provided constant throughput, which was suitable for the voice traffic but not for data traffic since it had a bursty nature. Ethernet was a technology that has proven itself time and again for the data communications. In spite of this reliability, Ethernet could not be used directly in the telecommunication networks due to some inherent differences, such as the lack of synchronization signals, even though it delivered higher capacity at a lower cost per bit [2]. Finally, with the introduction of Carrier Ethernet, these shortcomings were solved and operators were able to unify the voice and data communication on the same network without any additional supporting protocols [1].

With the introduction of LTE, the network adopted an all-IP approach, which finally unified the voice and data duality. This unification is realized by transitioning the voice telephony to the IP side. Network design was also greatly simplified with this unification, as the network now implemented and operated only one stack. Perhaps the biggest advantage of IP architecture was that all interfaces became IP compliant, and physical infrastructure became transparent and interchangeable. These developments made it possible to change the backhaul medium quite easily and opened new possibilities to implement wireless backhauling. To give an example, using the same interface for all connections effectively meant that theoretically, a BS could connect to another BS since both used the same protocols and procedures.

For GSM/GPRS networks, only wireless backhaul solution was microwave radio at

this time. Although backhauling using satellite was possible, it was too expensive and cumbersome to be of wide use. At that time, passive optical networks (PON) and mmWave radio were seen as the future backhaul technologies enabling higher datarates with better cost-efficiency [1]. However, while the mmWave technology has the potential to support Gbps-level speeds, it was not fit for cellular use due to its power consumption and hardware characteristics up until 5G era [3].

For LTE and LTE-Advanced networks, relays were introduced as a new network element to improve network coverage and range by employing wireless backhauling. 3GPP TR 36.806 [4] detailed the relay architectures. There were two types of relays: Type 1 relays had cells and identities of their own and had the same radio resource management (RRM) mechanisms as the eNBs, Type 2 relays did not have identities and some parts of the RRM was controlled by their donor eNB [5]. Type 2 relays supported L1 (decode-and-forward) and L2 relaying, which did not require the implementation of the whole RRM stack and consequently made these type relays simpler. Type 1 relays also supported L3 relaying and even self-backhauling, which we will mention later.

Both Type 1 and Type 2 relays were dependent on donor eNBs for their operations, and while Type 1 was capable of independent operations with a unique identity, it was still classified as a relay, not as an eNB with wireless backhaul capabilities. This shortcoming also shows the expectation of the network maintainers from relays: they were seen as a way to extend network coverage. This is more evident in the later version of the relays in Release 11 where their operation scenarios were defined as coverage extension and indoor coverage [6].

While the proposed concepts did have some good points and were a potential game-changer for contemporary networks, LTE relays did not see widespread use. There are two major reasons for this: the first reason is that in pre-5G networks, frequency was a very scarce resource. LTE networks were stuck to sub-6 GHz frequencies for access, and this weakened the LTE relay concept significantly simply because using the same frequencies for backhaul meant taking away from the access side which was already strained. The second reason is that the networks did not meet the densification projections that were made for LTE networks. In a sense, the networks did not get

dense enough to justify the deployment of wireless relays and allocation of spectrum from access to backhaul side.

To make eNB's capable of wireless backhauling, self-backhauling concept was introduced in LTE-Advanced. This new concept gave eNB's full relay capability: in essence, self-backhauling enabled eNB's to use the same radio resources for both access and backhaul usage. Self-backhauling improved spectrum efficiency through resource reuse, and cost efficiency through hardware and management tools reuse. However, the self-backhaul brought about a few challenges of its own, such as access-backhaul interference, and scheduling between access and backhaul [7]. Gamboa and Demirkol [8] have experimented using the LTE self-backhauling solution and have shown improvements in both the coverage and downlink bitrate of the network, despite self-interference and in-band communication reducing the effectiveness of the overall setup.

The challenges introduced with self-backhauling and relays will be handled with the new technology introduced in 5G. Bhushan *et al.* highlight network densification as the key mechanism for future wireless evolutions [9]. This densification is twofold: spatial densification and spectral aggregation. The spatial densification is achieved by increasing the number of antennas per node as well as increasing density of base stations deployed in a given area. The spectral aggregation is achieved when larger amounts of bandwidths are used. However, these concepts are beneficial only if the backhaul can support the denser network. The details of these developments and challenges will be given in the next subsection.

2.2 Network Densification in 5G

The three major use-cases of the 5G, namely eMBB, URLLC and mMTC, has shaped the 5G network architecture. These three use-cases necessitate the development of different technologies and concepts to meet the service requirements, sometimes with conflicting targets. The network architecture has become increasingly denser over the years, from 4-5 BS/km^2 in 3G to an anticipated 40-50 BS/km^2 in 5G [10]. This densification cannot be achieved with the existing macrocell architecture, and small

cell architecture is introduced to realize the densification.

There are some initial challenges related to the densification. First challenge is the cost of deployment for SBSs. Fiber backhauling solutions may be effective in terms of the connection quality it provides, but not only it is costly, but installing fiber is also a tedious process which takes time and effort. Moreover, there may be some infrastructural challenges for cabling certain areas, such as roads or densely populated urban centers. Costs can be significantly reduced by employing wireless backhauling on small cells as they do not require any extra infrastructure. Furthermore, since wireless access and backhaul can use the same medium and protocols, there is further cost reduction through hardware reuse.

Second challenge related to the densification is the performance. Densification can only be realized with wireless backhauling, but the performance of wireless backhauling with sub-6 GHz frequencies is nowhere near fiber. In wireless backhaul, a new type of interference is introduced into the system: namely the access-backhaul interference, which reduces the performance of both links. In releases 10 and 11, enhanced inter-cell interference coordination (eICIC) and further enhanced inter-cell interference coordination (feICIC) were introduced to reduce interference in order to extend the coverage range in heterogeneous networks (HetNets) [11, 12, 13]. Traditional fiber backhaul solutions do not suffer from any interference problems and perform adequately well in any situation.

Even if the interference problem is solved, the spectrum used in LTE is not in a capacity to satisfy the requirements of 5G networks. With the introduction of mmWave spectrum in Release 15, the wireless backhaul finally has the potential to reach the performance of its fiber counterpart. With the beam-based air communication in mmWave spectra, the access-backhaul interference no longer becomes the limiting problem, as they are noise-limited unlike the interference-limited LTE networks [14].

Even though the mentioned challenges are still not solved, densification and wireless backhauling bring new capabilities to the future networks. For example, the developments on wireless backhaul significantly improves the performance of satellite and aerial networks. Furthermore, wireless backhauling makes rural connectivity easier by increasing the flexibility in deploying the networks with reduced or no infrastruc-

ture. Wireless backhaul also aids mobile edge computing cases by making it easier to deploy edge servers closer to the users.

The developments in wireless backhaul significantly aid the densification of the network. Deploying small cells in lamp posts, on ground vehicles, and UAVs increase the number of alternative cell locations or overall deployment schemes, which in turn makes achieving densification easier. Recent research is also focused on using UAVs for a myriad of purposes, from bringing service to remote areas such as highways [15] to acting as MEC servers for IoT computation offloading [16].

2.3 Aerial Network Elements

HAPS and UAV are aerial platforms that can arbitrarily change their locations and act as mobile network components. Their high altitude enable them to maintain line of sight for reliable communications. Moreover, they can move closer to the UEs to allow transmission with lower power. Mobility of these solutions introduce completely new opportunities to improve the effectiveness of the networks or solve old problems efficiently enough to be considered by the service providers. However, just like the wireless backhaul, HAPS/UAV usage have challenges of their own such as node placement, air-to-ground channel modeling and resource management [17, 18, 19].

2.3.1 The State of the Art

Azizi *et al.* propose to use profit maximization as a utility function to solve the joint radio resource allocation, air base station (ABS) altitude determination, and user association in a heterogeneous network of ground and air base stations [20]. Since the given problem is a MINLP, it is decomposed into three subproblems, which are iteratively solved until convergence. One of these subproblems is an INLP, which is slightly challenging compared to the other two subproblems which are continuous. This INLP is solved without any relaxation methods. The proposed method has linear complexity in terms of users, which is a significant improvement compared to the similar methods having cubic or higher complexities. Results show that macro aerial base stations (MABS) can operate at higher altitudes than small aerial base sta-

tions (SABS) and cover a larger area. However, lower operating altitude means lower path-loss for SABSs. If all of the optimization variables are used jointly (altitude determination, power allocation, sub-carrier allocation, and user association), a network profit increase of 47% can be achieved. Moreover, compared to having a single ground base station, employing two MABS and two SABS almost doubles the overall network profit. These performance improvements are more evident in suburban & urban areas, and decrease for dense & high-rise urban areas.

Jaffry *et al.* cover the moving networks in 5G in their work [21]. In this work, non-terrestrial networks supported backhaul is highlighted as a future direction in this research area. UAV/HAPS are envisioned to give wide area coverage with wider area and near-LoS coverage. UAVs' low cost also allow them to act as relays with ease of deployment and extended operation times. These UAVs can provide backhaul links to ground-based networks. While this provides a new degree of freedom, there are also challenges related to three-dimensional channels that these UAVs will have to use. Even though UAV/HAPS are mentioned only as a future direction, this work is very comprehensive in the area of land-based moving networks with challenges, use-cases and applications and any we encourage all interested readers to consult [21] for these topics.

Wang *et al.* investigate the successful content delivery performance in integrated UAV-terrestrial networks [22]. In the proposed model, caching-enabled UAVs are used to offload the bursty traffic from the terrestrial cellular networks. UAVs use self-backhauling and share the spectrum resources with the terrestrial network. Wang *et al.* derive a closed-form expression for the achievable rate of the mmWave backhaul link for UAVs under general and noise-limited cases. Then, the authors analyze the minimum cache hit probability to achieve a certain backhaul rate requirement. Finally, using stochastic geometry, Wang *et al.* also analyze the successful content delivery probabilities with an approximated LoS model. Numerical results show that adding UAVs with LoS increases the successful content delivery probability. Self-backhauling with mmWave is shown to achieve adequate results, but with increased UAV density the achievable rate is reduced. Minimum cache hit probability is positively correlated with the backhaul rate and UAV density. Caching-enabled UAV network also outperforms the terrestrial-only network in terms of successful content

delivery probability. Increasing UAV density increases the successful content delivery probability up to a certain point, after which the probability decreases because of the LoS interference. The optimal UAV height is also shown to decrease with higher UAV densities.

Wu *et al.* study the joint optimization of user scheduling and trajectory for UAV networks with NOMA usage [23]. K-means clustering is used to partition the users. The problem of joint scheduling and trajectory is a non-linear non-convex optimization problem, and block coordinate descent method is used to divide the original problem into two blocks, namely the multi-user communication scheduling block and UAV trajectory block based on NOMA. Wu *et al.* propose an iterative algorithm to alternately solve the subblocks. Simulation results show that the proposed NOMA method outperforms OMA, NOMA simple circular trajectory and NOMA static position methods in terms of system max-min rate.

Dao *et al.* study the aerial radio access networks (ARAN) in their survey [24]. Dao *et al.* first describe the ARAN architecture and its fundamental features related to 6G networks. Then, the authors analyze ARANs from several perspectives such as energy consumption, latency, transmission propagation, and mobility. Dao *et al.* mention energy refills, network softwarization, mobile cloudization, data mining, and multiple access methods as technologies that enable the success of ARANs. Event-based communications, aerial surveillance, smart agriculture, urban monitoring, health care, and networking in underserved areas are given as the application areas of ARANs. As a final note, Dao *et al.* highlight the open research areas and trends towards 6G ARANs.

Cao, Lien and Liang propose a deep reinforcement learning scheme for intelligent multi-user access to non-terrestrial BSs [25]. Cao *et al.* use a centralized agent deployed in non-terrestrial BSs to train a deep Q-network (DQN), and UEs independently make their access decision based on DQN's input. The authors use a UE-driven scheme, which eliminates the need to retrain the system when the number of UEs is changed. Cao *et al.* also design a new long short-term memory (LSTM) network to capture the time-dependent feature of non-terrestrial BSs. Simulation results show that system throughput increases with a higher number of deployed UEs/non-terrestrial BSs. The proposed deep reinforcement learning scheme outperforms RSS-

based, Q-learning-based, UCB learning-based and random methods in terms of system throughput. The proposed scheme also performs significantly less handovers than other methods.

Mozaffari *et al.* present a tutorial on UAVs in wireless communications [26]. In their tutorial, Mozaffari *et al.* present the potential applications of UAVs, key research directions on these potential applications, and open problems of these research directions. Coverage and capacity extension, disaster-related deployments, connectivity enhancement, 3D MIMO-, IoT-, and cache-enabling are highlighted as potential applications for UAVs as base stations. Flying backhaul is also highlighted as another potential application to enable cost-effective, reliable, and high-speed wireless backhaul connectivity for terrestrial networks. Air-to-ground channel modeling, optimal deployment, trajectory optimization, cellular network planning and provisioning, resource management and energy efficiency, and drone UE usage are highlighted as challenges and open problems. Centralized optimization, optimal transport theory, machine learning, stochastic geometry, and game theory are proposed as the analytical frameworks to enable UAV-based communications.

Khamidehi and Sousa investigate the trajectory optimization problem for multi-aerial base station (ABS) networks [27]. The objective of the optimization problem is to maximize the minimum data rate of cell-edge users under ABSs power, backhaul link capacity, and collision avoidance constraints. ABSs are first partitioned into clusters with a modified K-means approach so that UEs are served by their associated ABS. Then, the optimization problem is divided into three sub-problems, namely power allocation, joint ABS-user association and subchannel asignment, and trajectory optimization. An iterative method with successive convex approximation is used to efficiently solve these subproblems. Simulation results show that increasing the maximum flight time also increases the max-min data rate since ABSs have more time to reduce the distance to the cell-edge users. With increased maximum flight time, ABSs also adjust their trajectories so that they can visit all associated users. Backhaul capacities are given as the bottleneck of the given problem, but after a certain increase in backhaul capacity, the bottleneck becomes the maximum transmit power constraint. Constraining the propulsion power of ABSs also increases the service time; without this constraint, ABSs cannot finish their mission with the available power.

Kurt *et al.* present a vision for future HAPS networks and state-of-the-art literature review [28]. Kurt *et al.* first present HAPS use-cases for next generation networks. HAPS-Mounted super macro base station (HAPS-SMBS) is proposed as a complementary solution to terrestrial systems. These HAPS-SMBSs are envisioned to support data acquisition, computing, caching, and processing. Use-cases for HAPS-SMBSs are given as supporting IoT devices, backhauling SBSs, covering unplanned user events, operating as aerial data centers, filling coverage gaps, supporting and managing aerial networks, supporting intelligent transport systems, and handling LEO satellite handoffs and providing seamless connectivity. Kurt *et al.* present a general view of HAPS system and its subsystems. Then, the authors give an overview of the channel models, radio resource management, interference management, and waveform design for HAPS systems. Kurt *et al.* also highlight the contributions of machine learning in design, topology management, handoff, and resource allocation problems of HAPSs. Finally, the authors present the challenges and open issues in two groups, namely next-generation (10 years) and next-next-generation (20 years) and provide possible roadmaps.

It is worth mentioning that machine learning approaches are widely used for UAV problems. UAVs form a distributed network and with machine learning, UAV networks can effectively become self-organizing. In the previous sections, we have mentioned numerous works using machine learning for UAV-related problems [29, 30, 31, 32, 15]. In the future, employing fast machine learning methods and distributive approaches will result in highly flexible UAV networks that require little to no management at all.

2.3.2 Open Issues and Challenges

Although there are many use-cases related with UAVs and HAPSs, there are also open issues prohibiting wide usage. For example, regulatory aspects for UAV/HAPS operation are not uniform and dependent on the location. Furthermore, frequency usage is also crucial for UAV/HAPS since it directly affects performance. As we mentioned before, higher frequencies have different licensing schemes, which may impose limitations on UAV/HAPS operations. Handling the regulatory aspects re-

garding UAV/HAPS is a challenge that needs to be solved.

The mentioned use-cases in [28] have different requirements and hence different problems. For instance, HAPS operating as a data center have different requirements than a HAPS extending coverage. Since there is a high diversity in terms of use-cases, system models also have different challenges, and integrating multiple of these use-cases in one design remains an important open research area.

Radio resource management with UAV/HAPS also remains an open challenge. Low computational overhead is desired with UAV/HAPS systems and machine learning approaches are promising in this regard. Moreover, serving different 5G use-cases such as URLLC and mMTC together is another challenge in terms of radio resource management. In such systems, objectives are completely different and changing the service type or supporting multiple use-cases at the same time requires a specific design. UAV/HAPS operations have to support multiple use-cases and switch between them seamlessly as needed, and radio resource management seems to be the key concept to realize this switch.

Network stability is another problem with UAV/HAPS usage. Since aerial systems have potentially high mobility and use aerial ad hoc technologies for interconnection, guaranteeing stability in such a dynamic network is difficult. Hierarchical systems formed between satellites, terrestrial networks and UAV/HAPSs are helpful in this regard, but highly dynamic topologies bring challenges to keep up the existing state of the network and high performance. To alienate the network stability problem and handle the dynamic nature of UAV/HAPS networks, routing protocols and network organization designs seem to be directions to handle different mobility patterns, traffic characteristics, and implement redundancy. To summarize, network stability remains an open research area for UAV/HAPS integration into contemporary networks.

Channel models for 3D designs are also not developed enough and requires investigation. Air-to-air and air-to-satellite links require further elaboration as the channel model directly affects system design and performance. Having accurate channel models can pave the way for the researchers to make performance evaluations to show that UAV/HAPS systems can perform well in their mentioned use-cases. Thus, channel modeling and performance evaluation remain open research challenges.

2.4 Flow Networks

In computer science, graph is a data structure consisting of vertices and edges that link them together. Graphs are used to represent a wide range of problems in computer science, and quite effective solutions are developed for these problems. Moreover, graph is a frequently used data structure by other areas as well. The reason is twofold: graphs are naturally representative of relationships between different components, and there are a myriad of developed methods whose effectiveness have been proven both in scientific and commercial context.

Flow networks are a type of directed graphs. Let G = (V,E) be a flow network with the capacity function c that indicates the maximum flow associated with an edge. There are two distinguished vertex types, namely source s and sink t. A flow f in G is a real-valued function satisfying three properties:

- Capacity constraint: For all $u, v \in V$, $f(u, v) \leq c(u, v)$
- Skew symmetry: For all $u, v \in V$, $f(u, v) = -f(v, u)$
- Flow conservation: For all $u \in V - \{s, t\}$, $\sum_{v \in V} f(u, v) = 0$

First property indicates that flows cannot exceed the capacity of their respective edges. Second property indicates that between two vertices, flow values must be the same with reversed signs. Third property indicates that except source and sink vertices, vertices cannot accumulate flows in them.

Flow networks are frequently used to model a source flow in a graph. For example, water pipe or electrical current modeling, supply logistics, communication networks, traffic and road capacity modeling can all be done using flow networks. Maximum-flow problem is the simplest and perhaps most used problem which finds the maximum flow that can be sent from the source to the sink while conforming to flow constraints.

2.4.1 Solutions to the Maximum-Flow Problem

There are various methods to solve the maximum-flow problem. Ford-Fulkerson method is an iterative method with several implementations. Another type of solution is push-relabel algorithms. Both solutions make use of a number of concepts that need introduction to grasp their inner workings. Residual networks, augmenting paths, and cuts are the aforementioned concepts.

Given a network and a flow, the residual network consists of the edges that can admit more flow. Mathematically, the residual capacity c_f is defined as the admissible flow, which is found by subtracting the allocated flow from the capacity:

$$c_f(u, v) = c(u, v) - f(u, v) \tag{2.1}$$

The set of residual edges E_f also need to be defined for the residual network definition:

$$E_f = \{(u, v) \in V \times V \text{ s.t. } c_f(u, v) > 0\} \tag{2.2}$$

The residual edges are the edges in G with positive residual capacity. Using the residual edges, we define the residual network $G_f = (V, E_f)$. In the residual network, an augmenting path p is a simple path from source to sink. By definition, each edge on the augmenting path admits additional positive flow without violating capacity constraints. The residual capacity of p is the minimum residual capacity on the augmenting path p, since that is the maximum amount of flow that can be admitted without violating capacity constraints.

A cut (S, T) of a flow network is a partition of V into S and $T = V - S$ such that $s \in S$ and $t \in T$. The net flow across the aforementioned cut is $f(S, T)$ and the capacity of the cut is $c(S, T)$. A minimum cut of a flow network is the cut with minimum capacity out of all possible cuts. Using these definitions, the max-flow min-cut theorem is defined as follows:

If f is a flow in a flow network $G = (V, E)$ with source s and sink t, then the following

conditions are equivalent:

1. f is a maximum flow in G
2. The residual network G_f contains no augmenting paths
3. $|f| = c(S, T)$ for some cut (S, T) of G

Ford-Fulkerson method finds an augmenting path p and increases the flow f on every edge on p by the residual capacity $c_f(p)$. If no augmenting paths can be found, then by max-flow min-cut theorem, the flow f is a maximum flow. Edmonds-Karp algorithm, which selects the augmenting paths using a breadth-first search, has a complexity of $O(VE^2)$. Another family of algorithms, namely the push-relabel algorithms, can be used to solve the maximum-flow problem as well. In terms of asymptotical speed, push-relabel algorithms include the fastest algorithms to solve the maximum-flow problems.

The approach in push-relabel algorithms is different than the Ford-Fulkerson method since push-relabel algorithms work on one vertex at a time, and use its neighbors in the residual network. Moreover, push-relabel algorithms do not maintain the flow conservation property during their execution. Instead, these algorithms maintain a preflow satisfying capacity constraints, skew symmetry, and a relaxed version of flow conservation allowing nonnegative total net flows in each vertex other than the source. Excess flow $e(u)$ is the total net flow at a vertex u.

Push-relabel algorithms make use of a height concept that allows pushing flows 'downhill', from a higher vertex to a lower vertex. Initially, height of the source s is fixed at $|V|$, and the sink t at 0. All other vertices initially have 0 height and these increase with time. Initially, the algorithm sends the capacity of the cut $(s, V - s)$ from the source. Suppose that every intermediate vertex has a reservoir that vertices use to store the excess flow, if any. This reservoir is used during the execution of the algorithm, whose contents will be pushed downhill as the execution progresses. To push the reservoirs downhill, "relabel" operation is used to increase the height of a vertex. After relabeling and pushing flows until they arrive to the sink, there comes a point when the maximum admissible flow is reached. At that point, the excesses that re-

main in the reservoirs are pushed back to the source so that the preflow adheres to the flow conservation constraint.

There are three operations that the push-relabel algorithm uses. The algorithm first initializes a preflow using InitializePreflow procedure given in Algorithm 1. The algorithm first sets all heights, excesses, and flows corresponding to edges to zero. Source height is set to the number of vertices. Finally, all flows from the source is filled fully by the capacity, and the reverse flows and excesses are updated accordingly.

Algorithm 1 Preflow Initialization

```
1:  procedure INITIALIZEPREFLOW(G, S)
2:      for each vertex u ∈ V[G] do
3:          u.height ← 0
4:          u.excess ← 0
5:      for each edge (u, v) ∈ E[G] do
6:          f[u, v] ← 0
7:          f[v, u] ← 0
8:      for each vertex s ∈ S do
9:          s.height ← |V[G]|
10:         for each vertex u ∈ Adj[s] do
11:             f[s, u] ← c(s, u)
12:             f[u, s] ← −c(s, u)
13:             s.excess ← s.excess − c(s, u)
14:             u.excess ← c(s, u)
```

After initializing the preflow, push-relabel algorithm performs push or relabel operations until there are none available. Algorithm 2 defines how the push operation is performed. Push assumes that vertex u has a positive excess, the residual capacity of (u, v) is positive, and u is one height higher than v. The operation finds the admissible flow to (u, v) and pushes this flow by adding this number to the flow (and subtracting it from the reverse). Excess values of vertices are also updated accordingly.

If there are no available push operations, it is either the case that there exist no augmenting paths in the residual network, or vertex heights do not allow push operations. If first case holds, then the algorithm finishes. Otherwise, relabel operations are ap-

Algorithm 2 Push

1: **procedure** PUSH(u, v)

Require: $u.excess > 0, c_f(u,v) > 0, u.height = v.height + 1$

2: $d_f(u,v) \leftarrow min(u.excess, c_f(u,v))$

3: $f[u,v] \leftarrow f[u,v] + d_f(u,v)$

4: $f[v,u] \leftarrow -f[u,v]$

5: $u.excess \leftarrow u.excess - d_f(u,v)$

6: $v.excess \leftarrow v.excess + d_f(u,v)$

plied so that augmenting paths can be used to push the flows. Algorithm 3 updates the height of u so that a push operation can be performed. To update the height, a vertex v with the minimum height is chosen among the vertices that are adjacent to u in the residual graph.

Algorithm 3 Relabel

1: **procedure** RELABEL(u)

Require: $u.excess > 0, (u,v) \in E_f$ and $u.height \leq v.height\ \forall v \in V$

2: $u.height \leftarrow 1 + min\{v.height : (u,v) \in E_f\}$

The generic implementation of push-relabel algorithm has an asymptotic bound of $O(V^2E)$. Relabel-to-front algorithm defined in Cormen *et al.*'s book Introduction to Algorithms' Chapter 26.5 [33], which also makes use of the same basic concepts, improves this bound to $O(V^3)$. To understand the algorithm, concepts of admissible edges and neighbor lists have to be defined. An edge (u,v) is admissible if push operation can be called on its vertices. Neighbor list $N[u]$ for a vertex contains exactly the vertices for which there may be a residual edge (u,v).

The main operation of relabel-to-front algorithm is defined in Algorithm 4. Discharge pushes all excess flow of a vertex u through admissible edges to neighboring vertices until u becomes inadmissible. In the algorithm, $current[u]$ gives the current vertex under consideration in $N[u]$, and $next - neighbor[u]$ gives the vertex that comes after v in $N[u]$, which is either another vertex or $NULL$. Discharge method tries to push excess flow to any available admissible edge via push method. If there are no admissible edges available, then the method calls relabel to increase the height of u so that the excess flow can be pushed. This process continues until all excess flow is

pushed from u.

Algorithm 4 Discharge

```
1:  procedure DISCHARGE(u)
2:      while u.excess > 0 do
3:          v ← current[u]
4:          if v = NULL then
5:              RELABEL(u)
6:              current[u] ← N[u].head
7:          else if c_f(u, v) > 0 and u.height = v.height + 1 then
8:              PUSH(u,v)
9:          else
10:             current[u] ← next-neighbor[v]
```

Relabel-to-front algorithm is defined in Algorithm 5. Similar to push-relabel algorithm, relabel-to-front algorithm first initializes preflow and gets all vertices except source and sink in a list L. The algorithm also initializes all current neighbors for vertices in L. After initialization, the algorithm picks the first vertex in L as u and starts a loop. In this loop, the initial height of u is saved and then DISCHARGE(u) is called. If a RELABEL operation is made in this call, then u is moved in front of L. Finally, the vertex after u in L is chosen as u and the loop continues until u becomes $NULL$.

Relabel-to-front algorithm is important for this thesis as the proposed algorithms make use of relabel-to-front algorithm heavily. The methods and algorithms mentioned in this chapter will frequently be referenced in later sections when explaining the proposed algorithm in the thesis.

2.5 Related Work

Having mentioned the general concepts that are seen in most papers of wireless backhauling area, we now present the reader our survey. Further sections will present the challenges of wireless backhauling, state-of-the-art in these topics, main learning points, and open issues regarding these areas.

Algorithm 5 Relabel-to-Front

```
1:  procedure RELABEL-TO-FRONT(G, 𝒮, 𝒯)
2:      INITIALIZEPREFLOW(G, 𝒮)
3:      L ← V − {𝒮, 𝒯}
4:      for each vertex u ∈ V[G] − {𝒮, 𝒯} do
5:          current[u] ← N[u].head
6:      u ← L.head
7:      while u ≠ NULL do
8:          old-height ← u.height
9:          DISCHARGE(u)
10:         if u.height > old-height then
11:             current[u] ← N[u].head
12:         u ← next[u]
```

For this purpose, we have highlighted six main topics: resource allocation, deployment, scheduling, performance evaluation, coverage, and security. Each of these topics have their own section with the same structure defined in the previous paragraph. These topics are especially important because their solutions are directly related to the effective realization of wireless backhauling. The specific points regarding their importance to the wireless backhauling are highlighted in the respective sections.

Wireless backhauling is an area that shares many challenges with traditional networking. Deployment, security, and coverage are examples of such challenges. Furthermore, there are challenges more specific to wireless backhauling due to different characteristics, such as scheduling of small cell backhaul traffic or deployment of small cells for high performance. Resource allocation seems to be the paramount problem for wireless backhauling as the dynamic management of resources is required to effectively use the scarce resources available. Finally, performance evaluation works are significant since these show the feasibility of wireless backhaul usage in the field without any performance losses, while also highlighting the necessary adaptations for effective field usage.

2.5.1 Resource Allocation

The first challenge before the implementation of any wireless backhaul system is the resource allocation problem. Resource in this context can be defined as anything that the network manages to perform its operation. Examples can be the bandwidth, energy, links and their capacities, or time-slots. Networks use these resources to optimize their performance. Frequently, the bandwidth or energy allocation has to be carefully adjusted to get the best possible performance from a wireless backhauled network.

User association is another frequently seen aspect of the resource allocation problems. For HetNets, a UE can get service from multiple BSs at any given time since SBSs and their associated MBSs often have overlapping coverage areas. Because of this, user association becomes another decision which can be used to optimize overall network performance. In this regard, approaches may be to maximize the average sum rate or the worst performing UE. User association is also used for load balancing of various resources, meaning that sub-optimal associations may take place. User association is mostly taken into account with another parameter such as transmit power allocation or spectrum allocation/scheduling.

2.5.1.1 State of the Art

In [34], Bonfante *et al.* use massive MIMO to construct a two-layer network in which small cells use wireless access and backhaul links. The two-layered system consists of mMIMO-BSs and small cells that use a fixed allocation of access and backhaul slots. All small cells receive their respective backhaul traffic from the base stations employing mMIMO. These BSs are connected to the core network with high-capacity wires and serve only the small cells, whereas the small cells only serve UEs and do not take part in backhaul of other small cells' traffic. For their system, the parameter $\alpha \in [0, 1]$ defines how the available time slots were divided between access and backhaul usage. Two different deployment scenarios are considered for small cells. In the first scenario, the small cells are randomly and uniformly distributed to the BS area; whereas in the second one, which the authors named as ad-hoc deployment, they

are deliberately placed closer to UEs to increase the throughput.

Tang *et al.* optimize beamforming and power allocation to maximize energy efficiency in CoMP networks doing simultaneous wireless information and power transfer (SWIPT) [35]. There are two types of users in these networks, namely the energy harvesting users (e.g. IoT devices) and information decoding users (e.g. laptops, mobile phones). Tang *et al.* separately consider beamforming and power allocation since the combined problem is non-convex. The authors first propose a zero forcing approach for beamforming to suppress interference for all nodes. The authors also propose a partial zero forcing approach for beamforming, which exploits interference as an energy source for energy harvesting users, while at the same time cancelling the interference of information decoding users. Simulation results show that partial zero forcing outperforms the zero forcing method. Moreover, both methods converge to their global optimals after a number of iterations. Increasing the number of small cells and information decoding users also increases the energy efficiency. Increasing the number of energy harvesting users increases the energy efficiency up to a certain number, after which efficiency sharply decreases for zero forcing, and decreases mildly for partial zero forcing.

In [36], Zhang *et al.* have formulated a problem to optimize the energy efficiency of wireless backhaul bandwidth and power allocation problem in a small cell. The network in question is a two-tiered network in which the MBS uses mMIMO to communicate with the small cells whereas small cells use OFDM to communicate with the UEs. The aforementioned problem is formulated as a nonlinear programming problem, which is shown to be non-convex. Therefore, the authors have decomposed the problem into two convex subproblems for wireless backhaul bandwidth allocation and power allocation. Two versions of the algorithm are developed, namely the optimum iterative algorithm and low-complexity algorithm, which fixes the bandwidth allocation factor by the value calculated from the equal power allocation. It is shown that the proposed iterative algorithm gets better results in all metrics than the low-complexity one, whereas they both outperform the existing benchmark.

Nguyen *et al.* study the joint design of downlink transmit beamforming and power allocation in two-tier wireless backhauled small-cell heterogeneous networks [37].

Nguyen *et al.* first formulate the problem of maximizing the energy efficiency of the wireless backhauled small cell HetNets. The proposed power consumption model considers adaptive decoding power at each small cell access point (SAP), making the model more appropriate. The problem is non-convex and NP-hard, and Nguyen *et al.* propose an exhaustive search based on branch-and-bound algorithm for global optimal solution. However, this solution has high complexity, and Nguyen *et al.* also propose a low-complexity algorithm based on first order Taylor convex approximation (FOTCA). As the final step, Nguyen *et al.* also extend the problem formulation to consider the small cell selection that takes into account the impact of power to switch on/off the SAPs. Introducing this as a binary variable, the authors model this as a mixed integer second order cone programming (MISOCP) approximated problem. This problem is decomposed into subproblems by the ADMM approach and Nguyen *et al.* develop a distributed algorithm to solve this problem individually at every SAP. Numerical results indicate that the all proposed methods outperform the fixed power allocation scheme. Moreover, the small cell selection algorithm also outperforms others since selectively activating the SAPs results in significantly less power consumption. Second best algorithm is the adaptive decoding power. Finally, it is shown that having a high power budget usually results in the MBS consuming too much power and reducing the energy efficiency. Nguyen *et al.* suggest that having low budgets result in better energy efficiency since it does not allow excessive power allocation.

In [38], Saha, Afshang and Dhillon explore the bandwidth partition scenarios between access and backhaul side of the network in a full wireless configuration. The authors propose a network that used sub-6 GHz frequencies for control traffic and mmWave for data traffic. For the backhaul traffic of small cells, Saha *et al.* compare two different strategies. In the first strategy, the available bandwidth is partitioned equally to all small cells. In contrast, in the second strategy, the bandwidth is partitioned proportional to each small cell load. The authors report that the load-balanced partitioning always provides higher coverage than fixed partitioning. Furthermore, for a given partition strategy, the existence an optimal access-backhaul split maximizing the rate coverage probability is shown. As a third point, Saha *et al.* show that IAB-enabled network outperforms the macro-only network up to a certain point, beyond which the

performance gains disappear and they converge to the same point.

Huang *et al.* propose a fairness-based distributed resource allocation algorithm (FDRA) [39]. Huang *et al.* use stochastic geometry to model the two-tier HetNet and define the channel reuse radius based on the spectrum sensing threshold. FDRA is proposed to maximize the total throughput in the small cell tier while considering the outage probability and fairness. To model the real traffic load accurately, tidal effect and users' actual geographical distribution are jointly considered and improved FDRA is proposed to further improve the throughput in hot areas. Simulation results show that FDRA improves fairness compared to the benchmark method CSRA. In all metrics, IFDRA outperforms FDRA while FDRA in turn outperforms CSRA.

In [40], Pham *et al.* consider the effects of resource offloading for mobile edge computing on the small cells that use wireless backhauling to a MBS. The authors highlight that the resource offloading research mostly assumes there is backhaul capacity for their purpose, but the real-life situation does not conform to that assumption. The authors decompose this joint task offloading and resource allocation into two parts. In the first part, the decision to offload the computation based on a given bandwidth factor and computation resource allocation is made; whereas in the second part, wireless backhaul bandwidth and computation resource allocation is made for a fixed offloading decision. These two sub-problems are solved individually with an iterative algorithm for the solution of the original problem. Moreover, the authors have extend the problem into multiple scenarios, such as ultra-dense networks, machine learning usage for computation offloading, and partial offloading of the computation. The proposed algorithm is compared to the given problem's two extremes where all computations are either performed locally or offloaded to the mobile edge computing server. As expected, with an increasing number of UEs, fewer offloading decisions are made, since the available resources decrease with a larger number of UEs. Furthermore, the proposed algorithm achieves lower computational overhead since its decisions are made only when offloading is advantageous. The same performance improvements are also seen in UEs' transmit power because the algorithm considers the consumed power while offloading for its decision-making.

Song, Liu and Sun consider the joint radio and computational resource allocation

problem in NOMA-based MEC in HetNets [41]. Song *et al.* minimize the energy consumption while guaranteeing task execution latency. The optimization problem is solved using decomposition and iteratively employing alternating optimization. Simulation results show that partial offloading outperforms full offloading and local computing in terms of energy efficiency. Increasing task tolerance latency increases energy efficiency for all configurations.

In [42], Han *et al.* investigate the user association and resource allocation problem in the mMIMO-enabled HetNets. The network in question uses renewable energy but is also connected to the grid in case of deficiencies. Han *et al.* formulate a problem of network utility maximization subject to backhaul, energy, and resource constraints. The given problem is solved using first a primal decomposition, and then a Lagrange dual decomposition. In the first decomposition, the problem is divided into two parts: the lower-level problem i.e. resource allocation problem of BSs and higher level problem i.e. cell-association problem of users. The lower level problem is also divided into N subproblems for individual BSs. The higher level problem is solved using the solution of resource allocation problem. To make this a distributive algorithm, the Lagrange dual decomposition is used to divide the problem into dual and master problems. In the resulting distributive algorithm, the UEs calculate and choose the optimal BS allocation at their side and report this connection to BSs; whereas BSs solve the dual problem and then broadcast the next turn's information to users. Furthermore, Han *et al.* design a virtual user association and resource allocation (vUARA) scheme, which reduces the communication overhead over the air interface and removes the information leaks to other users. In this scheme, after the initial resource allocations are made and user associations are reported, the measurements are reported to the radio access network controller (RANC), in which vUARA resides. RANC simulates the iterative cycles and after finding an optimal user association and resource allocation, reports this to the users and BSs. The proposed algorithm outperforms the common user-association schemes, i.e., max-SINR association and range expansion association.

Le *et al.* consider network and weighted sum energy efficiency of dense HetNets [43]. The authors propose a resource allocation algorithm supporting both delay-sensitive users with minimum data rate requirements and delay-tolerant users with

data rate fairness considerations. The problem is formulated and solved in three different ways. Mixed-integer programming, time-sharing, and sparsity-inducing formulations are employed to tackle the non-convex functions with efficient sub-optimal solutions. All formulations make use of successive convex approximations to find the solutions iteratively. Numerical results show that the proposed formulations perform quite similar from an energy efficiency perspective in terms of static consumed power and maximum transmit power at femtocells. Mixed-integer programming formulation is shown to be superior in terms of energy efficiency to other approaches when users' QoS and number of femtocells are considered.

In [44], Omidvar *et al.* propose a hierarchical optimization method for HetNets with flexible backhaul. In flexible backhaul, only some network elements are connected to the core with fiber connection and the rest of the network must reach to these BSs via wireless connections with dynamic links. The authors formulate a two-timescale hierarchical RRM control problem for HetNets with flexible backhaul. For the overall problem, RRM control variables are divided in two as long-term and short-term control variables. For long-term control, flow control, routing control and discontinuous transmission control are used whereas for short-term control, link scheduling control is used. From these variables the problem is formulated. The problem is then divided into two sub-problems by using primal decomposition. The inner problem becomes the optimization of routing control and flow control under a fixed DTX control and link scheduling policy. For the outer problem, an iterative algorithm is proposed whose solution is shown to be optimal. At each subframe, the scheduling information is sent to the UE and UE sends back the instantaneous SINR value as feedback. At each superframe, the RRM system receives the average associated link rates of BSs and sends back the next DTX time-sharing scheme, routing and flow control updates. The proposed algorithm performs better than five chosen baseline methods in terms of network utility for different backhaul connected BS percentage and BS power. The algorithm converges faster than all baseline methods, has similar signaling overhead, is shown to be more cost-saving compared to the baseline methods and adds only little overhead in terms of CPU utilization.

Gao *et al.* propose a quantum coral reefs optimization algorithm for joint resource allocation and power control problem in D2D HetNets [45]. Gao *et al.* first propose

a model for cooperative D2D HetNets and then derive analytical formulas for the total throughput of this model. A combination of coral reefs optimization algorithm and quantum evolution are proposed as a novel quantum coral reefs optimization algorithm (QCROA) to optimize the resource allocation and power control problem. The strength of this approach is that the same algorithm can be used for different parameters by changing the parameter vector, which showcases the reusability of this algorithm. Simulation results show that the proposed algorithm outperforms all alternative methods in throughput under different maximum power, cellular unit and idle relay count constraints.

In [46], Liu *et al.* have considered the resource allocation problem on a two-tier network that uses mMIMO on the MBS and in-band full-duplex communications on small cells. User association and spectrum allocation in the aforementioned network scheme is formulated as a mixed-integer nonlinear programming (MINLP) problem. To solve this problem distributively and efficiently, the primary problem is divided into two sub-problems, namely the user association problem for fixed spectrum allocation, and spectrum allocation with given user association. Next, an iterative algorithm is developed to solve the original problem. The proposed algorithm is compared with the max SINR user association algorithm and shown to be outperforming it in terms of sum logarithmic rate. It is also shown that the optimal spectrum allocation outperforms any fixed spectrum allocation scheme.

Song, Ni and Sun investigate the distributed power allocation problem in non-orthogonal multiple access HetNets [47]. The power allocation problem is modeled as a two-stage Stackelberg game where MBS is the leader and SBS is the follower. MBS tries to maximize its throughput under power and minimum rate constraints whereas SBS tries to find the best response to MBS' actions. Simulations show that the proposed algorithm converges fast and is better in terms of throughput compared to Nash game, dedicated spectrum allocation, and orthogonal multiple access.

In [48], Bojic *et al.* have proposed a dynamic resource allocation algorithm for mobile backhauled networks. The authors state that the problem with mobile backhaul resource allocation is often the assigned capacity. If this is made statically, resources are wasted in lightly loaded areas, whereas areas with heavy loads will experience

congestion. For dynamic assignment, the size of the resource chunks assigned needs to be carefully considered; if the size is too large, resources will be wasted, if they are small, then too many signaling may occur. The authors propose a backhaul resource manager (BRM) that fairly allocates the resources based on the network topology, link capacities and resource allocations. BRM is linked with the network management system and/or with the backhaul nodes to receive the needed information. From the aforementioned information, the BRM will perform the capacity-aware path computation. It is shown that the dynamic algorithm outperforms SPF by 40% and CSPF by 20-30%.

Hu *et al.* propose a power allocation model to maximize energy efficiency in HetNets [49]. The problem is modeled as a probabilistic fractional programming optimization problem. Bernstein approximation and Dinkelbach method are used to transform the problem into a standard convex approximation problem. Simulation results show that the transmit power convergence is quite fast and outperforms the rate maximization-based robust power allocation algorithm in terms of energy efficiency. The proposed algorithm is slightly worse than the energy-efficiency based optimal power allocation algorithm, but the aforementioned algorithm considers perfect CSI, which is not available in the actual communication scenario.

In [50], Faruk *et al.* have thoroughly analyzed the energy consumption characteristics of macro and small cells. It is shown that wireless backhaul, especially with PtMP microwave links, results in high power consumption compared to other alternatives. BS equipment and cooling are also shown to be the most power-consuming tasks. Self-backhauling is shown to be an alternative to other wired and wireless backhaul methods due to its advantageous characteristics such as reduced deployment and maintenance costs, fast rollout and relaxed constraints compared to LoS deployments. The power consumption of small cells with fixed wireless access to MBS and self-backhauled small cells are compared. Since PtP links consume power independent from the load, they are shown to be more power-efficient than self-backhauling as the BS load increases. The break-even point is shown to be 50% load. It is given that after 55W power consumption on fixed wireless radio links, they become less-efficient than self-backhauling solutions. This result is more evident if the number of MBSs increase. Lastly, it is shown that the implementation of PtMP links where a single

PtMP unit serves several small cells results in a more power-efficient solution as the number of served small cells increase.

In [51], Wang, Hossain and Bhargava have delved into the topic of radio resource management of 5G small cell backhauling. General aspects such as cell association, RAT selection, resource and interference management, cell coordination, spectrum sharing and energy-efficient backhauling are explored. Then, some existing works about these aspects are briefly analyzed. Finally, a case study is conducted with a mMIMO-based in-band wireless backhauling network is given. It is shown that user rates can significantly be increased with backhaul-aware cell association and bandwidth allocation schemes.

Siddique, Tabassum and Hossain compare the performance of in-band full-duplex (IBFD) and out-of-band full-duplex (OBFD) backhauling [52]. Siddique *et al.* formulate a problem to maximize the minimum achievable rate at the small cells in a hybrid IBFD/OBFD setting. This problem is solved in a centralized fashion by transforming the original problem into the epigraph form. The authors also solve the individual IBFD and OBFD backhauling problems, which provides useful insights into the nature of these schemes. Then, solutions are found for two distributed backhaul spectrum allocation schemes, namely maximum received signal power (max-RSP) and minimum received signal power (min-RSP). Max-RSP scheme assigns a larger amount of backhaul spectrum to closer SBSs, offering higher backhaul rates to SBSs at the cost of lowering the available access spectrum. Min-RSP scheme conversely allocates larger backhaul spectrum to the SBSs that are farther away from the MBS. Numerical results are validated using Monte Carlo simulations. These simulations show some interesting results. Hybrid allocation scheme favors IBFD for low interference and OBFD for high interference settings. For OBFD scheme, access spectrum allocation seems to be following the same pattern for lower rank (i.e. closer) SBSs as well as the higher rank (i.e. farther) SBSs, but the backhaul spectrum required is higher for higher rank SBSs due to weak backhaul signal strength. Conversely, IBFD case lower rank SBSs are limited by backhaul interference. This starts to improve with higher rank SBSs, which require higher access rates i.e. more access spectrum for higher performance. It is shown that hybrid scheme outperforms both individual schemes by prioritizing IBFD for low interference and OBFD for high interference

scenarios.

In [53], Shariat *et al.* have decomposed the overall resource allocation problem between backhaul and access links to power and sub-channel allocation. The optimization is first done between one small cell and a UE, and then were generalized into multiple UEs. A novel algorithm is proposed for rate balancing that employed small cell grouping and resource slicing. The evaluation shows that joint optimization with rate balancing provides significant improvements for resource allocation between access and backhaul links. This method works not only for IAB systems, but it also shows significant improvements for fixed backhaul partitioning systems.

Khodmi, Benrejeb and Choukair develop a joint power allocation and relay selection algorithm for heterogeneous ultra-dense networks [54]. Khodmi *et al.* use the increment water-filling algorithm for power allocation and a non-linear programming algorithm for resource allocation. Simulations show that multi-hop relay improves the coverage and achieves the data rate requirements of cell-edge users compared to no relay or single-hop relay architectures.

In [55], Hao *et al.* have developed an energy-efficient resource allocation algorithm for two-tiered networks that use mMIMO on the MBS and OFDMA-based cellular frequencies on the small cells. A hybrid analog-digital precoding scheme at MBS is proposed. From this scheme, a problem for power optimization and subchannel allocation is formulated to maximize the energy efficiency of the HetNet subject to users' QoS and limited wireless backhaul. The formulated problem is a MINLFP problem, and the authors made use of transformations, namely MINLFP to DCP, and then DCP into convex optimization by appropriate approximation, to solve it iteratively. In simulations, the proposed algorithm is used with two structures: fully connected structure and subarray structure, for beamforming RF chains. If subchannel allocation is not considered, then the hybrid precoding scheme with subarray structure performs the best in energy efficiency, while having the lowest throughput of three. The energy efficiency of hybrid precoding with fully connected structure is negatively correlated with the number of RF chains, whereas its throughput is positively correlated. The results are similar when subchannel allocation is considered.

Mahmood *et al.* propose the adaptive capacity and frequency optimization method

for wireless backhaul networks [56]. Conventional network planning and its shortcomings are mentioned in this work. Mahmood *et al.* generate synthetic time series data based on real network data. This time series data is then used to train autoregressive integrated moving average (ARIMA) and multi-layer perceptron (MLP) models. These models are then used in the automated planning algorithm. Results show that MLP models especially have very low mean absolute percentage error (MAPE) and all models have performed below the error threshold of 10. The proposed methods are shown to achieve better capacity planning and optimization as well as reduce resource wastage.

In [57], Han and Ansari have developed a system for user association and power consumption on two-tiered HetNets with green energy and on-grid power usage. The proposed algorithm works in a distributed fashion and optimizes the use of green energy and the traffic delivery latency of the network. The system tries to make use of green energy wherever possible over on-grid power. Moreover, when forced to use on-grid power, the system balances the use of on-grid power and the traffic delivery latency. The traffic arrival process is modeled to be a Poisson process; therefore, base stations realize a M/G/1-PS (processor sharing) queue. The algorithm, which the authors have named GALA, is distributed to BS and users. The BS-side algorithm measures the traffic load and updates the advertising traffic load. In contrast, the user-side algorithm selects the optimal BS based on the advertised traffic load, BS's energy-latency coefficient and BS's green traffic capacity. The developed algorithm is compared with maximum rate algorithm and α-distributed algorithm, which consists of several optimization policies for user-BS associations to balance the flow level traffic among BSs. The GALA algorithm is shown to outperform both of the algorithms significantly in terms of power saving at the cost of negligible latency increases compared to the α-distributed algorithm.

Ma *et al.* investigate the user association and resource allocation problem of MIMO-enabled HetNets [58]. The network in question is a two-tiered HetNet consisting of a single mMIMO MBS and multiple small cells. Ma *et al.* formulate the problem of maximizing the α-fairness network utility via user association and resource allocation. The optimization problem is formed as a mixed-integer nonlinear programming problem. Then, this problem is solved using an algorithm based on the

Lagrangian dual decomposition method. Simulations show that the proposed method significantly outperforms the benchmark methods such as optimal UA and equal RA (where all UEs associated with a BS use the same resources) or max-SINR UA and optimal RA (UA is done solely by checking max-SINR), while falling short of the global optimum only by a very small margin.

Huskov *et al.* propose a smart backhaul architecture that adapts itself according to given task [59]. The system can adapt itself to certain equipment conditions for different equipment types defined previously. This development aims to save power by reducing the amount of unused resources, while keeping network at a high performance regardless of the load. The proposed system also makes use of a concept called multiple-input distributed-output (MIDO). MIDO aims to serve multiple devices on the same spectrum band providing interference alignment. Simulations show that the usage of MIDO with the proposed system outperforms standard MIMO systems 2.5 and 30 times for 2x2 and 8x8 configurations, respectively. However, Huskov *et al.* also highlight the fact that employing higher number of antennas at receiver side is unrealistic.

Zhang *et al.* design a two step resource allocation scheme for mmWave enabled IAB networks [60]. User association and transmit power allocation problem is iteratively solved. User association problem is solved using a distributed framework based on the many-to-many matching game. For this purpose, Zhang *et al.* design a novel backhaul capacity and interference aware matching utility function considering the interference penalty and backhaul capability simultaneously. For power allocation, convex approximation is used. This problem is then solved iteratively to converge to a KKT point. Simulations show that the proposed algorithm outperforms the random matching algorithm by up to 71.9% when the number of SBSs is 10.

Pu *et al.* formulate a resource allocation problem for mmWave self-backhauling networks [61]. The problem is modeled as a combinatorial integer programming problem. By introducing penalty factors, Pu *et al.* transform the problem into another equal problem which can then be solved by the Markov approximation method. The proposed algorithm maximizes the sum data rate and satisfies the minimum traffic demands of all users while satisfying the backhaul constraints of the individual SBSs.

Simulation results show that the achieved data rates in each UE is slightly bigger than the required data rate, and repeating BH traffic is minimized with the proposed algorithm. In terms of spectral efficiency, the proposed algorithm outperforms the benchmark methods of OBFD and access-backhaul split significantly.

Dai *et al.* tackle sum-rate maximization problem via joint user association and power allocation in H-CRANs [62]. Dai *et al.* solve the user association problem first in a central fashion. The power allocation problem is formulated as a generalized Nash equilibrium problem and then solved distributively using variational inequality theory. Simulations show that the proposed algorithm has considerably lower computational complexity and signaling overhead than the centralized max sum-rate method.

Xu *et al.* give a detailed survey of resource allocation problem of 5G HetNets [63]. Xu *et al.* first define different types of HetNets, and the resource allocation problems related with these types. For traditional HetNets, Xu *et al.* highlight transmit power, user association, and bandwidth allocation optimization as general problems. For OFDMA-based HetNets, subchannel allocation is another problem due to multiple orthogonal subcarrier usage. For NOMA-based HetNets, cross-tier and co-channel interference are problems to be considered alongside user fairness. For relay-based HetNets, relay selection and transmit power allocation are recurring problems. For H-CRANs, inter-cell interference, signal overhead reduction and baseband unit management are the problems to be tackled. For multi-antenna HetNets, interference cancellation and beamforming design are architecture-specific problems. Integration of new techniques and technologies such as IRS and ultra-dense networks require careful consideration for resource allocation problems to attain the best performance possible.

2.5.1.2 Lessons Learned

While resource allocation is by no means a problem unique to wireless backhauling, effective distribution of the resources is crucial for a performing network. With this in mind, we highlight the following points:

- Access/backhaul resource split is very important in terms of effective resource

usage, because employing a dynamic split makes the network adaptable to different scenarios. Employing dynamic and fast algorithms that can adapt to changing conditions is very useful for ultra-dense networks with wireless backhauling.

- Ultra-dense networking employs many small cells that a UE can connect to, meaning that at any time instant, multiple BSs are available for UEs. While this increases the network adaptability and redundancy, user association is important for network performance as the small cells usually do not have good range and require LoS, finding a suitable BS is important. From the network's perspective, the network can employ suboptimal user associations to keep up the average performance at a certain level or achieve a certain minimum.

- Transmit power allocation is important for the wirelessly backhauled networks as high air interface usage may result in inter- and intra-tier interference. Employing spatial multiplexing with the right transmit power allocation effectively nullifies the interference from the UE side, while BSs employ beamforming in tandem with power allocation to reduce the effect of interference.

- With the usage of MEC in 5G networks, computation offloading is a possibility that can result in power savings and effective computation. However, while wireless backhaul can allow offloading computation-heavy tasks, there may be cases in which offloading can induce performance penalties to the air interface, resulting in power wastage from the network side. Consideration of air interface is required for computation offloading decisions in wirelessly backhauled networks.

2.5.1.3 Open Issues and Challenges

Resource management remains an open challenge for future networks. 3D networking is a new concept that significantly changes the network design. Because of this, resource allocation schemes also require a new approach for 3D networking for the best performance. Spectrum and interference management also remains an important open issue for future networks [64].

With a myriad of new concepts being introduced in 5G networks and beyond, networks are inevitably getting more and more complex. While new concepts come with performance gains, they also introduce complexity to the design of a network. In terms of resource allocation, using multiple technologies increases the complexity of the problem. New models are required in this direction to allow better design; which, in turn, will allow better problem formulations that capture the requirements and result in better performance gains [63].

Spectrum sensing is another way to increase resource efficiency. The tradeoff between dynamic resource allocation and spectrum sensing capabilities remain an open research area that can result in better spectrum and resource efficiency.

The majority of the mentioned works do not use mmWave frequencies. Consequently, most mentioned works cannot meet the KPI demands of 5G communications in terms of throughput. While beamforming and mMIMO are technologies that can enable better performance with mmWave usage, resource allocation problems have to consider beamforming and mMIMO usage alongside the main problem, which adds more degrees of freedom and makes the problem even harder. Due to these points, mmWave frequencies and the resource allocation schemes require a more in-depth exploration.

Machine learning-based resource allocation is also promising since it allows autonomous management of networks. BSs can capture an enormous amount of data that can be used for training a model for user association, transmit power calculation, or spectrum allocation. While these are open research areas, the training process is a problem since there aren't enough resources for training in one BS. Therefore, either distributed algorithms are required for this task, or a new paradigm must be developed in order to realize machine-learning based dynamic resource allocation.

2.5.2 Deployment

Base station deployment is a problem not specific for wireless backhauling. Nonetheless, deployment plays a critical role when wireless backhaul is employed. Here, the challenge is where to place MBSs and small cells. Wireless backhaul enables more radical deployment strategies since infrastructure (or the lack of it) has less of an im-

pact compared to traditional backhaul. Similar to the traditional networks though, deployment directly impacts network performance. Therefore, innovative deployment schemes have to be employed to make wireless backhauling feasible.

Mesh, star and tree topologies are usually considered while deploying networks. Tree topologies are especially used with hierarchical networks, as well as with wireless backhauled networks. In such schemes, root node of the tree has fiber backhaul capabilities, and all other nodes send their data via their parent to the root node to be backhauled to network core. Mesh topology is also suitable to be used with wireless backhauling as it promotes redundancy via multiple paths. Finally, wireless backhauling also allows dynamic deployment with mobile base stations, which improves network's capabilities while making management and deployment harder.

While the deployment of small cells is a task performed before the actual deployment, moving elements such as UAVs and nomadic BS make this problem dynamic. Because of this, finding the optimal deployment in a fast manner becomes important as this increases the network adaptability and reduces the time that the network is affected by the condition changes in a bad way.

2.5.2.1 State of the Art

In [34], Bonfante *et al.* propose a two-layer network scheme where self-backhauling small cells serve UEs and MBSs using mMIMO backhaul their data. Small cells use fixed allocation of access and backhaul slots. Two different deployment scenarios are considered for small cells. In the first scenario, the small cells are randomly and uniformly distributed to the BS area; whereas in the second one, which the authors named as ad-hoc deployment, they are deliberately placed closer to UEs to increase the throughput. To reduce the inter-cell interference, small cells in ad-hoc deployment use a more directive antenna that serve only the closest UEs, thus reducing interference. Results show that the performance increases threefold when ad-hoc deployment is used compared to uniform deployment.

McMenamy *et al.* investigate mmWave backhaul network flow affected by the number of hops and deployment of fiber-connected BSs [65]. The authors first determine

which nodes will have fiber backhauling given the deployment while satisfying the constraint on the number of hops. Then, the authors maximize the overall flow while meeting the flow demand of every node and the constraint on the number of RF chains for every node in the network. Simulation results show that increasing the number of hops reduces the number of required fiber-connected BSs (BGWs), but also the total flow. Number of RF chains grows with the number of hops. If the number of BGWs are reduced, then the performance can be maintained by increasing the number of RF chains at the expense of power consumption.

Aftab *et al.* address UAV-BS placement problem with a novel machine learning based intelligent deployment mechanism [29]. An intelligent long short-term memory (LSTM) model is used to predict the deployment with a feedback mechanism for self-correcting. Simulation results show that deep learning-based approach outperforms the normal deployment in terms of outage probability. ML-based deployment also achieves far better energy efficiency compared to the normal deployment.

In [66], Fouda *et al.* have considered a scenario where UAV are deployed to a highly congested area where some UEs suffer low quality and some have no service at all. The network in question is a two-tiered network with gNB's that serve UEs in tandem with UAV small cells. The network uses in-band frequency division duplexing (FDD) as the operation mode. The UAVs' mobile capabilities are used as the main degree of freedom to maximize the overall system sum-rate. The authors have formulated an optimization problem to find the UAVs' optimal 3D deployment locations, UE power allocations, precoder design at gNB and UE-base station associations. Since these optimization parameters have different update intervals, the problem is decomposed into two parts: in the first part, average sum-rate is maximized by optimizing the UAV locations, access and backhaul link power allocations; whereas in the second part, precoder design and transmit power allocation is optimized to maximize the instantaneous sum-rate. Simulations show that UAV deployment increases the average sum-rate at the expense of performance of UEs that are close to gNBs. It is also shown that UAVs cannot support a high number of UEs as efficiently and distributing users between UAVs is key to a high-performance network.

Park, Tun and Hong propose initial deployment and trajectory optimization tech-

niques for stable communication between high speed train (HST) and UAVs [30]. Park *et al.* use Soft Actor-Critic (SAC) method of reinforcement learning for optimizing the UAV trajectory and Support Vector Machine (SVM) for optimal initial UAV deployment. Numerical results show that with less HST speed, less handovers occur and UAVs can have lower operating altitudes with higher data rates. Conversely, with faster HST speeds, UAVs have to increase their altitudes to keep the connection up.

Zhang *et al.* formulate a joint optimization problem of UAV deployment, caching placement and user association for maximizing the QoE of users [67]. The cache-enabling UAVs are deployed at peak hours to the hotspots to alleviate the pressure on the MBSs. The formulated optimization problem is divided into three subproblems. UAV deployment subproblem is solved via a swap matching based algorithm. Then, Zhang *et al.* solve the caching placement subproblem with a greedy algorithm, and user association problem with a Lagrange dual. The joint problem is solved with a novel low complexity iterative algorithm. Simulation results show that the proposed algorithm can reach the near optimal value of the exhaustive search within 0.02 mean opinion score (MOS). The proposed algorithm outperforms the random and classic algorithms in terms of UAV backhaul traffic offloading ratio, and average MOS, under varying user numbers, cache space, and UAV height.

Okumura and Hirata propose an algorithm to automatically plan the deployment of 300 GHz wireless backhaul links [68]. The algorithm first selects a pair of adjacent cells that can get a LoS connection. After deploying a master BS to the center rooftop, the algorithm deploys the slave BSs according the the previously mentioned pairs of adjacent cells to find a LoS path from every slave BS to the master BS. Simulation results show that this deployment scheme exceeds the standart demand of 100 Gbit/s even when the rain rate is 100 mm/hr.

In [69], Polese *et al.* have explained the integrated access and backhaul concept and explored some scenarios for its real-life usage. A thorough explanation of IAB is made, its architecture, network procedures, topology management, scheduling and resource multiplexing aspects are detailed. Then, an evaluation of an IAB mmWave network is made through simulations with different configurations. For routing the backhaul traffic, wired-first and highest-quality-first approaches are tested and wired-

first approach is found to be better in terms of throughput and latency. For IAB deployment scenarios, three alternatives are compared: namely, only donors scenario in which only some wired gNBs exist, IAB scenario where alongside the first scenario donors, some IAB nodes are available, and all wired scenario where it is assumed that the topology is the same with the second scenario but all backhaul connections are wired. All wired scenario outperforms the others as expected, but IAB scenario outperforms the donors-only scenario in average delays, 5th percentile throughput and average buffering in target UEs.

Hu *et al.* propose an intelligent UAV-BS deployment scheme based on machine learning [31]. Hybrid ARIMA-XGBoost model is used for predictions: ARIMA handles linear predictions and XGBoost is then applied on the residue of ARIMA. Simulation results show that the proposed hybrid prediction model has a prediction accuracy of 87% in the worst case. The proposed deployment scheme also significantly reduces the blocking ratio, increases the capacity, throughput, and resource utilization of the network.

In [70], Dahrouj *et al.* have proposed a hybrid system that uses RF as well as FSO links for wireless backhauling. The hybrid usage combines the low-cost, high-availability and high-coverage nature of RF and high-capacity, low-latency nature of FSO, bringing the best of both worlds. To reduce complexity, the problem is first relaxed such that only optical fiber links are allowed. Using the found planning, the found nodes are connected with fiber or hybrid RF/FSO links. To improve reliability, the planning is made so that the network is K-link disjoint. This is done by clustering the BSs together and iteratively finding the links between clusters. The authors show that the solution is equivalent to a maximum weight clique in the planning graph. The results indicate that as K increases, the network design is cheaper than fiber-only while maintaining the required reliability and key performance indicators (KPIs).

Liu *et al.* investigate a fast deployment strategy of UAVs using deep learning [32]. Liu *et al.* formulate a problem of finding the optimal UAV-BS position as fast as possible, while maximizing the sum of downlink rates in the network. The authors solve this problem by designing a geographical position information learning algorithm, which employs a dueling deep Q-network. Simulation results show that the

proposed algorithm finds the deployment position in seconds with little loss. While reinforcement learning requires training after any changes in the scenario, the proposed algorithm does not have such a requirement and achieves 96% of the optimal value. For different area sizes and UE densities, the proposed algorithm performs almost indistinguishably from reinforcement learning but with less time.

Zhang *et al.* optimize the BS and gateway densities of an mmWave backhaul network to improve network spatial throughput [71]. Low density makes the network backhaul-limited, but after a certain density, the performance gains disappear because the network becomes interference-limited due to the intra-tier interference of the newly added elements. The optimal densities are found numerically and the simulation results show that with the increasing network density, the spatial throughput first increases and then converges to a gateway density-related constant.

Kulkarni, Ghosh and Andrews [72] propose a network with k-ring deployment with wireless backhauling. In k-ring deployment, a BS with fiber connection is on the centre while all other BSs are backhauled wirelessly to this BS, and then to the core. Kulkarni *et al.* argue that with their design, instantaneous rate performance is independent of the scheduling, which makes the system noise-limited. The authors formulate an optimization problem for a fixed k that maximizes the end-to-end rate achieved by all UEs. For this purpose, Kulkarni *et al.* define the max-min rate as the maximum rate that can be achieved by the worst member of a routing strategy. The analysis is done initally for IAB, and then extended to orthogonal access and backhaul (OAB), which basically means different resources for access and backhaul. For IAB case, Kulkarni *et al.* employ highway routing, which designates the x and y axes that are centered on the fiber backhauled BS as the highways. The data is then distributed from these highways, even if this results in more hops than a direct path. If BSs on the highways have larger antenna gains than others, this scheme is justified. Nearest neighbor routing is used as the benchmark scheme. IAB solution requires the global load and access rate information, which makes its solution more complex. OAB is simpler to implement since it does not require such information. Kulkarni *et al.* present results in four main fields. First, it is shown that with higher number of rings, the performance drops significantly since the network is backhaul-limited. Second comparison is made between full-duplex and half-duplex modes. Full-duplexing

shows significant potential in terms of data rate if the self-interference is kept below -100 dB. However, Kulkarni *et al.* also note that these results do not seem practical. Third comparison is made between OAB and IAB. It is shown that OAB can slightly outperform IAB if the access/backhaul resource assignment is made according to the load. If this assignment is made in a fixed fashion regardless of the expected load, then IAB is the better choice. Finally, dual and single connectivity are compared in terms of performance. It is shown that, performance gains with dual connectivity is higher if there are load imbalances in single connectivity cases.

Zhang and Ansari investigate the backhaul-aware uplink communications in a full-duplex drone BS-aided HetNet problem with the objective to maximize the total throughput while minimizing the number of deployed BSs [73]. The authors decompose the problem into three parts, namely (i) joint user association, power, and bandwidth problem, (ii) drone BS placement problem, (iii) determining the number of drone BSs to be deployed. The first problem is solved using an approximation algorithm. The second problem is solved using exhaustive search, while the third problem is solved using linear programming. Simulation results show that the blocking ratio of the proposed algorithm outperforms the benchmark methods in terms of total throughput and the blocking ratio.

Saadat, Chen and Jiang consider multi-hop backhauling for ultra-dense networks [74]. The authors propose a deployment scheme that allows multiple backhaul paths to MBSs, helping overcome the blockage problem to some extent and improve overall reliability. The proposed method for backhaul is called the multipath multi-hop (MPMH) backhaul. Saadat *et al.* assume that the small cell density is much higher than the user density, and each SBS only serves one user. User-serving SBSs relay their backhaul traffic via the closest SBS to the MBS. The backhaul routing is based on minimizing the number of hops. The proposed MPMH backhaul is compared with direct backhauling, multihop backhauling and multiple-association backhauling that employs multiple paths between sender and receiver. Simulation results show that MPMH backhaul outperforms all other alternatives regardless of the inter-site distance. It is also shown that multi-hop scheme significantly improves LoS probability. Finally, as the inter-site distance increases, multipath scheme shows significant gains over multi-hop scheme. Needless to say, MPMH backhaul incorporates the advan-

tages of both schemes together.

Nasr and Fahmy analyze the scalability of ring and star topologies in mmWave backhaul networks [75]. The network in question has one MBS with fiber connectivity and multiple small cells with star or ring topology. Mesh topology uses multiple hops to relay traffic to the MBS, whereas in star topology, every small cell has direct connection to the MBS. Simulation results show that initially, star topology performs better but as transmit SNR increases, mesh topology performance increases faster. Similar results can be observed for different small cell sizes as well. Increasing small cell numbers almost linearly increases the performance for both topologies.

2.5.2.2 Lessons Learned

We highlight the following points for the deployment problem:

- UAVs are used for different purposes, such as content caching and performance improvement. UAV deployment directly affects the performance gains from such use cases. It is shown that with correct deployment, UAVs can be used as small cells with better freedom and adaptability.

- With the new ultra-dense network architectures, hierarchical designs seem to be advantageous. Using tree structures allow better flow designs for backhauling, while mesh structures allow better flexibility and redundancy by employing multiple links for backhauling. Mesh structures employ every small cell as both an access node and a backhaul node for other small cells, whereas tree structures make use of small cells that perform backhauling most of the time, especially for nodes that have higher heights in the tree structure.

- Machine learning approaches are widely employed with UAV deployment problems, and with solid results. Machine learning approaches are also shown to be quite fast, which is a significant step to realize a self-organizing network with UAVs. Compared to the optimization approaches using linear programming and similar formulations, machine learning approaches give flexibility to the network to adapt to different conditions rapidly.

2.5.2.3 Open Issues and Challenges

Current research on UAVs do not make use of mmWave frequencies. In the future, UAVs can use mmWave frequencies for better performance. The deployment of mmWave-using UAVs is an open research problem that requires investigation. This problem has different facets such as vertical modeling, beam management, and antenna design.

While UAVs and other mobile cell options provide flexibility for 5G networks, current approaches are still not fast enough to be used for self-organizing networks on a grand scale. Current formulations are often optimization problems that require considerable time for solving. Consequently, these solutions do not use the mobility capabilities to the fullest extent. Faster and more adaptable approaches are required for self-organizing networks and this problem remains an open challenge. While some machine learning approaches are performing in this area, there is still room for improvement, especially for larger-scale self-organizing networks.

UAV and mobile cell mobility also presents problems for mesh and tree topologies, as the moving cells alter the topology and adaptation to these movements are required. This problem can be solved in a central or distributed fashion. While each seem to have their own advantages, distributed approaches seem to have an edge since the cells have a high volume of base stations. The adaptability of topologies is an open challenge that needs an effective solution for a moving network with high performance.

2.5.3 Coverage

In 5G, HetNets, and small cell deployments such as femto- or picocells are envisioned to be employed for expanding cell coverage. As we mentioned in previous sections, these technologies employ wireless backhauling and their performance is directly linked with the wireless backhaul quality. Because of this, coverage enhancements via wireless backhaul are frequently analyzed in the literature.

2.5.3.1 State of the Art

Lukowa *et al.* evaluate the performance of IAB with central scheduling [76]. The authors consider IAB with mmWave for coverage extension. Lukowa *et al.* also propose two different TDD schemes, namely Split TDD, and Flexible TDD. In Split TDD, data flows are split in the time domain for all cells, which avoids downlink-to-uplink interference. In Flexible TDD, any BS can be scheduled for downlink or uplink flow in any time slot, allowing effective multiplexing of downlink and uplink slots. Beamforming is used to cancel out the interference and increase the coverage. Power allocation for backhaul links are also controlled to protect the uplink access from strong backhaul interference. Results show that while Split TDD achieves high downlink throughput, Flexible TDD outperforms it even though interference is arguably higher. The deployment of relay nodes also provides significant throughput gains in the 5^{th} percentile users for both downlink and uplink.

Qu, Li and Zhao study the coverage problem in device-to-device relay networks [77]. Qu *et al.* formulate the coverage problem to maximize the overall system downlink rate. The proposed problem is transformed into a 0-1 integer programming problem and the authors propose a novel algorithm that uses a greedy approach to find the optimal solution. Simulation results show that the relay node covers a larger area as it moves outward from the cell center. This is because the required bandwidth for device-to-device communication gets less and the allocated resources can serve more UEs. As the relay location grows, spectral efficiency goes up as the relay covers more UEs, but then goes down as the backhaul link deterioration overpowers the coverage increase. The proposed scheme is also shown to outperform the existing SINR-based scheme.

Singh and Singh use full-duplex amplify and forward relays to enhance the coverage probability and the transmission capacity of D2D enabled cellular networks [78]. The authors derive closed form expressions for coverage probability and transmission capacity as a function of D2D user density and SIR threshold. Singh *et al.* also analyze the effect of D2D user density, relay node density and D2D link distance on the coverage probability and transmission capacity. Numerical results show that for cellular users, coverage probability is improved with relay nodes as they enhance

the SIR at receiver. For lower SIR threshold, coverage probability is higher and vice versa. For D2D users, same phenomenon can be observed albeit with higher coverage probabilities than the cellular users in all cases. The transmission capacity also increases with the introduction of relay nodes for both cellular and D2D users, whereas it decreases with higher SIR thresholds. The transmission capacity initially increases for higher relay node densities, but after a certain density value, any increase results in a decrease in the transmission capacity since the gains are overpowered by the interference introduced with the high relay density.

In [79], Sharma, Ganti and Milleth have analyzed the coverage of a two-tier network, in which small cells are connected to the MBSs wirelessly and use in-band full-duplex communication. MBSs also serve the UEs, meaning that they have to manage the resource split for access and backhaul. Coverage formulation is made for both IBFD and FDD for both macro and small cells. These formulations are then compared in simulations, which are shown to be on-par with the numerical results found. FDD is shown to have a better coverage probability than IBFD for different configurations of small cell signal-to-interference thresholds and small cells per macro cell ratios. The authors have also introduced bias to the system to connect to small cells even if MBS gives better SIR values. For high bias values, coverage as well as average rate significantly decreases. IBFD suffers from inter-tier interference and low coverage but provides higher rates.

Khan *et al.* use UAVs to outsource network coverage according to a desired QoS requirement [80]. In the proposed system, UAVs belong to the operators and the operators may use other operators' UAVs to provide coverage. The authors propose a reputation-based auction mechanism to model the interaction between the outsourcing operators and the serving UAVs. To enforce the service level agreement, Khan *et al.* propose a blockchain-based system using support vector machines (SVM) for real-time autonomous and distributed monitoring of the UAV service. Simulation results show that increasing the number of monitoring nodes also increases the probability of true classification and reduces the false alarm probability. Ergodic capacity is shown to be positively correlated with the UAV cost for deviation. Furthermore, if the UAV moves towards the area that requires coverage, service quality increases but this also increases the cost since the old service area gets worse service.

Kamel, Hamouda and Youssef investigate the uplink coverage and ergodic capacity of mMTC considering an ultra-dense network environment [81]. Kamel *et al.* model the small-scale fading using a general $\alpha - \mu$ channel model and the Rayleigh fading, and consider the direct MTC access mode where MTC nodes directly connect to small cells. Analysis on the proposed system model reveals that the impact of high density of small cells in UDNs relaxes the requirements on the maximum transmit power, which in turn reduces the complexity, and increases the battery life and cost savings. Simulation results show that the uplink coverage is significantly high at relatively low coverage thresholds. 80% of the time, MTC nodes experience an SINR of at most 0 dB. As the number of MTC nodes increase, the coverage performance declines since the limited resources are allocated to a higher number of nodes. Higher power truncation threshold results in a significantly higher coverage performance. As the power truncation threshold gets smaller, the effect of MTC node density on the coverage performance also decreases. This is because the network becomes noise-limited at low power truncation threshold. Increasing the small cell density results in a significant coverage performance gain, but after a certain density, the coverage probability reaches a maximum. Increasing the MTC node density at low small cell density results in a sharp decrease in coverage probability, but this decrease becomes milder at higher small cell densities.

Zaidi *et al.* propose a network where UAVs are used as decode-and-forward relays to extend the coverage [82]. NOMA is applied on two different set of users, namely the ground users and the UAVs. UAVs are used to forward the data to cell-edge users. Simulation show that the outage probability is reduced with increasing SNR and reducing the value of threshold data rate. However, while increasing the transmit power of the UAVs reduces the outage probability, this is not possible as UAVs are power-limited. Increasing the altitude of UAVs also increases the outage probability. Finally, it is shown that for the same parameters, NOMA outperforms OMA systems by a significant margin.

Jo *et al.* employ a joint method of "iteration and convex optimization based on power control (PC)" and "advanced range expansion (RE) technique" to overcome the trade-off between the throughput maximization and coverage optimization in 5G ultra-dense HetNets [83]. The authors derive the transmission power of optimal cov-

erage in terms of minimal handover failure rates. Jo *et al.* try to minimize the sum of handovers occuring too early or too late. Too early handover is defined as when a handover occurs due to higher signal power from a target BS but actually the received signal is enough to keep up the connection alive. Too late handover is defined as when handover never occurs until the receiver power from serving BS decreases below a certain threshold that is not high enough for connection. Jo *et al.* propose an iterative power allocation algorithm for the derived power problem. Numerical results show that the proposed algorithm shows 2-5% better fairness performance in idle mode coverage compared to the benchmark methods. Moreover, RLF rate is shown to converge to zero for the proposed algorithm whereas benchmark methods do not converge. The proposed method also achieves almost the same result as the power control algorithm that only focuses on maximizing the throughput.

Shokry *et al.* use UAVs as a cell-free network to provide coverage to vehicles on a highway with no infrastructure [15]. Shokry *et al.* formulate the UAV trajectory decisions as a Markov decision process and use deep reinforcement learning to learn the optimal trajectories of the deployed UAVs to maximize the vehicular coverage. Deep deterministic policy gradient is adapted to solve the continuous control task. Simulations show that the proposed method consistently outperforms random dispatching, fixed dispatching and hovering methods. The proposed method provides the desired coverage with fewer UAVs compared to other methods. The authors also show that the proposed method inherently reduces the energy consumption as well since adding an energy penalty results in a reduction of only 16% in energy consumption.

Khoshkholgh *et al.* [84] propose a large-scale aerial-terrestrial HetNet and specifically focus on its coverage performance. The authors evaluate the coverage probability as the complementary cumulative distribution function of the SIR ratio. From there, Khoshkholgh *et al.* conduct Monte Carlo simulations. Results indicate that increasing the percentage of aerial BSs do not always result in a coverage increase. Moreover, high-rise environments can support a higher aerial-to-ground BS ratio since high-density blockages in such environments result in the dominance of NLOS interference. Furthermore, the authors also report that due to the aggressive interference in UAV communication, densifying the network is especially destructive, regardless of the environment. However, this effect can be mitigated by replacing some

ground BSs while introducing aerial BSs. Finally, adjusting the height at which aerial BSs reside can increase the coverage probability, especially at sub-urban environments in which the overall performance is lowest.

Azimi-Abarghouyi *et al.* [85] perform coverage analysis for finite wireless networks. The authors model the network using stochastic geometry and Poisson point processes. The given setup is highlighted as being useful for mmWave communications, indoor, and ad hoc networks. The coverage probability is mathematically derived with a lower bound as well. Numerical results indicate that there is an optimal distance of the user in terms of coverage probability. The tightness of the bounds are validated with the numerical results. Finally, it is shown that coverage probability is improved when the path loss exponent is larger.

Siddique, Tabassum and Hossain analyze the performance of spectrum allocation schemes for IBFD and OBFD backhauling [52] and the coverage rate for these schemes. The authors propose two schemes, namely maximum received signal power (max-RSP) and minimum received signal power (min-RSP). It is shown that IBFD and OBFD schemes work well in different environments in terms of coverage. For min-RSP, SBSs that are farther from the MBS have higher coverage rates while closer ones have lower coverage rates. Conversely, for max-RSP, the closer SBSs have higher coverage rates. In terms of duplexing, IBFD scheme has better overall coverage rate than OBFD scheme with both algorithms. In max-RSP case, OBFD backhauling outperforms IBFD for the farthest SBSs only. Whereas in min-RSP case, OBFD backhauling outperforms IBFD only for the closest SBSs.

2.5.3.2 Lessons Learned

We highlight the following points for network coverage with wireless backhauling:

- In terms of coverage, small cell usage is the best solution to cell-edge performance issues to date. Employing SBSs to extend the coverage is a solid approach that will be employed frequently in future networks.
- While mmWave communications have less range than the traditional sub-6 GHz frequencies, ultra-dense networking is made possible with mmWave us-

age. In contemporary networks, deploying new BSs to improve coverage often bring their own problems such as inter-cell interference. Using mmWave frequencies with beamforming and mMIMO mitigates most of these problems, and allows new deployments to extend network coverage.

- In terms of coverage, machine learning usage is very promising for future networks for a multitude of tasks such as mitigation of coverage holes or interference. While the performance of these ML algorithms is solid in terms of network KPIs, more research on such ML algorithms are required to make these algorithms perform faster, allowing better response times to changing network conditions.

2.5.3.3 Open Issues and Challenges

Coverage analysis is a tool that researchers have frequently used to show the performance gains of small cell usage. Moreover, the chosen papers act as proof that various concepts such as IBFD or aerial-terrestrial networks perform well. Though coverage analysis is a general problem of networking, it is nevertheless important to show that the performance of newer concepts are adequate.

Coverage enhancements are often made at the expense of capacity or throughput. While coverage enhancements are very important for ubiquitous connectivity targets for future cellular networks, these enhancements need to be able to satisfy the demands of 5G networks. This is also due to the inherent characteristics of the frequencies used for wide coverage; lower frequencies have better range but worse capacity. This tradeoff is still an open research problem and its solution will directly allow the operators to increase the coverage as well as the performance with the existing hardware and infrastructure.

Network deployment is an area closely linked with coverage. Massive coverage gains can be made at the table with the correct network deployment designs. Small cells are at the heart of this phenomenon, and the points that we mentioned for network deployment also hold for coverage to some extent.

Although we will mention the rural connectivity in a later section, bringing ubiq-

uitous coverage while meeting the demands of 5G use-cases is a very challenging task. Doing this in a rural context is even harder. While wireless backhauling makes ubiquitous connectivity possible, there are challenges specific to rural connectivity in terms of both coverage and performance. Therefore, ubiquitous coverage is still an open challenge.

Satellites, UAVs, and HAPSs are also other network elements whose use are made feasible with wireless backhauling. Currently, their main use-case is ubiquitous connectivity and expanding the coverage of existing networks. Consequently, the challenges related to satellites, UAVs, and HAPSs are also in some sense open challenges related to coverage.

Another point with wireless backhauling is that the traditional coverage models are often 2D models. With mmWave usage and integration of UAVs, HAPSs, and satellites, these models will no longer be accurate. Because of this, new 3D models for coverage are required for efficient network designing [86].

Infrastructure sharing is another concept that can result in enhancement of coverage quality and availability [86]. This concept not only reduces deployment and management costs, but it also aids in coverage/capacity balance. Network slicing and other technologies that can aid in infrastructure sharing in an effective manner should be investigated in the future.

CHAPTER 3

USER ASSOCIATION AND ROUTING PROBLEM IN HETNETS WITH UAVS AS SMALL CELLS (UAV-UAR)

In this chapter, the problem of user association and routing in HetNets with UAV-SCs (UAV-UAR) is introduced. First, the system model for IAB HetNet with UAV-SCs is introduced. The developed model can also be presented as a flow network. This flow network presentation is then used to formulate the UAV-UAR problem. Finally, using the aforementioned formulation, a mixed-integer linear programming (MILP) problem is proposed to solve the UAV-UAR problem.

3.1 The UAV-UAR Problem Definition

Consider a cellular network with UEs, MBSs and UAV-SCs. The network uses mmWave frequencies to support high data rates. However, mmWave frequencies have considerably lower ranges than sub-6 GHz frequencies due to high path loss and sensitivity to blockage, foliage, and atmospheric attenuation. In the network model, if a UE is close enough to an MBS, it will have a strong link, and consequently, high capacity. This will not always be the case, and after some threshold distance, the link capacity will be even worse than that of sub-6 GHz frequencies. The link deterioration problem can be solved by increasing the network density via introducing more BSs. UAV-SCs come to the rescue by acting as BSs to serve UEs close to them.

While MBSs can send backhaul traffic back to core network easily via fiber links, UAV-SCs do not have such capabilities. To operate as BSs, UAV-SCs have to send their traffic to MBSs either directly, or via multiple hops using other UAV-SCs. In the network model, MBSs serve UEs and UAV-SCs by acting as their core network entry

point. UAV-SCs serve not only UEs but also potentially other UAV-SCs by providing them access to an MBS.

The rationale behind using UAV-SCs instead of MBSs is twofold. Firstly, MBSs are much costlier in terms of time and money than UAV-SCs due to their installation procedures. Secondly, for some scenarios (e.g. rural deployments), deploying MBSs can simply be impossible due to nonexistent power links, or terrain conditions. MBSs and UAV-SCs use mmWave frequencies to support high capacity connections. While for MBSs, the advantage is better UE performance, UAV-SCs also make use of significantly better backhaul links, which in turn increases their capabilities to serve UEs and give a service that is comparable to the MBSs in terms of performance.

All elements in the network use the same carrier frequencies for communication, integrating access and backhaul. This integration normally results in access-backhaul interference, but also promotes frequency and hardware reuse, which makes it desirable. MmWave links use massive MIMO antenna arrays and beamforming to realize point-to-point links. Directive nature of the links significantly mitigate the losses related to mmWave frequencies, while also making the networks noise-limited in nature instead of interference-limited. Due to beamforming, any interference or blockage simply breaks the link, but the probability of such an occurrence is negligible. Consequently, access-backhaul interference ceases to be a limitation for the proposed network.

In the proposed network, there is a high chance that a UE has multiple available links to different BSs. Choosing which of these available BSs serve the UE directly affects the network throughput. Since short-distance links will possibly be the ones that allow the highest capacity, choosing the link with the highest capacity might be a sound action. However, consider a case where the closest UAV-SC has the highest link strength, but it cannot backhaul the traffic to the core network as its available backhaul hops are fully used. In such a case, the limiting factor is not the actual link strength, but the available backhaul capacity for each BS.

To sum up, there are two problems related to the network throughput: the first one is to choose which BS to serve a UE, and the second is to find backhaul routes to send the user traffic from the UAV-SCs to MBSs. These problems will be the basis for

Table 3.1: Variables used in problem formulation are presented in this table.

$f[i,j]$	flow from vertex i to vertex j
$c(i,j)$	capacity of the edge from vertex i to vertex j
$a_{i,j}$	binary value indicating whether flow from vertex i to vertex j is used or not
d_i	throughput demand of i'th UE
S	number of UEs
A	number of UAVs
T	number of MBSs
$\mathcal{S}$	set of vertices representing UEs
$\mathcal{A}$	set of vertices representing UAVs
$\mathcal{T}$	set of vertices representing MBSs

UAV-UAR problem, which will be detailed in the next section.

3.2 MILP Formulation for UAV-UAR

In this section, we will first define the two problems mentioned in the last section with more detail. These detailed definitions will then form the basis for the UAV-UAR problem definition. The variables and parameters used for the problem formulation can be found in Table 3.1.

The first problem, namely choosing one of the available BS's to serve the UE, is a standard user association problem. A UE s_i has a throughput demand d_i. To satisfy this demand, s_i has to choose a BS $b \in \{\mathcal{A}, \mathcal{T}\}$, and the capacity of the link between s_i and b is $c_{s_i,b}$. If b is a MBS i.e. $b \in \{\mathcal{T}\}$, then the available capacity is the same as the link capacity. Otherwise, b is a UAV-SC (i.e. $b \in \{\mathcal{A}\}$), the link capacity provides an upper bound to the available capacity. This is because while MBSs can directly send the incoming traffic using fiber backhaul links, UAV-SC's backhaul is not guaranteed in this way, and uses wireless links. If the sum of capacities of available wireless links is smaller than the link capacity between the user and UAV-SC, then the backhaul capacity becomes the limiting factor.

The backhaul capacity brings us to the second problem, namely routing the traffic from UAV-SCs to MBSs. Unlike UEs, UAV-SCs can use multiple paths and multiple MBSs to backhaul their traffic. UAV-SC's are capable of concurrently serving users and sending backhaul traffic. For this purpose, UAV-SCs first use any single-hop links to MBSs. If these are not enough, then other UAV-SCs are queried for any available backhaul routes. Finally, the sum of mentioned links gives the backhaul capacity and how much user throughput can be supported. By routing traffic to the UAV-SCs that are closer to MBSs, backhaul capacity can be greatly increased. A good solution of routing problem will allow for a better capacity available to the users from UAV-SCs.

Combining the aforementioned two problems, we now define the UAV-UAR problem. Given a topology in the form of a graph $G = (V, E)$ where V is the set of network elements and E is the set of links between them, the objective is to find the *(UE,BS)* tuples representing which user is served by which BS and flows $f[i, j]$ between every network element that maximizes the total throughput achieved by UE's in the network. More formally:

Let $G = (V, E)$ be a flow network where $V = \{\mathcal{S}, \mathcal{A}, \mathcal{T}\}$ is a set of vertices and E is its edges. $\mathcal{S} = \{s_1, s_2, \ldots, s_S\}$ is the set of vertices representing the UEs, $\mathcal{A} = \{a_1, a_2, \ldots, a_A\}$ is the set of vertices representing the UAV-SCs and $\mathcal{T} = \{t_1, t_2, \ldots, t_T\}$ is the set of vertices representing the MBSs with fiber connectivity. Source vertices are the vertices in the set $\mathcal{S}$ and sink vertices are in the set $\mathcal{T}$. For the edges $(s, u) \in E$, the capacity $c(s, u)$ is a positive value that represents the link capacity in the cellular network. The joint user association and routing problem can be formulated as follows:

$$\max_{f,a} \sum_{j \in V} f[i,j] a_{i,j}, \forall i \in \mathcal{S} \tag{3.1}$$

$$\text{s.t.} \sum_{j \in V} a_{i,j} = 1, \forall i \in \mathcal{S} \tag{3.2}$$

$$\sum_{j \in V} f[k,j] = \sum_{j \in V} f[j,k], \forall k \in \mathcal{I} \tag{3.3}$$

$$f[j,l] \leq c(j,l), \forall j,l \in V \tag{3.4}$$

$$\sum_{j \in V} f[i,j] \leq d_i, \forall_i \in \mathcal{S} \tag{3.5}$$

$$\sum_{j \in V} f[i,j] \geq 0, \forall_i \in \mathcal{S} \tag{3.6}$$

Objective 3.1 aims to maximize the total flow going out of all vertices in $\mathcal{S}$, which are source vertices, i.e. UEs. The objective uses flow values in conjunction with $a_{i,j}$ values to enforce user association. Constraint 3.2 represents the user association in such a way that every UE can only use one of its available links to a BS. Constraint 3.3 represents the flow conservation quality for intermediate nodes, i.e. the incoming sum of flows to an intermediate node must be equal to the outgoing sum of flows. As intermediate nodes cannot accumulate flows in them, any incoming flow must be sent via edges that do not send any flows. Constraint 3.4 represents the flow capacity constraint, i.e. a flow cannot be assigned a value that is greater than its capacity. If users have an available flow that is greater than their demand, only the demand should be allocated, as denoted in Constraint 3.5. Finally, Constraint 3.6 states that UEs cannot have incoming flows.

Since the variables $a_{i,j}$ are binary and $f[i,j]$ are non-integer finite values, the defined problem can be formulated as a mixed-integer linear programming (MILP) problem. To give an idea about the search space, consider $S = 100, I = 20, T = 2$, in which case there is a total of 2662 links, assuming a topology where every user is connected to every BS (22 links for each user, a total of 2200) and every BS has a connection to all other BSs (21 links for every BS, a total of 462). There is a total of 2662 $f[i,j]$ and 2200 $a_{i,j}$ values to be assigned. For simplicity, assuming that $f[i,j]$ can only take discrete values in the range $[0, 10000]$, which is larger and continuous in reality, the problem complexity becomes $10000^{2662} 2^{2200} \approx 10^{11310}$. In terms of input variables,

assuming $S \gg A, T$, the search space complexity is $O(10000^{(S+I+T-1)IT} 2^{S(I+T)}) = O(10^{4SIT})$.

The MILP formulation has the advantage of guaranteeing optimal solution, meaning that this approach results in a high-performance network using the full potential of all available network elements. However, the cost associated with the MILP formulation in terms of execution time and memory usage is quite high as well, to such an extent that it is not suitable for real-time use. Because of this, while the MILP model is suitable to show the performance ceiling of a network in a time snapshot, it is inadequate for a dynamic network where UAV mobility is quite high. We will alleviate the problem of the MILP formulation with a novel algorithm that has significantly better performance in terms of execution time, with minor sacrifices in throughput performance. Chapter 4 will introduce heuristic solutions with the aforementioned characteristics.

CHAPTER 4

RELABEL-TO-FRONT ALGORITHM BASED SOLUTIONS TO THE UAV-UAR PROBLEM

While the UAV-UAR problem can be modeled as an optimization problem, there are also other approaches available. As we mentioned before in Chapter 2, flow networks can also be used to model cellular networks. Using a flow network allows the formulation of UAV-UAR problem as a maximum-flow problem, which has a significantly better performance in terms of execution time. However, the formulation requires some work as the UAV-UAR problem has some constraints that does not fit into maximum-flow problems by default.

As we mentioned in Chapter 2, push-relabel methods have the best time complexity among algorithms for solving the maximum-flow problem. As the basis of our algorithm, relabel-to-front algorithm defined in [33] is chosen. This algorithm is explained in detail in Chapter 2. In this chapter, we will first formulate the UAV-UAR problem as a maximum-flow problem. During this formulation, we will show the modifications that we make to extend the relabel-to-front algorithm to fit the UAV-UAR problem. Later, on top of the relabel-to-front algorithm, we will introduce two heuristic methods that use the relabel-to-front algorithm as a base, but have better results in total throughput achieved.

4.1 Adaptation of the Relabel-to-Front Algorithm to the UAV-UAR Problem

Recall the mathematical definition in Chapter 3: Let $G = (V, E)$ be a flow network where $V = \{\mathcal{S}, \mathcal{A}, \mathcal{T}\}$ is a set of vertices and E is its edges. $\mathcal{S} = \{s_1, s_2, \ldots, s_S\}$ is the set of vertices representing the UEs, $\mathcal{A} = \{a_1, a_2, \ldots, a_A\}$ is the set of vertices

representing the UAV-SCs and $\mathcal{T} = \{t_1, t_2, \ldots, t_T\}$ is the set of vertices representing the MBSs with fiber connectivity. Source vertices are the vertices in the set $\mathcal{S}$ and sink vertices are in the set $\mathcal{T}$. For the edges $(s, u) \in E$, the capacity $c(s, u)$ is a positive value that represents the link capacity in the cellular network.

The objective is to maximize total flow from UEs to MBSs. If we define $\mathcal{S}$ as the set of source nodes and $\mathcal{T}$ as the set of sink nodes, the problem at hand becomes a multi-source, multi-sink problem. By adding consolidated auxiliary nodes for sources (denoted *supersource*) and sinks (denoted *supersink*), this problem is transformed into a maximum-flow problem. Auxiliary nodes are connected only to their counterparts with edges that have infinite capacity.

While this problem can be solved using relabel-to-front algorithm to find the maximum total flow, the solution will not be valid as in its default form, the algorithm employs all available links from source nodes to allocate flows. In other words, the algorithm violates the user association constraint. To make sure that the found solution conforms to all problem constraints, minor changes are necessary. After finding a solution that maximizes the flows, if we eliminate the flows from source nodes in such a way that only one outgoing link is used, then the solution becomes valid and in this form, the algorithm can be used to solve the UAV-UAR problem. Since we eliminate edges that violate the user association constraint, we name this algorithm Relabel-to-Front-and-Eliminate (RTF-E).

RTF-E can be seen in Algorithm 6. The operation begins by calling the RELABEL-TO-FRONT algorithm. After it has returned a solution, the algorithm starts iterating all source vertices representing UEs in line 3. To select the ideal flow, the variable *ideal* is initialized in line 4. Then, the algorithm iterates through all outgoing flows from s in lines 5-9. The ideal flow is defined in such a way that (i) if there is at least one flow greater than the s's demand, then it is the flow with the least capacity, (ii) if all flows have capacities less than s's demand, then it is the flow with the most capacity. Lines 6-7 perform the operation for condition (i) and lines 8-9 for condition (ii). After finding the ideal flow, the flows are iterated again in line 10-14. All flows except *ideal* are cancelled in lines 11-12. Line 13-14 checks that if the allocated flow is greater than the demand, then it reduces the allocated flow to the s's demand,

removing the excess.

Algorithm 6 Relabel-to-Front-Eliminate

1: **procedure** RTF-E(G, $\mathcal{S}$, $\mathcal{T}$)
2: RELABEL-TO-FRONT(G, $\mathcal{S}$, $\mathcal{T}$)
3: **for** each vertex $s \in \mathcal{S}$ **do**
4: $ideal \leftarrow NULL$
5: **for** each vertex $i \in Adj[s]$ **do**
6: **if** $f[s,i] > f[s,ideal] \wedge f[s,ideal] < d_s$ **then**
7: $ideal \leftarrow i$
8: **else if** $f[s,ideal] > f[s,i] > d_s$ **then**
9: $ideal \leftarrow i$
10: **for** each vertex $i \in Adj[s]$ **do**
11: **if** $i \neq ideal$ **then**
12: $f[s,i] \leftarrow 0$
13: **else if** $f[s,i] > d_s$ **then**
14: $f[s,i] \leftarrow d_s$

In terms of complexity, line 2 takes $O(V^3)$ time since it is the default implementation of RELABEL-TO-FRONT algorithm. For loop at line 3 takes $O(V)$ time, and in this loop, all flows of a UE is iterated twice. A UE can be connected to at most $A + T$ vertices, but if we take this as V for simplicity, the complexity becomes $O(V)$ for the first loop. Second loop makes the same iteration with $O(V)$. Therefore, loop at line 3 has a complexity of $O(V(V + V))$, which is $O(V^2)$. The overall complexity becomes $O(V^3 + V^2)$, which is effectively $O(V^3)$. The complexity of RTF-E is the same as relabel-to-front, and much better than the exponential complexity of the MILP formulation.

In terms of correctness, relabel-to-front algorithm produces a valid flow network as a solution. For UAV-UAR problem, only user association constraint is violated. In the algorithm, the loop in lines 10-14 ensures that only one flow has nonzero value for every vertex representing a UE. Except the variable *ideal*, all flows are set to zero in line 12. Therefore, RTF-E algorithm solves the UAV-UAR problem. Note that in the second loop, skew symmetry constraints are violated. However, to calculate the end result, we are only interested in the outgoing flows from source nodes and not the

overall graph. This does not mean that the result is invalid, as validating the graph is not necessary since any updates will only reduce the assigned flows. Thus, the links used will remain the same, making the result valid. As mentioned before, this solution is not guaranteed to be an optimal one.

Arguably, RTF-E is a relatively easy adaptation to the relabel-to-front algorithm. In the next sections, we will propose different heuristics with better performance.

4.2 Heuristic Preflow Initialization

While RTF-E is an easy implementation, it does no considerations in terms of performance except keeping the link with the most allocated flow for every UE. In this section, we propose a heuristic to allocate preflows in such a way that the initial assignment chooses the best flow from all available flows. We will first explain the rationale behind the proposed heuristic, and then explain the algorithm in detail. Finally, we will perform the complexity and correctness analysis of the proposed heuristic.

In the relabel-to-front algorithm, first step is to initialize preflows from source(s). After this step, the algorithm tries to discharge all excess flows accumulated in the intermediate vertices. This operation solves the routing part of the UAV-UAR problem. Since source vertices are used only for preflow initialization and cancelling any allocated preflow that could not be discharged to the sink(s), user association problem must be solved in the preflow initialization.

When initializing preflows, relabel-to-front algorithm allocates all possible flows from every source node. To realize user association, this process has to be changed in such a way that only one preflow is allocated. This makes the initial state of the flow network conforming to the user association constraint. Since the algorithm does not perform any operations on source vertices except pushing allocated flows back if needed, the flow network does not violate the user association constraint. Another point to consider is the effect of the preflow initialization on throughput. Since this decision is final, choosing a flow is basically determining an upper bound on the user's throughput. This is because during the execution of the algorithm, the allocated flow is guaranteed to not go up, but may be pushed back by some amount if it cannot be

discharged to one or more sink vertices. For this purpose, we propose a heuristic for choosing the preflow. If there are no edges with a capacity that can satisfy the user's demand, then the edge with the most capacity is chosen. Otherwise, if there are a number of edges that has more capacity than the user's demand, among these edges, the edge with the least capacity is chosen.

The rationale behind the heuristic is explained as follows. If no edges can fully satisfy the demand, then choosing the one with the most capacity ensures that the potential upper bound is the highest available. For the second case, if we have multiple alternatives available, choosing the one with the least capacity leaves us room for potential improvements. We will employ this characteristic in the third algorithm that we introduce later in this chapter.

Algorithm 7 defines the procedure for heuristic preflow initialization. The procedure first initializes height, excess, and flow values for all vertices in lines 2-6. Starting from line 7, the algorithm iterates through source vertices with the variable s and initializes their heights and finds the ideal preflow. Line 8 initializes the height, and in lines 10-14, the algorithm iterates through all outgoing edges from s and finds the ideal flow according to the conditions that we described in the previous paragraph. Lines 11-12 handle the first condition and lines 13-14 handle the second condition. After the ideal flow is found, the algorithm first finds the flow value for the ideal flow. This is done in such a way that if it can satisfy the demand fully, then it is initialized to the user's demand (lines 15-16). Otherwise, the initialization uses the whole edge capacity (lines 17-18). Lines 19-22 make the assignment of flows and excesses.

Algorithm 7 Heuristic Preflow Initialization

```
1:  procedure HEURISTICINITIALIZEPREFLOW(G, S)
2:      for each vertex u ∈ V[G] do
3:          u.height ← 0
4:          u.excess ← 0
5:      for each edge (u, v) ∈ E[G] do
6:          f[u, v], f[v, u] ← 0
7:      for each vertex s ∈ S do
8:          s.height ← |V[G]|
9:          ideal ← NULL
10:         for each vertex u ∈ Adj[s] do
11:             if c(s, u) > c(s, ideal) ∧ c(s, ideal) < d_s then
12:                 ideal ← u
13:             else if c(s, ideal) > c(s, u) > d_s then
14:                 ideal ← u
15:         if c(s, ideal) ≥ d_s then
16:             cap ← d_s
17:         else
18:             cap ← c(s, ideal)
19:         f[s, ideal] ← cap
20:         f[ideal, s] ← −cap
21:         s.excess −= cap
22:         ideal.excess += cap
```

The complexity of HEURISTICINITIALIZEPREFLOW has a complexity of $O(V^2)$. The first for loop (lines 2-4) iterates through all vertices, which has $O(V)$ complexity, and second for loop (lines 5-6) iterates through all edges with $O(E)$ complexity. Final for loop (lines 7-22) iterates through all source vertices, and in this loop all adjacent vertices are iterated again in lines 10-14. Thus, the final for loop as a complexity of $O(V^2)$. For a fully connected graph, $E \approx V^2$, and the final complexity becomes $O(V + V^2 + V^2)$, which is equal to $O(V^2)$.

Algorithm 8 Relabel-to-Front with Heuristic Preflow Initialization

```
1:  procedure RTF-H(G, 𝒮, 𝒯)
2:      HEURISTICINITIALIZEPREFLOW(G, 𝒮)
3:      L ← V − {𝒮, 𝒯}
4:      for each vertex u ∈ V[G] − {𝒮, 𝒯} do
5:          current[u] ← N[u].head
6:      u ← L.head
7:      while u ≠ NULL do
8:          old-height ← u.height
9:          DISCHARGE(u)
10:         if u.height > old-height then
11:             current[u] ← N[u].head
12:         u ← next[u]
```

Using HEURISTICINITIALIZEPREFLOW(), we now define the Relabel-to-Front-Heuristic (RTF-H). The algorithm 8 is basically the same as relabel-to-front algorithm defined in Algorithm 5, but in line 2, we use our own heuristic preflow initialization. In terms of correctness, this algorithm solves the UAV-UAR problem. Since the relabel-to-front algorithm only violates the user association constraint, and line 2 solves this problem as it only initializes one nonzero flow per user, the final solution conforms to the UAV-UAR problem constraints. Heuristic preflow initialization has a complexity of $O(V^2)$, and the algorithm itself also has a complexity of $O(V^3)$, meaning that the overall complexity of this algorithm is also $O(V^3)$. In other words, introducing the heuristic preflow initialization incurs no performance loss compared to the relabel-to-front algorithm.

While RTF-E provides a solution to the UAV-UAR problem, RTF-H is expected to perform better in terms of total achieved throughput. RTF-E guarantees that it keeps the best flow from all remaining outgoing flows of a user. However, during the execution of the algorithm, since all users flood the network with all available links, the redundant capacity has to be discharged to the MBSs. While canceling the redundant flows afterwards results in a valid solution for UAV-UAR problem, the redundant capacity may result in less useful capacity being discharged. This in turn may result in

a non-ideal flow or the ideal flow with a less than optimal value being chosen. RTF-H alleviates both these problems: Since it only allocates one flow per UE, the algorithm does not handle any redundant flows. Furthermore, the flow selection selects the best flow available. Thus, RTF-H algorithm starts its execution with the best upper bound and is more likely to be closer to this upper bound in its final state.

4.3 Iterative Relabel-to-Front Algorithm

RTF-H is certainly an improvement over RTF-E, but as we mentioned before when we explained the heuristic, there is room for improvement. Recall that if there are multiple links available that can fully satisfy user demand, then the heuristic chooses the link with the least capacity. If during the execution of the algorithm the assigned flow gets reduced, other available alternatives may provide better results.

Consider this example: suppose a user s has a demand d_s and there are two links to vertices a, b available with $c(s, a) > c(s, b) > d_s$. The heuristic initially chooses the vertex b and assigns the flow $f[s, b] \leftarrow d_s$. Yet, after the algorithm finishes, the flow has a value that is less than d_s. In this case, there is another alternative available with a higher upper bound. The same argument can be made for cases where there is no edge with the capacity higher than the user's demand. In this case, $d_s > c(s, a) > c(s, b)$, and the vertex a is chosen with the flow $f[s, a] \leftarrow c(s, a)$. Suppose after the algorithm finishes, $c(s, b) > f[s, a]$, in which case the higher bound that the vertex b provides has the possibility of improving the assigned flow.

In both cases, while the alternatives available are not guaranteed to perform better than the assigned flow in the end, their upper bound are higher than the final result. Consequently, if we iteratively try other available alternatives for vertices whose demands are not fully satisfied, the total achieved throughput may actually increase. The algorithm can iterate until it founds a solution that is less than the previous one. We call this algorithm Relabel-to-Front-Iterative (RTF-I), and present this as our final heuristic method for UAV-UAR problem.

RTF-I algorithm uses RTF-H as its main operation, but performs the alternative flow selections between every call. In the initial call to RTF-H, the algorithm initializes

preflows as usual, but for subsequent calls, preflow initialization is not performed in RTF-H. After RTF-H finds a solution, RTF-I has to reset the graph state for intermediate and sink nodes in the same way as preflow initialization does. For source vertices, RTF-I has to scan every edge to determine if there is an alternative with a higher upper bound available. If so, the old flow is cancelled and new flow is initialized. After this is done for every source vertex, RTF-H is called again to find a solution on the same graph, but with a different set of preflows.

We will now explain the operation of RTF-I, as defined in Algorithm 9. The algorithm first makes a call to RTF-H. This call finds an initial solution and the graph state for the iterative algorithm to work on. The algorithm has a while loop in lines 3-30 that runs until RTF-H calls cannot improve the solution. The algorithm iterates through all vertices with the variable v in lines 4-29. In this iteration, height, excess, and flow states are reset for all intermediate and sink vertices in lines 25-29. For source vertices, we first find the assigned flow in line 7. An if condition checks whether the assigned flow is nonzero and less than the demand. Note that there can only be one nonzero flow of a source vertex. If the flow is less than the demand, then the algorithm searches for a better alternative in lines 9-13. This search is performed on all edges of the source vertex v except the edge of the assigned flow. In this search, the heuristic that RTF-H uses to select the ideal flow is directly used. Namely, if there are no edges that can fully satisfy d_v, then the edge with the highest capacity is chosen. This search is made in lines 10-11. Otherwise, if there are edges with capacities higher than d_v, then among those edges, the edge with the least capacity is chosen in lines 12-13.

There is the possibility that the found ideal capacity is lower than the actual assigned flow. Line 14 checks this condition and if this is the case, no changes are made to the assigned flow and the loop breaks (line 25). However, if the capacity is higher, then the old flow is reset and excess value is reduced accordingly for the receiving vertex in lines 19-20. If the ideal edge has a capacity higher than d_v, then d_v is assigned as the flow in lines 15-16. Otherwise, the flow fully uses the edge capacity in lines 17-18. Lines 21-24 make the assignment to the flows and excesses as necessary. After this operation is performed for every source vertex, RTF-H is called in line 30. However, in this call, RTF-H does not make any preflow assignment as we have already performed that step in the algorithm manually.

Algorithm 9 Iterative Relabel-to-Front with Heuristic Preflow Initialization

```
1:  procedure RTF-I(G, 𝒮, 𝒯)
2:      RTF-H(G, 𝒮, 𝒯)
3:      while RTF-H finds a better result than previous one do
4:          for each vertex v ∈ V do
5:              if v ∈ 𝒮 then
6:                  for each vertex i ∈ Adj[v] do
7:                      if d_v > f[v, i] > 0 then
8:                          ideal ← NULL
9:                          for each vertex j ∈ {Adj[v] − i} do
10:                             if c(v, j) > c(v, ideal) ∧ c(v, ideal) < d_v then
11:                                 ideal ← j
12:                             else if c(v, ideal) > c(v, j) > d_v then
13:                                 ideal ← j
14:                         if c(v, ideal) > f[v, i] then
15:                             if c(v, ideal) ≥ d_v then
16:                                 cap ← d_v
17:                             else
18:                                 cap ← c(v, ideal)
19:                         i.excess −= f[v, i]
20:                         f[v, i], f[i, v] ← 0
21:                         f[v, ideal] ← cap
22:                         f[ideal, v] ← −cap
23:                         v.excess ← cap
24:                         ideal.excess += cap
25:                         break
26:             else
27:                 v.height ← 0
28:                 v.excess ← 0
29:                 for each flow f[v, i], i ∈ 𝒜, 𝒯 do
30:                     f[v, i] ← 0
31:         RTF-H(G, 𝒮, 𝒯)
```

Recall that RTF-H has a complexity of $O(V^3)$. RTF-I starts with a RTF-H call, which has $O(V^3)$ complexity. Then, the while loop iterates through all vertices with the for loop at line 6. In this iteration, there are two nested for loops that iterate all flows of a source vertex. There can be $|V-1|$ connections at most, meaning that the complexity becomes $O(V^3)$ for the loop at line 6. There is another for loop at line 29 that goes through all flows of a non-source vertex. This also has a complexity of $O(V)$. Thus, the for loop at line 4 has a complexity of $O(V^3+V^2)$. In the while loop at line 3, there is a call to RTF-H, bringing the total complexity of the inner part of the while loop to $O(V^3+V^2+V^3)$. To find how many times the while loop can be executed at most, we must find how many alternative flows the algorithm tries. In the worst case, a source vertex must try every connection one by one, meaning that the while loop will be called V times. Combining all these figures together, RTF-I has a complexity of $O(V^3+V(V^3+V^2+V^3))$, which is $O(V^4)$.

In this chapter, we introduced three methods based on relabel-to-front algorithm that all solve the UAV-UAR problem, with significantly better runtime performance, but no optimality guarantee. In next chapters, we will present the simulator setup, empirical results of the proposed algorithms, and make a detailed analysis in terms of various parameters.

CHAPTER 5

SIMULATION RESULTS AND DISCUSSION

In the previous chapters, we defined the UAV-UAR problem, provided an MILP model and three heuristic algorithms based on the relabel-to-front algorithm and flow network modeling. In this chapter, we will validate and assess the performance of mentioned methods with simulations. First, we will generate the structures that will represent the network and explain this process in detail. Later, we will use generated topologies to perform Monte Carlo simulations, which will be used to analyze the effect of various parameters on network performance. After all mentioned steps, we will present the result of our simulations. The results will be presented in multiple subsections, and each subsection will focus on the effect of one parameter on the UAV-UAR problem. We will present the results grouped by the scenarios in every subsection.

5.1 Simulation Scenarios

We identified two scenarios for UAV-UAR problem that fits its use-cases. 3GPP introduces numerous use-cases and scenarios alongside their design parameters and equations that define operation-related parameters such as LoS probability or path loss. Two main scenarios interest us in this document: namely rural macro (RMa) and urban macro (UMa) use-cases [87].

As we mentioned in Chapter 2, aerial network technologies are frequently mentioned for enabling and/or improving rural connectivity. It is quite logical to apply UAV-UAR problem in this context. There are two main motivations in a rural context: the first motivation is bringing service using wireless backhauling, and the second

Table 5.1: Variables used in simulations with their range of values.

Variable	Explanation	Values
x	horizontal length of the operation area	2000 m for RMa, 1000 m for UMa
y	vertical length of the operation area	2000 m for RMa, 1000 m for UMa
f_c	center carrier frequency	28 GHz and 60 GHz
f	bandwidth	1.5 GHz and 3 GHz
G_t	transmitter antenna gain	8 dBi
G_r	receiver antenna gain	8 dBi
P_t	transmit power	[10,35] dBm
rec	receiver sensitivity	-98 dBm
d_i	throughput demand of i'th UE	[150,1200] Mbps
S	number of UEs	[200,500]
A	number of UAVs	[10,50]
T	number of MBSs	[2,4]
h_{UAV}	UAV height	[30,300] m
h_{UE}	UE height	1.5 m
h_{MBS}	MBS height	35 m for RMa, 25 m for UMa

one is to make this service high-capacity and mmWave, using 5G infrastructure. In a wide area, using some fixed MBS's, we will try to serve a number of users using UAV's as their BS's, employing wireless backhauling. For this use-case, we will use 28 GHz carrier frequency, which is in FR2 frequency band for 5G. While this carrier frequency will not supply very high throughput, it is nevertheless more reliable than higher frequency bands, and has less free space path loss. We will also use 60 GHz frequencies for comparison.

Due to low LoS probability, urban scenarios are often very dependent on smart deployment schemes to keep LoS connections. While the UAV-UAR problem is not related with deployment, it can nevertheless be applied in urban scenarios to maximize throughput. Even though the algorithm will perform optimizations in terms of

user association and routing, compared to the RMa scenario, deployment is a more effective factor on throughput in UMa scenario. We will use the same frequencies for UMa scenario as well. In urban scenarios, UAV's can be employed to increase the LoS probability and link strength by shortening the links between UEs and BSs.

For both scenarios, we will start with a fixed deployment scheme. All network elements will have randomized coordinates at first, which will also determine the link strengths between them. The topology generation will be explained in the next section. Link capacities will be determined using path loss and LoS probability equations from 3GPP [87].

5.2 Topology Generation

After defining the scenarios that will be used for simulations, we need example topologies that will be used to solve the UAV-UAR problem. We require a number of UEs, UAV-SCs, and MBSs to form a network. This network will be conforming to the network model that we introduced in Chapter 3 to define the UAV-UAR problem on. To generate such a network, we first need the locations of the aforementioned network elements. For this purpose, we define Algorithm 10, which will first generate the coordinates for network elements.

The coordinates alone are not enough to form a network. Using the locations, we will then introduce link and link capacities will be calculated using link budgets. Path loss calculations will be done according to the formulas given by 3GPP [87] for our scenarios. LoS probabilities are also needed to calculate path losses, which are also given in 3GPP [87]. Using path losses, antenna gains, and noise, we calculate the received power, and we then find the link capacity in terms of Mbps.

The algorithm GENERATETOPOLOGY starts by randomly generating coordinates in line 2. x, y values use the values in Table 5.1 while z values are determined by height values in the aforementioned table depending on the type of vertex initialized. In three loops, source, intermediate, and sink vertices are generated. In lines 10-20, using the generated coordinates, distances between two nodes are found. Then, using this distance, LINKBUDGET finds the capacity using a path loss and LoS probability

Algorithm 10 Topology Generation

```
1:  procedure GENERATETOPOLOGY(S, A, T)
2:      coordinates ← Generate |V| random 3D coordinates
3:      vertices ← [ ]
4:      for i in 0 to S do
5:          vertices.append(VERTEX(TYPE=SOURCE))
6:      for i in 0 to A do
7:          vertices.append(VERTEX(TYPE=INTERMEDIATE))
8:      for i in 0 to T do
9:          vertices.append(VERTEX(TYPE=SINK))
10:     for i in 0 to |V| do
11:         v1 ← vertices[i]
12:         c1 ← coordinates[i]
13:         for j in 0 to |V| do
14:             v2 ← vertices[j]
15:             c2 ← coordinates[j]
16:             if v1.type ≠ SOURCE ∧ v2.type ≠ SOURCE then
17:                 distance ← EUCLIDEANDISTANCE(c1, c2)
18:                 if distance > 0 then
19:                     capacity ← LINKBUDGET(distance)
20:                     (v1, v2) ← capacity
21:                     (v2, v1) ← capacity
```

specific for the scenario, distance, and other parameters. If both vertices are not source vertices, two-way edges are generated with the found capacity at lines 20 and 21. If both are source vertices, we do not generate any links as UEs do not connect to each other in cellular networks. In the end, we have a graph which represents a network. Using this graph, the implemented algorithms will add necessary data structures such as flows and solve the UAV-UAR problem.

5.3 Monte Carlo Simulations

After generating a topology as defined in the previous section, we will perform Monte Carlo simulations to compare the results of MILP and RTF-based solutions. To solve the MILP formulation, we use Gurobi solver [88] for Python. RTF-based algorithms are implemented in Python. For a number of steps, we generate a topology, and using this topology, find a MILP-based and a RTF-based result. Using multiple such simulations, we find an averaged result.

In these simulations, we will first present the performance of RTF-based heuristics, namely RTF-E, RTF-H, and RTF-I. After this, we will give a comparison of RTF-I and MILP optimizer. We will perform simulations to see the effect of UAV height, receiver sensitivity, transmit power, and different user demands. For all simulations, our performance metric will be the total throughput and its ratio to the total demand.

5.4 Simulation Results

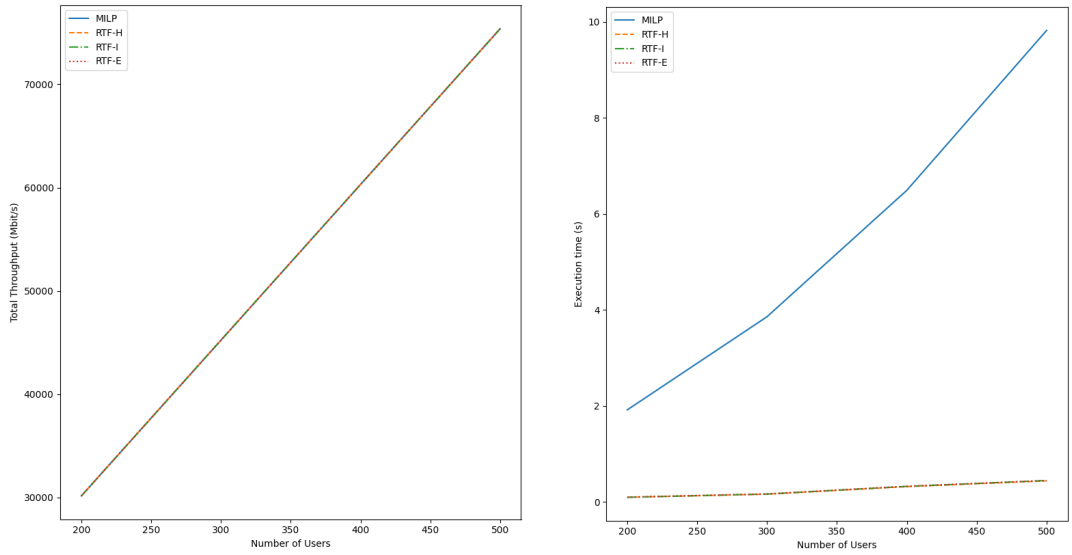
This section presents the results of our simulations. In the first subsection, a comparison of all RTF-based heuristics will be given. In the later subsections, we will observe the effect of four parameters, namely different user demands, different transmit powers, different receiver sensitivities, and different UAV operation heights.

The results that we present will be based on three metrics, namely total satisfied throughput, satisfied throughput percentage, and associated user percentage. Total satisfied throughput percentage denotes the total user throughput that the network manages to handle and is used for MILP and RTF comparison figures. The second

metric, namely the satisfied throughput percentage, is in figures that show the effect of different simulation parameters under different user numbers. We use percentage instead of total throughput simply to be able to normalize results based on user numbers and compare results under different user numbers. The final metric, namely the associated user percentage, shows how many users managed to connect to the network and have their data delivered to MBSs in any way. we denote connection as the satisfaction demand of user being greater than 20%. Under this threshold, the user is considered not connected to the network.

5.4.1 Performance of RTF-Based Heuristics

First of all, we will show the performance of proposed heuristics and MILP solution. For this demonstration, we have two cases: RMa scenario with 28GHz carrier frequency, and UMa scenario with 28 GHz carrier frequency. UAV and MBS numbers are fixed to 30 and 4, respectively.



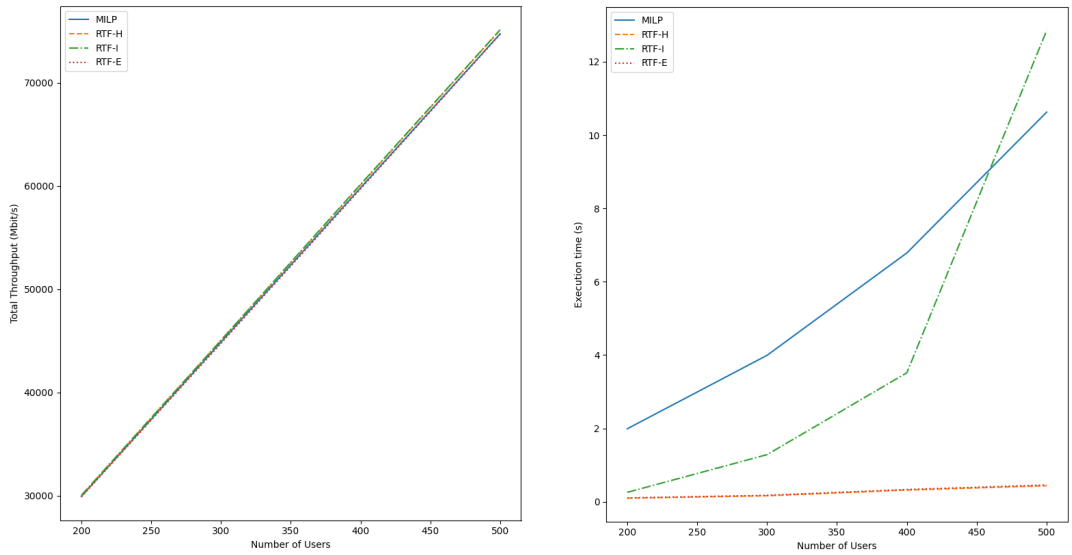
(a) Total throughput

(b) Execution time

Figure 5.1: Comparison of RTF-based heuristics and MILP optimization in terms of total throughput and execution time is shown.

Figure 5.1 shows three heuristics and MILP optimization in two subfigures. Subfigure 5.1a shows the number of users versus total throughput, whereas Subfigure 5.1b shows the number of users versus execution time. In Subfigure 5.1a, MILP result denotes the optimal result that can be found. Note that all RTF results are almost identical to that of MILP, which is a testament to the performance of proposed heuristics as they work almost optimal in this case. Real gains can be observed in Subfigure 5.1b, where the execution time of the algorithms are shown. MILP execution times begin at 2 seconds and go as high as 10 seconds with the highest user count, whereas all RTF methods have less than 0.2 second execution time with the highest user count, which amounts to a speedup of 50 times.

Figure 5.2 shows the results in the same format but for UMa scenario with 28GHz carrier frequencies. In Subfigure 5.2a, in terms of throughput, except RTF-I, all results are optimal whereas RTF-I is quite close to optimal. The real difference is evident in Subfigure 5.2b where the execution time performance can be seen. In this subfigure, while RTF-E and RTF-H have significantly outperformed MILP, RTF-I has significantly worse execution time than other two heuristics. Even so, RTF-I still outperforms MILP until user count exceeds 500. The reason this happens is that when the number of users gets higher, there are significantly more alternatives for RTF-I to try. Furthermore, more users means more demand to satisfy, which increases the problem difficulty, further degrading the execution time. These two characteristics combined results in a significant performance loss, especially if RTF-I cannot attain the maximum throughput in the first try.



(a) Total throughput

(b) Execution time

Figure 5.2: Comparison of RTF-based heuristics and MILP optimization in terms of total throughput and execution time is shown.

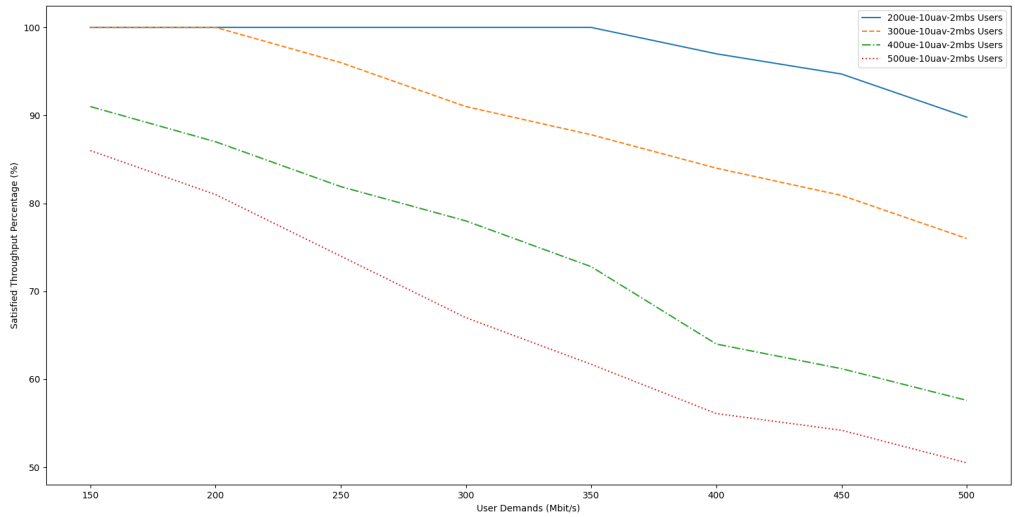


Figure 5.3: Comparison of different user demands and user numbers in terms of satisfied percentage is shown for RMa scenario and 28 GHz carrier frequency.

5.4.2 Effect of User Demand

In this section, we will show the effect of user demand on the network performance. For this purpose, demands ranging between CAT4 and CAT9 are used.

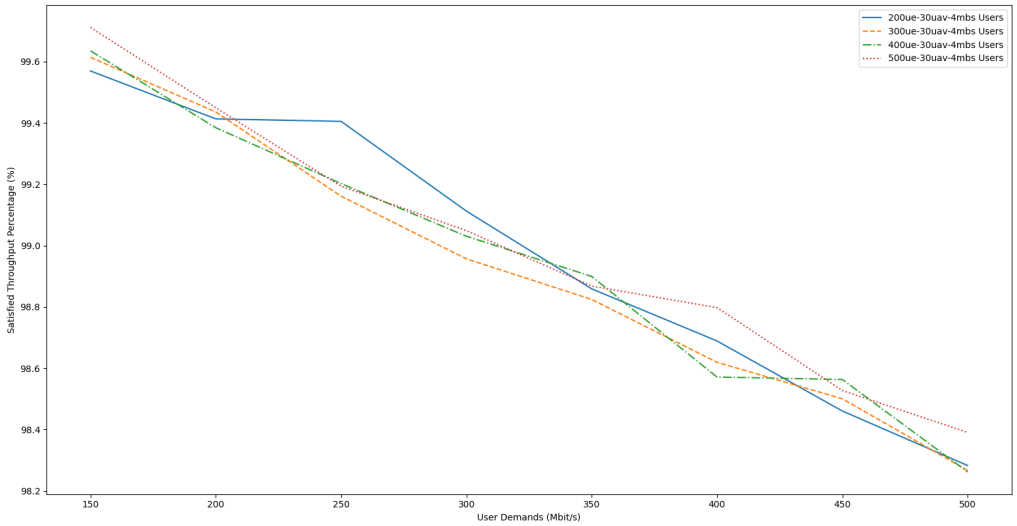


Figure 5.4: Comparison of different user demands and user numbers in terms of satisfied percentage is shown for UMa scenario and 28 GHz carrier frequency.

For the effect of different user demands in RMa scenario, we present Figure 5.3. 28 GHz carrier frequency, 10 UAVs and 2 MBSs are simulated. This figure presents the effect of increased user demand quite nicely. For every figure, the satisfied percentage drops when the user demand is increased. However, while for 200 users the satisfaction is still above 90%, 300 users have a 17% drop, and this trend gets worse for 400 and 500 users with 30% and 37% drops, respectively. For a fixed number of UAVs and MBSs, the network can support only so much total demand, and after a certain threshold, cannot fully satisfy users because of the limited backhaul capacity. As the effect of demand increase becomes much higher with higher user counts, consequently, the drop in satisfied percentage is also sharper.

Figure 5.4 shows different user demands and how much of them RTF-H managed to satisfy. This graph is for UMa scenario and 28 GHz carrier frequency. UAV and

MBS numbers are 30 and 4, respectively. The drop is almost linear for increasing user demand. Moreover, the number of users does not seem to have any effect whatsoever, the figures being more or less the same for different users. This suggests that another parameter affects the performance here, namely the deployment of UAVs and MBSs.

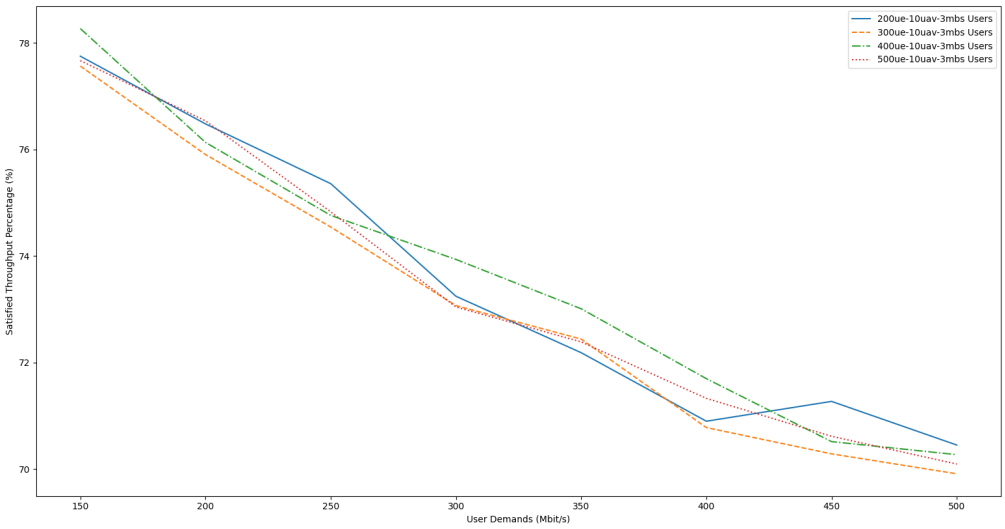


Figure 5.5: Comparison of different user demands and user numbers in terms of satisfied percentage is shown for UMa scenario and 60 GHz carrier frequency.

Figure 5.5 shows different user demands for UMa scenario and 60 GHz carrier frequency. UAV and MBS numbers are 10 and 3, respectively. Same as the previous figure, there is a negative correlation between the demand and satisfied percentage. The number of users also seems to have no effect again, as all users have attained more or less the same percentage.

5.4.3 Effect of Transmit Power

In this subsection, we will investigate the effect of transmit power on the network performance.

5.4.3.1 RMa Scenario

Next two simulations are performed using the RMa scenario.

In Figure 5.7, the effect of different transmit powers on RMa scenario can be seen. For this simulation, 28 GHz carrier frequency is used. UAV and MBS numbers are 30 and 4, respectively. In terms of throughput, it can be seen from Subfigure 5.7a that both RTF-H and MILP achieve optimal results. Subfigure 5.7b shows us interesting information in terms of execution time. The increase in transmit power results in an increase in the execution time. This can be seen especially in the MILP results, as the execution time roughly doubles when the transmit power is increased from 30dBm to 40dBm. This increase can be observed in the RTF-H execution times as well, but their values are significantly less than the MILP ones. The reason for this increase is that transmit power and number of links are positively correlated. This increase in links also results in an increase in execution time as MILP/RTF takes into account more flows.

In Figure 5.8, the effect of different transmit powers on 60 GHz carrier frequency can be seen. UAV and MBS numbers are 30 and 3, respectively. In terms of throughput, same as the last simulation, RTF achieves the optimal results of MILP. In terms of execution time, values are significantly less than the ones in Figure 5.7, since there is one less MBS. However, even though carrier frequencies are different, same trend can be observed in the execution time, as it increases with higher transmit powers.

To show the effect of transmission power alone, we present Figure 5.6. The figure shows transmit power versus satisfied percentage demand. UAV and MBS counts are 10 and 3, respectively. From the figure, it can be seen that as the number of users goes up, the satisfied demand percentage goes significantly lower for all values of transmit power. For 200 users, 25dB seems to be enough to satisfy the demand, but for higher user numbers, the improvement still stands with higher transmit power. Furthermore, as the number of users are increased, the effect of transmit power seems to be closer. This shows that especially for lower number of UAVs and MBSs, increasing transmit power might be a solution to increase network performance.

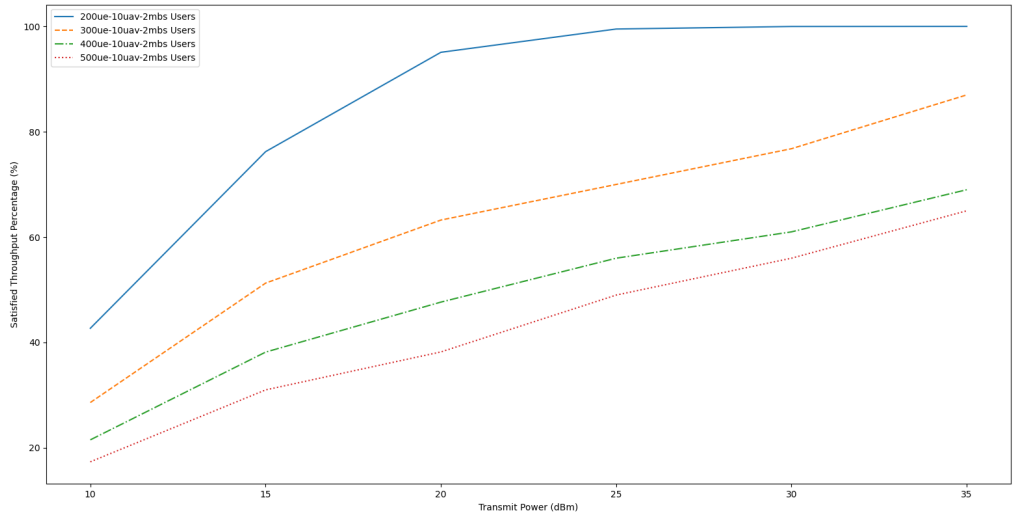
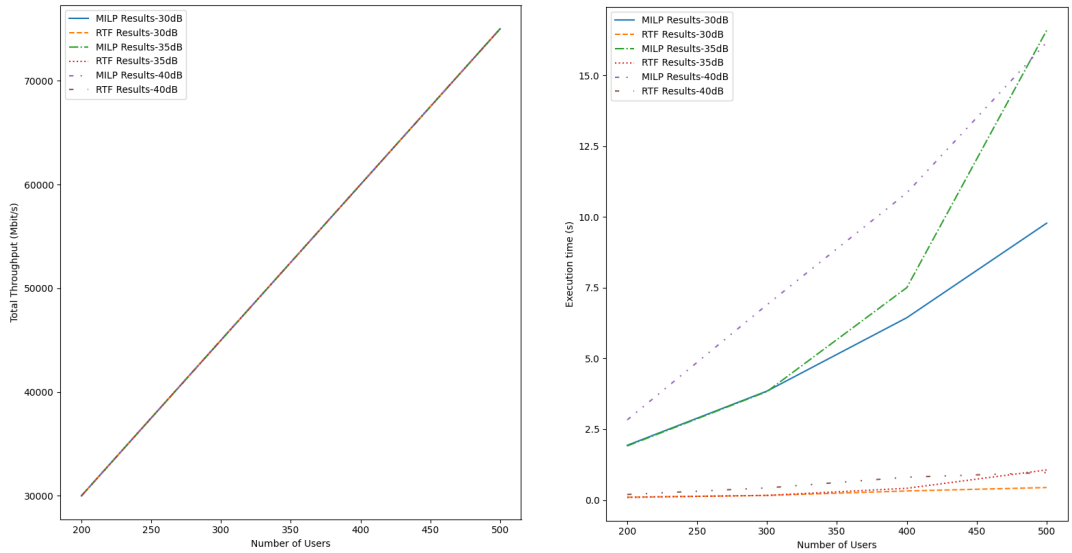


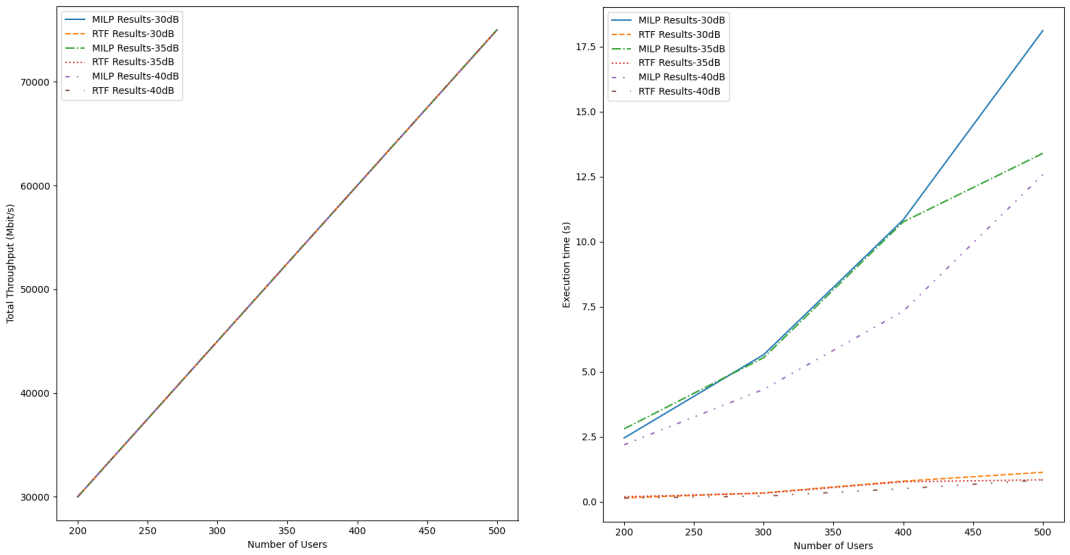
Figure 5.6: Effect of transmit power on RMa scenario and 28 GHz frequency is shown.



(a) Total throughput

(b) Execution time

Figure 5.7: Comparison of different transmit powers on RMa scenario and 28 GHz frequency in terms of total throughput and execution time is shown.



(a) Total throughput

(b) Execution time

Figure 5.8: Comparison of different transmit powers on RMa scenario and 60 GHz frequency in terms of total throughput and execution time is shown.

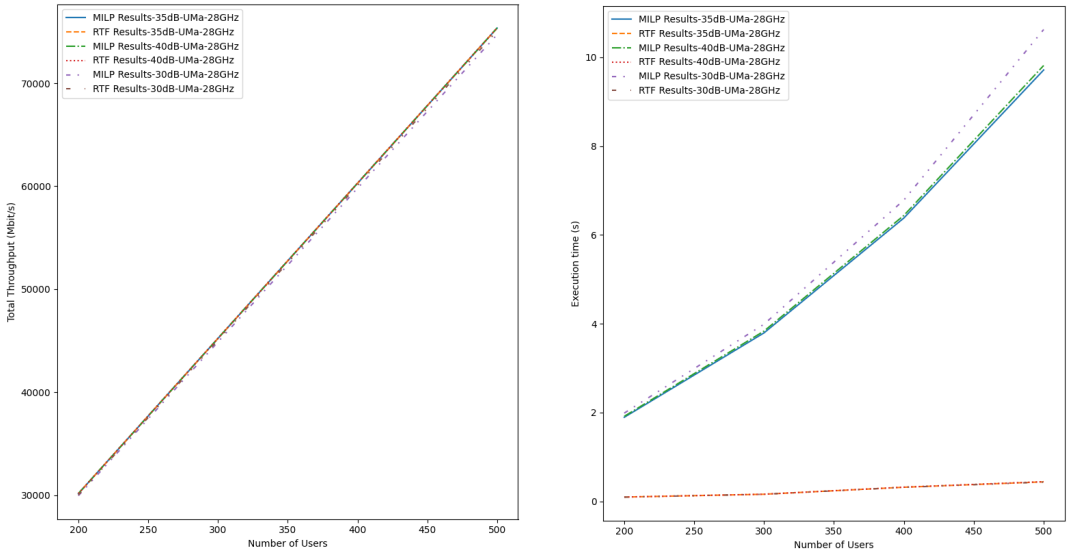
5.4.3.2 UMa Scenario

Next two simulations are performed using the UMa scenario.

In Figure 5.9, the effect of different transmit powers with 28 GHz carrier frequency can be seen. For this simulation, UAV and MBS numbers are fixed as 30 and 4, respectively. In terms of throughput, the increase in transmit power did not result in any apparent increase as the lowest value already attains optimal throughput. In terms of execution time, RTF-H outperforms MILP as low as two times and as high as 10 times.

In Figure 5.11, the effect of different transmit powers with 60 GHz carrier frequency can be seen. For this simulation, UAV and MBS numbers are fixed as 50 and 4, respectively. Similar to the previous result, the optimal throughput is achieved. However, since the UAV counts in this simulation are significantly higher than the previ-

ous one, the difference in execution time is also significantly higher, so much so that RTF-H outperforms MILP by roughly 130 times.



(a) Total throughput

(b) Execution time

Figure 5.9: Comparison of different transmit powers on UMa scenario and 28 GHz

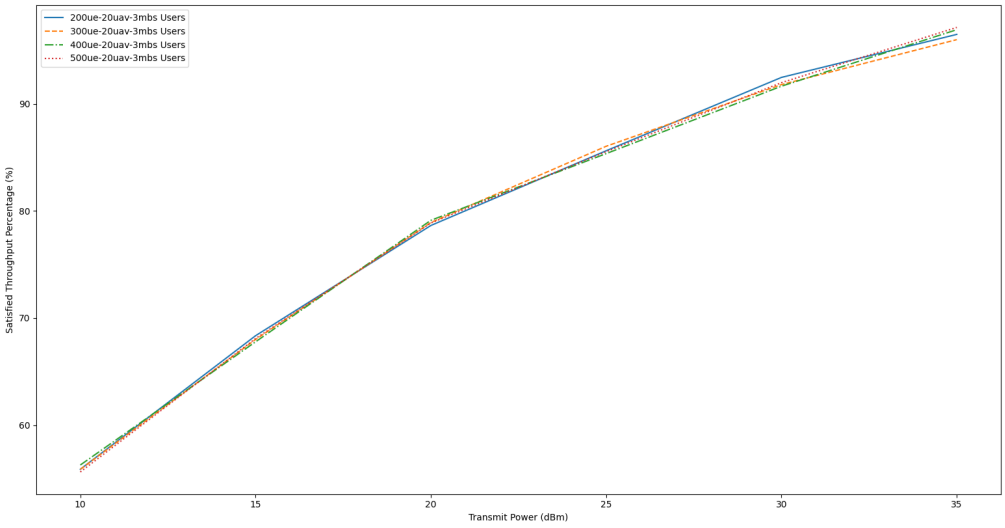
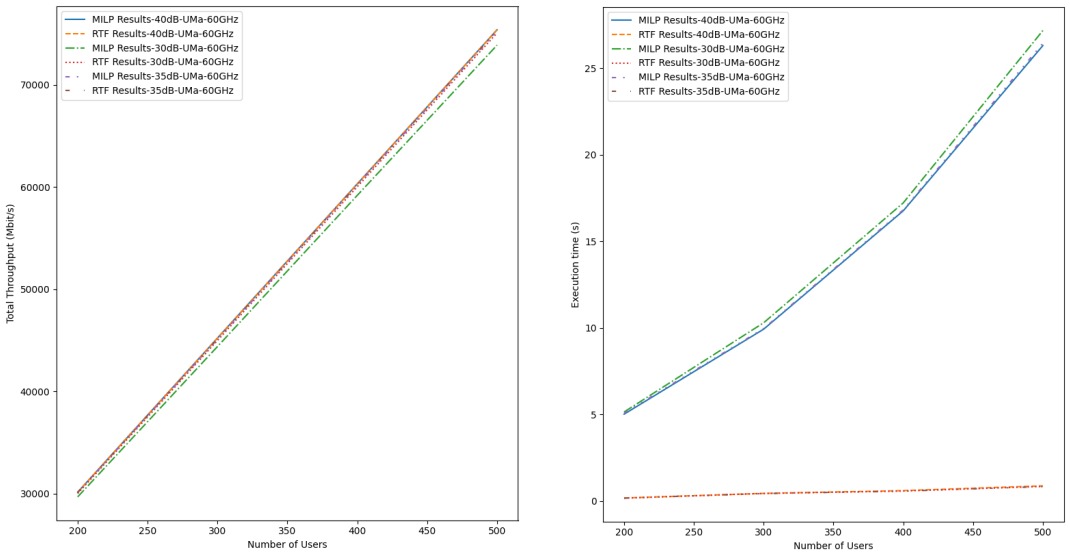


Figure 5.10: Effect of transmit power on UMa scenario and 60 GHz frequency is shown.



(a) Total throughput

(b) Execution time

Figure 5.11: Comparison of different transmit powers on UMa scenario and 60 GHz frequency in terms of total throughput and execution time is shown.

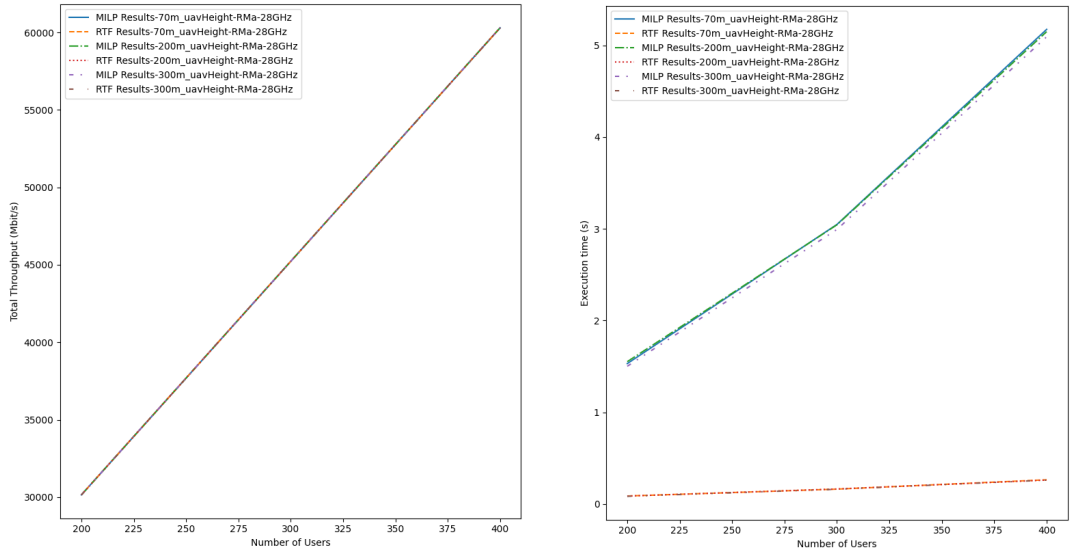
Similar to RMa scenario, we present Figure 5.10 to show the individual effect of transmit power. The figure shows transmit power versus satisfied percentage demand. UAV and MBS counts are 20 and 3, respectively. Unlike the RMa scenario, however, the results are stacked together, but they still show the positive correlation between the transmit power and satisfied percentage. After 30 dBm, the satisfied percentage goes above 90%, which is a very good increase from the initial point of 55%.

5.4.4 Effect of UAV Altitude

In this subsection, we will investigate the effect of UAV altitude on network performance. We have tried three values for UAV altitudes, namely 70, 200, and 300m.

5.4.4.1 RMa Scenario

In Figure 5.12, a RMa scenario network with 28GHz operating frequency can be seen. For this simulation, UAV and MBS numbers are fixed in 30 and 3, respectively. In Subfigure 5.12a, it can be seen that all simulations have attained the maximum throughput. For execution time, Subfigure 5.12b shows that irrespective of the UAV altitude, RTF-based methods significantly outperform the MILP optimization.



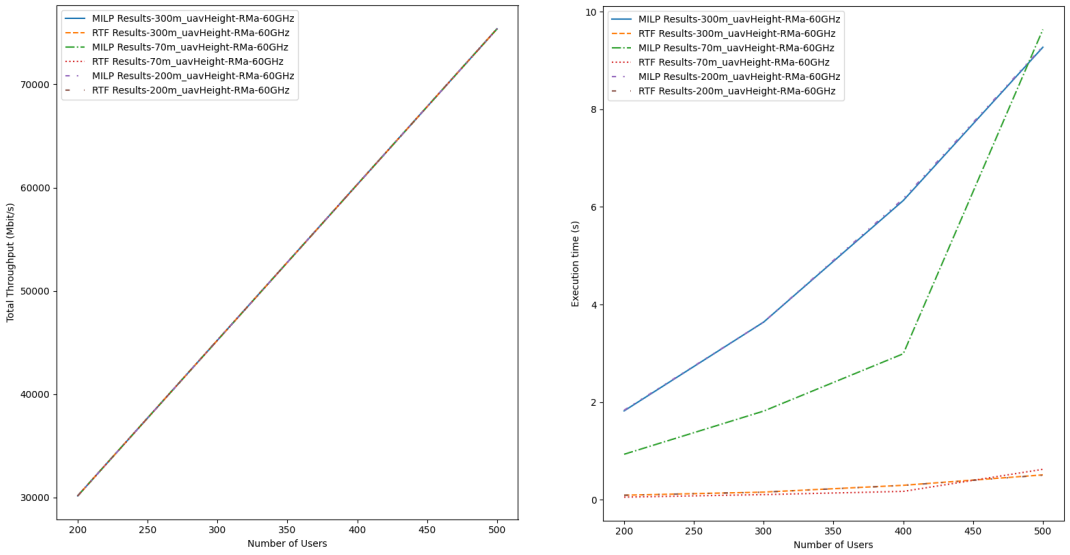
(a) Total throughput

(b) Execution time

Figure 5.12: Comparison of different UAV altitudes on RMa scenario and 28 GHz frequency in terms of total throughput and execution time is shown.

In Figure 5.13, same simulation is repeated with 60 GHz carrier frequency, and number of UAVs and MBSs are 30 and 3, respectively. In Subfigure 5.13a, it can be seen that RTF results have achieved optimal throughput even with 300m UAV altitudes. What's different from the 28 GHz results can be seen in Subfigure 5.13b. Compared to the 28 GHz simulations, execution times are worse for both cases. This is because for 60 GHz carrier frequency, LoS is more difficult to establish when the distance is increased, which in turn increases the execution time as there are fewer links that can

handle user traffic. This is true for both access links and UAV-UAV and UAV-MBS backhaul links.



(a) Total throughput

(b) Execution time

Figure 5.13: Comparison of different UAV altitudes on RMa scenario and 60 GHz frequency in terms of total throughput and execution time is shown.

Figure 5.14 shows the effect of UAV altitude more in-depth. 28 GHz carrier frequency is used, with the number of UAVs and MBSs being 10 and 2, respectively. The effect of increasing the UAV altitudes becomes more evident when the number of available BSs is reduced. When the number of users is 200, UAV altitude does not affect achieved throughput as severely, just a mere 7%. 300 users has a slightly worse drop of 13%. The trend gets worse for higher user numbers with 400 users suffering a 24% drop and 500 users a 35% one. As the UAV altitude is increased, the effect is worse when the number of users are more.

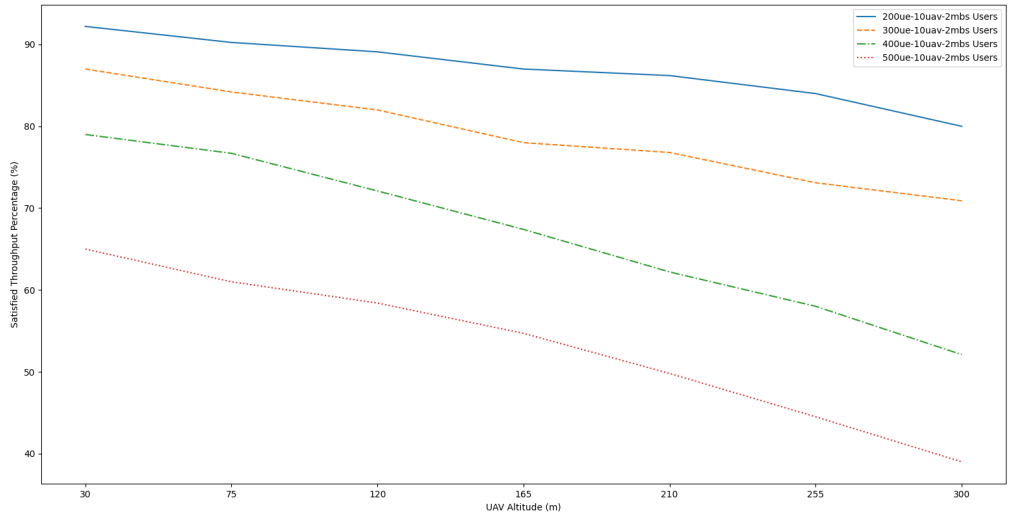
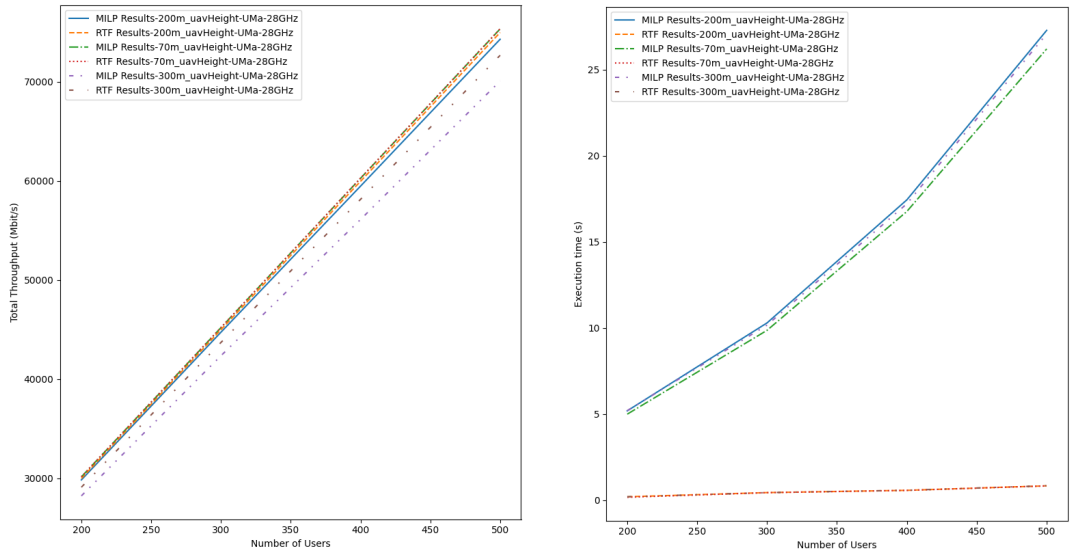


Figure 5.14: Effect of UAV altitude on RMa scenario and 28 GHz frequency is shown.



(a) Total throughput

(b) Execution time

Figure 5.15: Comparison of different UAV altitudes on UMa scenario and 28 GHz frequency in terms of total throughput and execution time is shown.

5.4.4.2 UMa Scenario

In Figure 5.15, different UAV altitudes are tested in UMa scenario with 28 GHz carrier frequency. Number of UAVs and MBSs are 50 and 4, respectively. In Subfigure 5.15a, it can be seen that achieved throughput reduces with increasing UAV altitude. In terms of execution time, compared to the RMa scenario simulations, the performance gains are even more evident in Subfigure 5.15b, RTF-H outperforming MILP by 25 and 135 times for 200 and 500 users, respectively.

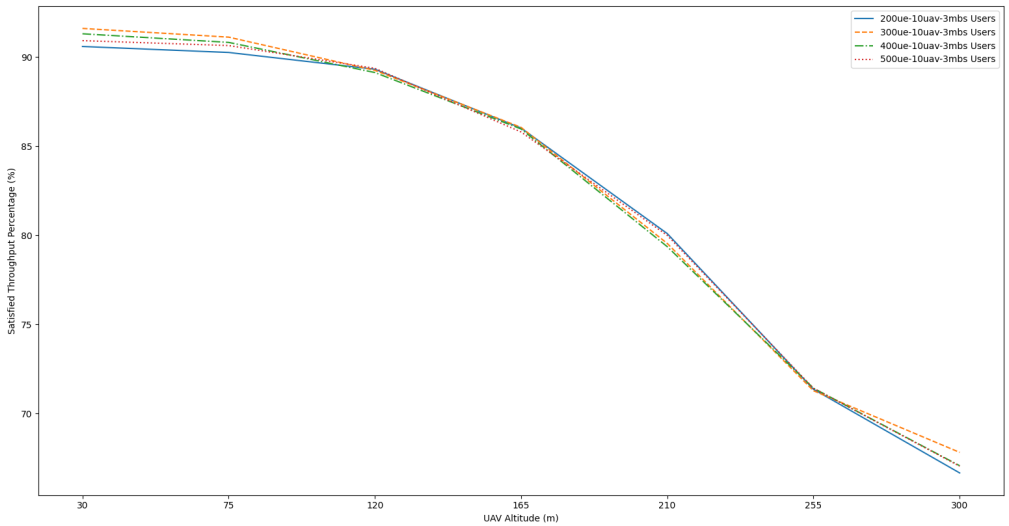


Figure 5.16: Effect of UAV altitude on UMa scenario and 28 GHz frequency is shown.

Figure 5.16 shows the effect of UAV altitude for UMa scenario more in-depth. 28 GHz carrier frequency is used, with the number of UAVs and MBSs being 10 and 3, respectively. Compared to RMa scenario, the results are once more stacked similar to the other parameters. Furthermore, while the drop in RMa scenario is smoother, UMa scenario has a sharper drop, especially after 120 m. The number of users seems not to be having an inverse effect as the results are more or less the same for all user numbers. This shows that the network is not limited in terms of total link capacity, but something else, which is deployment, as we mentioned before for UMa scenario.

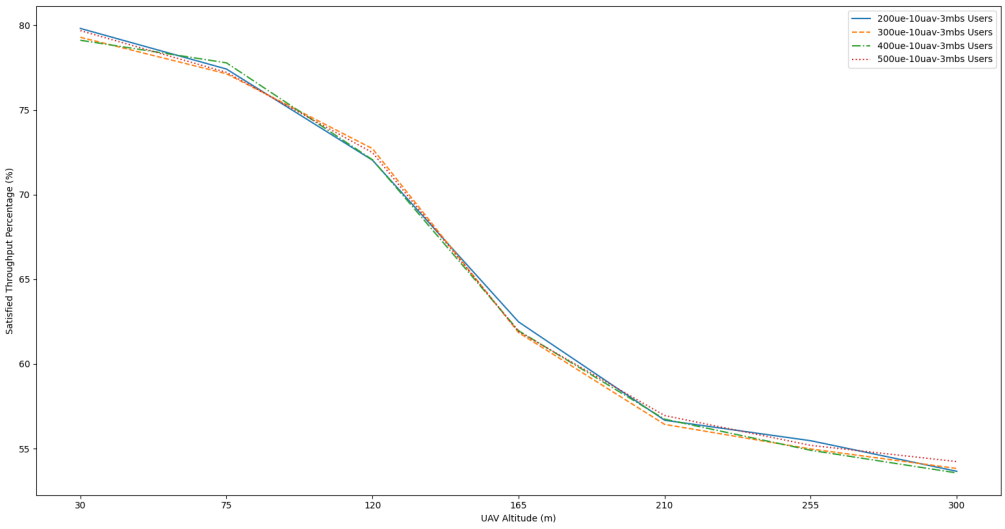


Figure 5.17: Effect of UAV altitude on UMa scenario and 60 GHz frequency is shown.

Figure 5.17 repeats the last simulation but with 60 GHz carrier frequency. UAV and MBS counts are the same as well. Compared to Figure 5.16, satisfied percentages are shown to be even lower. Also, the drop in satisfied percentage is sharper. This is mostly because 60 GHz frequency is affected more by deployment, which significantly degrades the performance in this case as Monte Carlo simulations randomly assign locations for UAVs. Number of users also does not seem to have an adverse impact as the results are more or less the same as before.

5.4.5 Effect of User Association

For UMa scenario simulations, the results do not seem to be on par with the RMa scenario. To prove our hypothesis of deployment affecting the results, we present the same results with more resolution and associated user percentage instead of satisfied throughput. If our hypothesis is true, the associated user percentage should be the same for a given parameter set even when the number of users is changed.

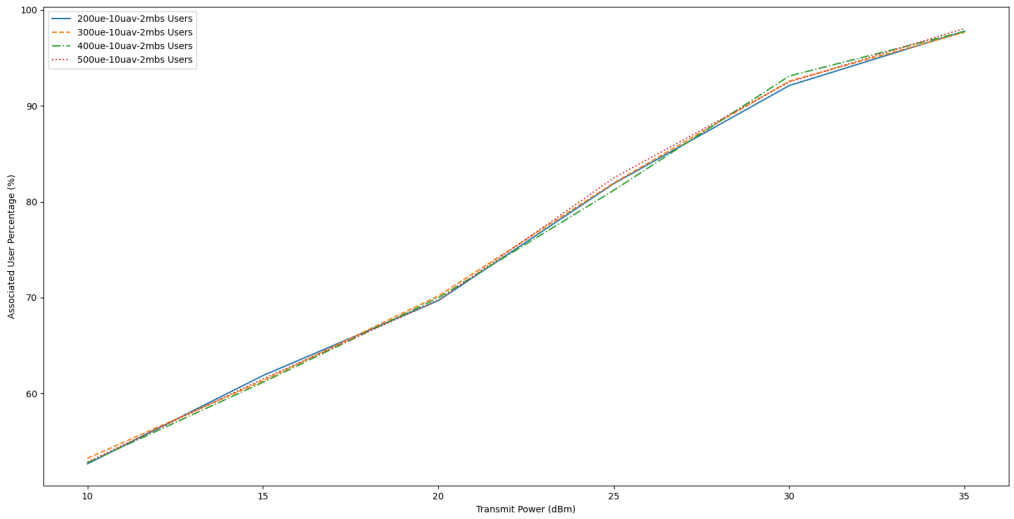


Figure 5.18: Associated user percentages with different transmit powers in UMa scenario and 28 GHz carrier frequency is shown.

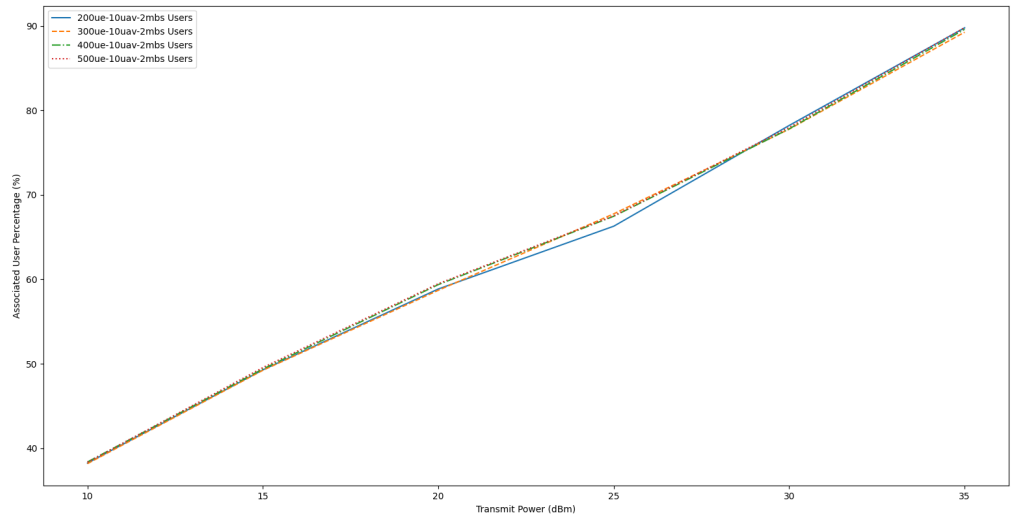


Figure 5.19: Associated user percentages with different transmit powers in UMa scenario and 60 GHz carrier frequency is shown.

Figures 5.18 and 5.19 show associated user percentages for different transmit powers

for UMa scenario. Notice that the percentages are the same for a fixed transmit power. In a backhaul-constrained network, it is expected that for a fixed set of parameters, this percentage drops when the users are increased. This is because when the backhaul capacity is fully allocated, new users cannot be served and user association percentage drops. However, since these two figures show otherwise, backhaul links are available, but the users cannot find any suitable connection to any of the available base stations. Hence, the percentage remains fixed because of random deployment. Same trend can be observed for UAV altitudes (Figures 5.22 and 5.23) and user demands (Figures 5.20 and 5.21) but in the opposite manner since these two parameters are negatively correlated with network throughput whereas transmit power is positively correlated.

Compared to the UMa cases, Figure 5.24 shows the user association percentage on different transmit powers in RMa scenario. The difference is evident; having more users directly affects the performance and associated user percentage. Same trend can also be observed in Figures 5.25 and 5.26.

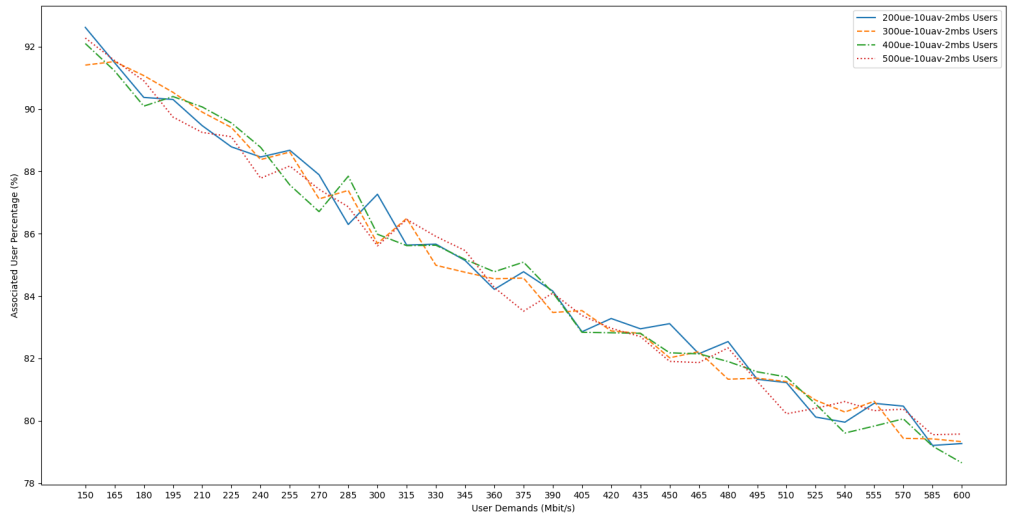


Figure 5.20: Associated user percentages with different user demands in UMa scenario and 28 GHz carrier frequency is shown.

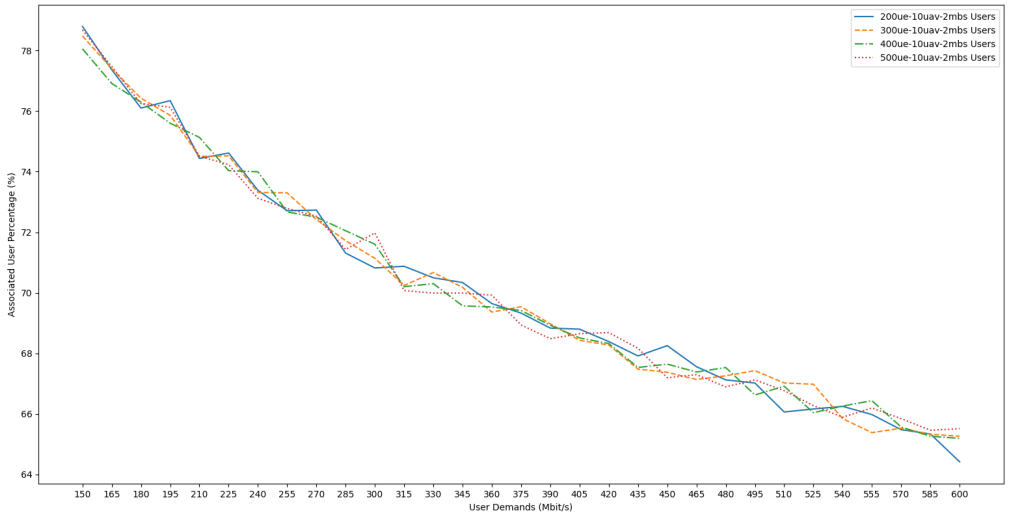


Figure 5.21: Associated user percentages with different user demands in UMa scenario and 60 GHz carrier frequency is shown.

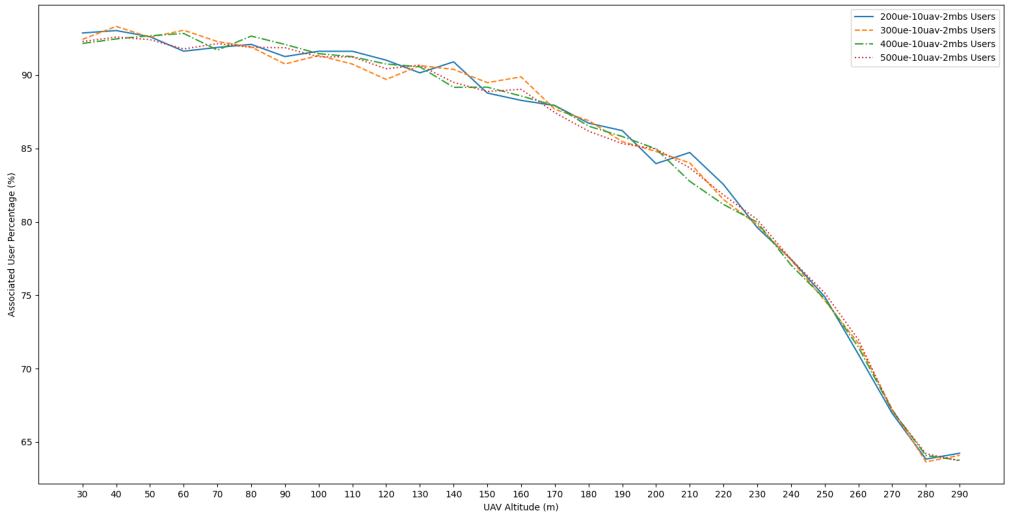


Figure 5.22: Associated user percentages with different UAV altitudes in UMa scenario and 28 GHz carrier frequency is shown.

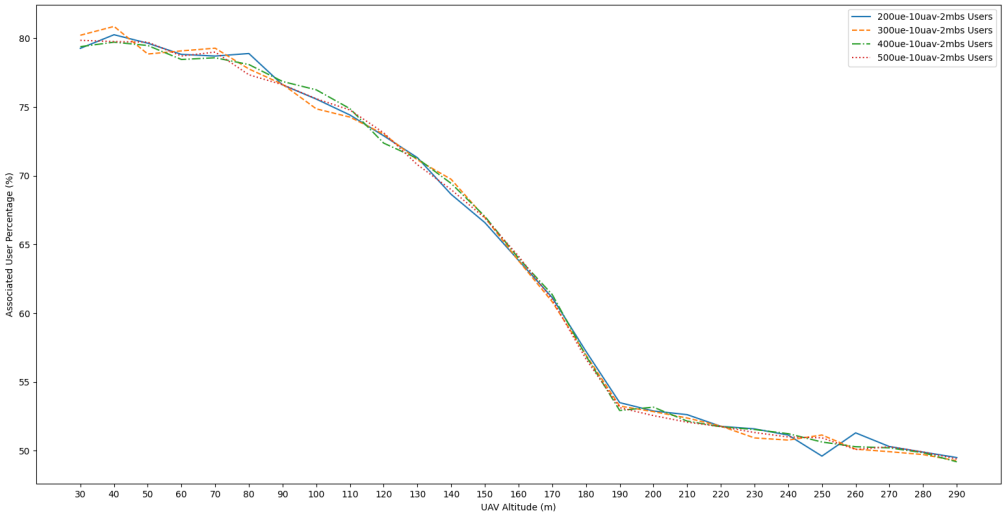


Figure 5.23: Associated user percentages with different UAV altitudes in UMa scenario and 60 GHz carrier frequency is shown.

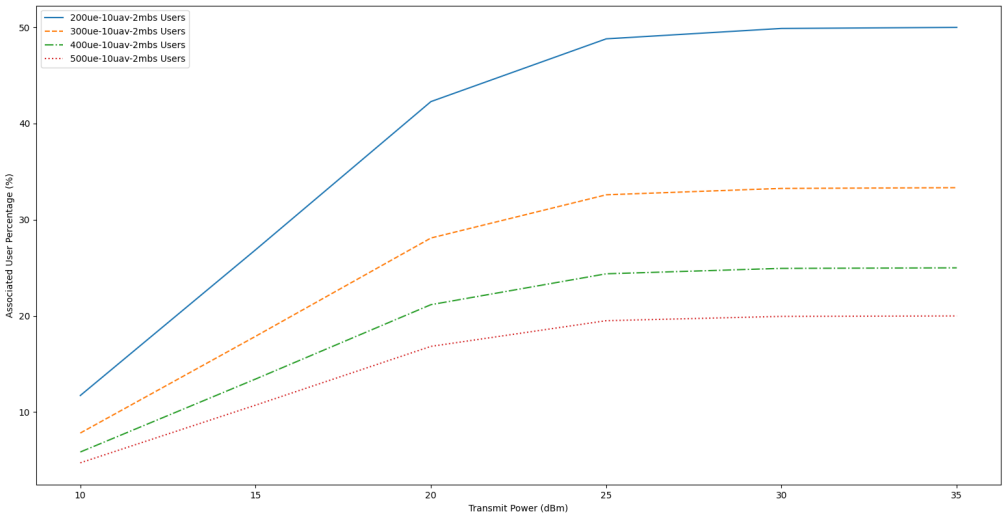


Figure 5.24: Associated user percentages with different transmit powers in RMa scenario and 28 GHz carrier frequency is shown.

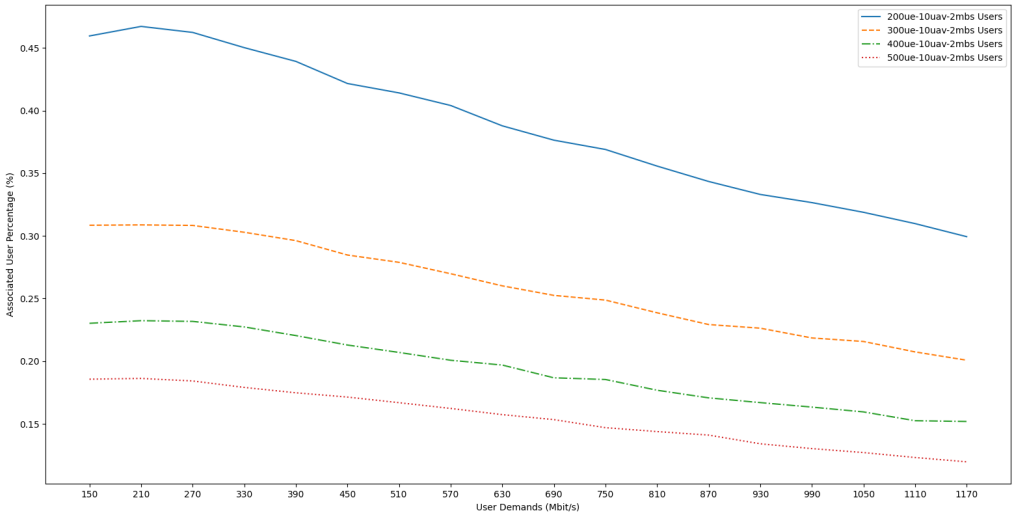


Figure 5.25: Associated user percentages with different user demands in RMa scenario and 28 GHz carrier frequency is shown.

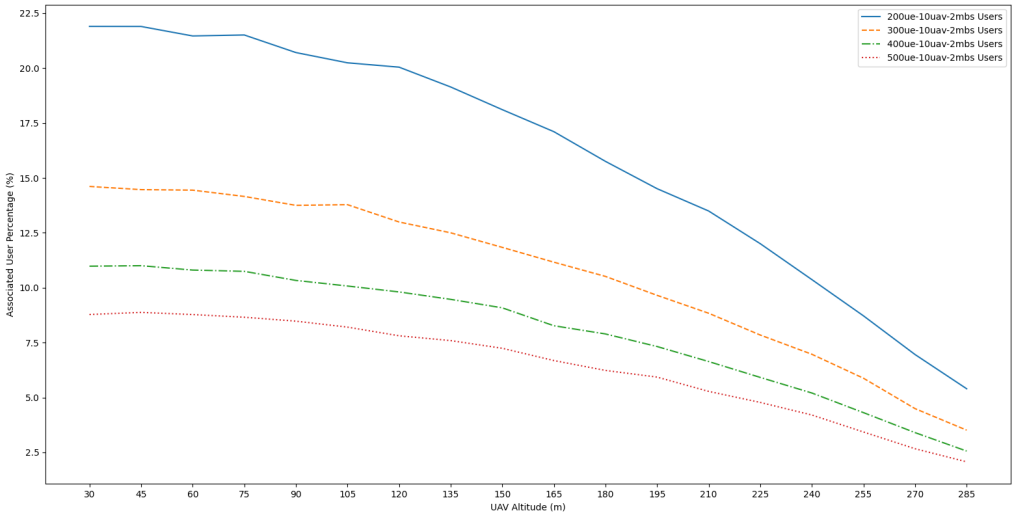


Figure 5.26: Associated user percentages with different UAV altitudes in RMa scenario and 28 GHz carrier frequency is shown.

5.5 Discussion

From the simulations, there are some lessons learned, and in this section, we summarize the comments made in the previous section.

In terms of the performance of proposed heuristics, all three heuristics achieve the optimal throughput that the MILP finds. The advantage in terms of execution time is also quite significant for RTF-H and RTF-E; these methods usually outperform MILP by 50 times or more. Furthermore, most of the time their execution times are less than 100 ms, meaning that they are very suitable for dynamic or even real-time networks and their operations. RTF-I does a worse job than these two heuristics in terms of execution time when it cannot find an optimal solution in the initial iterations, and because of this, RTF-H is often found to be better than RTF-I, since the best result is almost always the initial result found by RTF-H. Even though there are alternatives available, not only they do not result in an improvement, but since they also result in a worse network, the execution of RTF algorithm takes more time to allocate all flows. Nevertheless, the performance gains from the heuristics remain significant.

In terms of scenarios, RMa scenario is found to be a better benchmark than UMa scenario as it is less strict in terms of LoS probability and its effect on performance. 28 GHz carrier frequency is also better than 60 GHz carrier frequency in terms of performance, as higher frequencies often require dedicated and specific channel models and beamforming calculations. Currently, 28 GHz seem to be a better alternative as it also allows users and base stations to use the same frequency. This is not possible for 60 GHz as there are no antenna suitable for user devices in this frequency yet. Even so, both scenarios and frequencies are shown to be performing in a specific set of parameters, but careful planning is required to maximize the performance.

While the effects of different parameters on throughput is more evident on RMa scenario, the UMa scenario does not show this trend. We explained the reasoning behind this as the problems in deployment and user association. While the RMa scenario gives better and more realistic results with Monte Carlo simulations, UMa scenario is affected much more severely by deployment, and randomizing deployment in Monte Carlo simulations blurs the effects of the parameters that are changed between simu-

lations.

In terms of parameters, as we expected, UAV altitude and user demand are negatively correlated with throughput, whereas transmit power is positively correlated. In terms of transmit power, 30-35 dBm seem to be the sweet spot that enables solid performance for even low numbers of UAV-SCs and MBSs. In terms of UAV altitude, the performance drop seems to be sharper for 60 GHz as it is more susceptible to free space path loss and LoS probability with higher distances. Altitudes up until 150 m seem to be performing well after which we can observe a sharp performance drop. Finally, user demand seems to have a linear effect when the number of users are the same, but when these two numbers are both increased, since the increase in total throughput is multiplied, the effect of higher demand becomes more severe.

CHAPTER 6

CONCLUSION

In this chapter, we conclude this work and briefly discuss the possible future extensions.

6.1 Conclusion

In this thesis, we firstly present background information on cellular networks, wireless backhauling, aerial networks, flow networks, and maximum flow problem. Building on top of this information, we introduce a network model, on which we define the UAV-UAR problem. Firstly, we give a MILP formulation of this problem which is optimal, but has a large runtime.

Since our network assumes UAV's used as small cells, they are capable of 3D mobility, and the MILP based solution cannot handle it due to its large runtime requirement. For this purpose, we resort to the relabel-to-front method, and develop three distinct RTF-based heuristics for UAV-UAR problem. RTF-E employs the RTF method and then eliminates all but the connection edge from all user's outgoing edges. RTF-H employs a heuristic preflow initialization in which the user association constraints are conformed from the beginning, which removes the need to eliminate any edges that do not satisfy the aforementioned constraint. Finally, RTF-I employs RTF-H iteratively by trying alternative edges that have a higher upper bound than the achieved result.

Using Monte Carlo simulations, we show that RTF-based heuristics outperform the MILP-based optimization in terms of execution time, while achieving comparable

throughput results that are often within 5%. In terms of execution time, RTF-based methods outperform MILP solution by roughly 20 and 1000 times, at worst and at best, respectively. RTF-based solutions often have execution times less than 100ms, and these figures go only as high as 200ms, which are still suitable for dynamic networks, whereas MILP methods have exeuciton times as high as 30s in worst cases.

We present results in terms of four parameters, namely the user demand, UAV height, receiver sensitivity, and transmit power. As expected, user demand and UAV height are negatively correlated with total satisfied throughput as increasing the demand increases the link load whereas UAV height degrades the link capacity. Conversely, receiver sensitivity and transmit power are positively correlated with the satisfied throughput, albeit receiver sensitivity has only a minor impact. This is because links that are broken because of receiver sensitivity are often NLoS, which has minute capacities often less than 1 Mbps anyway. For transmit power, correlation is strong as it increases, link capacities also increase and new links can also be established.

To conclude, this thesis shows that the proposed RTF-based heuristics produce a valid solution for UAV-UAR problem. RTF-based methods have quite fast execution time, which makes them adequate to handle mobility. Furthermore, their performances are comparable to optimal, meaning that there is no performance loss while improving capabilities to handle mobility.

6.2 Future Work

First of all, while using Monte Carlo simulations for obtaining results, we have randomly generated topologies. Especially for UMa scenario, deployment is very important to significantly improve LoS probability. Chapter 2 mentions various works on deployment that aim to maximize network throughput by maximizing LoS. While our main focus is routing and user association, using a deployment scheme will significantly result in more performance, especially in cases where LoS probability is affected heavily with distance.

Another extension can be made in the air interface modeling. We assume fixed throughput per second, on which more realistic improvements can be made. There

are various scheduling schemes that govern the air interface, which can improve the realism and applicability of our work to cellular networks. Moreover, discrete event simulators can be employed to set up a fully capable 802.11ay [89] network. Not only this will improve the realism of simulations as these simulators fully implement network stacks on all network elements, but they also have other alternatives available, giving a myriad of possibilities to compare or test.

Finally, while our RTF-based heuristic methods are capable of handling mobility of both UEs and UAV-SCs, this is not tested in this work. As an extension, modeling mobility of network elements and showing the network performance under certain types of mobility can be made. However, to model the mobility, topology also has to be modeled as while UAV-SCs are flying and height topology is relatively easy to model, UEs often use roads and are more susceptible to interacting with each other, which requires significant consideration.

REFERENCES

[1] S. Chia, M. Gasparroni, and P. Brick, "The next challenge for cellular networks: backhaul," *IEEE Microwave Magazine*, vol. 10, no. 5, pp. 54–66, 2009.

[2] P. Briggs, R. Chundury, and J. Olsson, "Carrier ethernet for mobile backhaul," *IEEE Communications Magazine*, vol. 48, no. 10, pp. 94–100, 2010.

[3] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 101–107, 2011.

[4] 3GPP, "Evolved universal terrestrial radio access (e-utra); relay architectures for e-utra (lte-advanced)," Technical Report (TR) 36.806, 3rd Generation Partnership Project (3GPP), 04 2010. Version 9.0.0.

[5] 3GPP, "Evolved universal terrestrial radio access (e-utra); further advancements for e-utra physical layer aspects," Technical Report (TR) 36.814, 3rd Generation Partnership Project (3GPP), 03 2017. Version 9.2.0.

[6] 3GPP, "Evolved universal terrestrial radio access (e-utra); relay radio transmission and reception," Technical Report (TR) 36.826, 3rd Generation Partnership Project (3GPP), 07 2013. Version 11.3.0.

[7] A. Mourad, "Self-backhauling in 5G," in *5G-PPP Workshop on 5G RAN Design, Air Interface Design and Integration, Valencia, Spain*, 2016.

[8] J. Gamboa and I. Demirkol, "Softwarized lte self-backhauling solution and its evaluation," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, 2018.

[9] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. T. Sukhavasi, C. Patel, and S. Geirhofer, "Network densification: the dominant theme for wireless evolution into 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 82–89, 2014.

[10] X. Ge, S. Tu, G. Mao, C. Wang, and T. Han, "5G ultra-dense cellular networks," *IEEE Wireless Communications*, vol. 23, no. 1, pp. 72–79, 2016.

[11] N. Group *et al.*, "R1-113871: Study on pico coverage with cre offset," in *3GPP TSG RAN WG1 Meeting*, vol. 67.

[12] N. Docomo, "Performance of eicic with control channel coverage limitation," *R1-103264, 3GPP Std., Montreal, Canada*, vol. 5, p. 27, 2010.

[13] M. Mobility, "R1-124951: Maintenance of up-to-date sib1 in cre zone," in *3GPP TSG RAN WG1 Meeting*, vol. 79.

[14] S. Singh, M. N. Kulkarni, A. Ghosh, and J. G. Andrews, "Tractable model for rate in self-backhauled millimeter wave cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 10, pp. 2196–2211, 2015.

[15] M. S. Shokry, D. Ebrahimi, C. Assi, S. Sharafeddine, and A. Ghrayeb, "Leveraging uavs for coverage in cell-free vehicular networks: A deep reinforcement learning approach," *IEEE Transactions on Mobile Computing*, 2020.

[16] N. Cheng, F. Lyu, W. Quan, C. Zhou, H. He, W. Shi, and X. Shen, "Space/aerial-assisted computing offloading for iot applications: A learning-based approach," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 5, pp. 1117–1129, 2019.

[17] I. Bekmezci, O. K. Sahingoz, and Ş. Temel, "Flying ad-hoc networks (fanets): A survey," *Ad Hoc Networks*, vol. 11, no. 3, pp. 1254–1270, 2013.

[18] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Modeling air-to-ground path loss for low altitude platforms in urban environments," in *2014 IEEE global communications conference*, pp. 2898–2904, IEEE, 2014.

[19] D. Orfanus, E. P. de Freitas, and F. Eliassen, "Self-organization as a supporting paradigm for military uav relay networks," *IEEE Communications Letters*, vol. 20, no. 4, pp. 804–807, 2016.

[20] A. Azizi, S. Parsaeefard, M. R. Javan, N. Mokari, and H. Yanikomeroglu, "Profit maximization in 5G+ networks with heterogeneous aerial and ground base stations," *IEEE Transactions on Mobile Computing*, 2019.

[21] S. Jaffry, R. Hussain, X. Gui, and S. F. Hasan, "A comprehensive survey on moving networks," *IEEE Communications Surveys & Tutorials*, 2020.

[22] W. Wang, N. Cheng, Y. Liu, H. Zhou, X. Lin, and X. Shen, "Content delivery analysis in cellular networks with aerial caching and mmwave backhaul," *IEEE Transactions on Vehicular Technology*, 2021.

[23] X. Wu, Z. Wei, Z. Cheng, and X. Zhang, "Joint optimization of uav trajectory and user scheduling based on noma technology," in *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, IEEE, 2020.

[24] N.-N. Dao, Q.-V. Pham, N. H. Tu, T. T. Thanh, V. N. Q. Bao, D. S. Lakew, and S. Cho, "Survey on aerial radio access networks: toward a comprehensive 6g access infrastructure," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 1193–1225, 2021.

[25] Y. Cao, S.-Y. Lien, and Y.-C. Liang, "Deep reinforcement learning for multi-user access control in non-terrestrial networks," *IEEE Transactions on Communications*, vol. 69, no. 3, pp. 1605–1619, 2020.

[26] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, "A tutorial on uavs for wireless networks: Applications, challenges, and open problems," *IEEE communications surveys & tutorials*, vol. 21, no. 3, pp. 2334–2360, 2019.

[27] B. Khamidehi and E. S. Sousa, "Trajectory design for the aerial base stations to improve cellular network performance," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 1, pp. 945–956, 2021.

[28] G. K. Kurt, M. G. Khoshkholgh, S. Alfattani, A. Ibrahim, T. S. Darwish, M. S. Alam, H. Yanikomeroglu, and A. Yongacoglu, "A vision and framework for the high altitude platform station (haps) networks of the future," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 729–779, 2021.

[29] A. Aftab, N. Ashraf, H. K. Qureshi, S. A. Hassan, and S. Jangsher, "Block-ml: Blockchain and machine learning for uav-bss deployment,"

[30] Y. M. Park, Y. K. Tun, and C. S. Hong, "Optimized deployment of multi-uav based on machine learning in uav-hst networking," in *2020 21st Asia-Pacific*

Network Operations and Management Symposium (APNOMS), pp. 102–107, IEEE, 2020.

[31] J. Hu, H. Zhang, Y. Liu, X. Li, and H. Ji, "An intelligent uav deployment scheme for load balance in small cell networks using machine learning," in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, IEEE, 2019.

[32] J. Liu, Q. Wang, X. Li, and W. Zhang, "A fast deployment strategy for uav enabled network based on deep learning," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1–6, IEEE.

[33] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms*. MIT press, 2002.

[34] A. Bonfante, L. G. Giordano, D. Lopez-Perez, A. Garcia-Rodriguez, G. Geraci, P. Baracca, M. M. Butt, M. Dzaferagic, and N. Marchetti, "Performance of massive mimo self-backhauling for ultra-dense small cell deployments," in *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–7, 2018.

[35] J. Tang, A. Shojaeifard, D. K. So, K.-K. Wong, and N. Zhao, "Energy efficiency optimization for comp-swipt heterogeneous networks," *IEEE Transactions on Communications*, vol. 66, no. 12, pp. 6368–6383, 2018.

[36] H. Zhang, H. Liu, J. Cheng, and V. C. M. Leung, "Downlink energy efficiency of power allocation and wireless backhaul bandwidth allocation in heterogeneous small cell networks," *IEEE Transactions on Communications*, vol. 66, no. 4, pp. 1705–1716, 2018.

[37] T. M. Nguyen, A. Yadav, W. Ajib, and C. Assi, "Centralized and distributed energy efficiency designs in wireless backhaul hetnets," *IEEE Transactions on Wireless Communications*, vol. 16, no. 7, pp. 4711–4726, 2017.

[38] C. Saha, M. Afshang, and H. S. Dhillon, "Integrated mmwave access and backhaul in 5G: Bandwidth partitioning and downlink analysis," in *2018 IEEE International Conference on Communications (ICC)*, pp. 1–6, 2018.

[39] X. Huang, D. Zhang, S. Tang, Q. Chen, and J. Zhang, "Fairness-based distributed resource allocation in two-tier heterogeneous networks," *IEEE Access*, vol. 7, pp. 40000–40012, 2019.

[40] Q. Pham, L. B. Le, S. Chung, and W. Hwang, "Mobile edge computing with wireless backhaul: Joint task offloading and resource allocation," *IEEE Access*, vol. 7, pp. 16444–16459, 2019.

[41] Z. Song, Y. Liu, and X. Sun, "Joint radio and computational resource allocation for noma-based mobile edge computing in heterogeneous networks," *IEEE Communications Letters*, vol. 22, no. 12, pp. 2559–2562, 2018.

[42] Q. Han, B. Yang, G. Miao, C. Chen, X. Wang, and X. Guan, "Backhaul-aware user association and resource allocation for energy-constrained hetnets," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 1, pp. 580–593, 2017.

[43] N.-T. Le, L.-N. Tran, Q.-D. Vu, and D. Jayalath, "Energy-efficient resource allocation for ofdma heterogeneous networks," *IEEE Transactions on Communications*, vol. 67, no. 10, pp. 7043–7057, 2019.

[44] N. Omidvar, A. Liu, V. Lau, F. Zhang, D. H. K. Tsang, and M. R. Pakravan, "Optimal hierarchical radio resource management for hetnets with flexible backhaul," *IEEE Transactions on Wireless Communications*, vol. 17, no. 7, pp. 4239–4255, 2018.

[45] H. Gao, S. Zhang, Y. Su, and M. Diao, "Joint resource allocation and power control algorithm for cooperative d2d heterogeneous networks," *IEEE Access*, vol. 7, pp. 20632–20643, 2019.

[46] Y. Liu, L. Lu, G. Y. Li, Q. Cui, and W. Han, "Joint user association and spectrum allocation for small cell networks with wireless backhauls," *IEEE Wireless Communications Letters*, vol. 5, no. 5, pp. 496–499, 2016.

[47] Z. Song, Q. Ni, and X. Sun, "Distributed power allocation for nonorthogonal multiple access heterogeneous networks," *IEEE Communications Letters*, vol. 22, no. 3, pp. 622–625, 2018.

[48] D. Bojic, E. Sasaki, N. Cvijetic, T. Wang, J. Kuno, J. Lessmann, S. Schmid, H. Ishii, and S. Nakamura, "Advanced wireless and optical technologies for

small-cell mobile backhaul with dynamic software-defined management," *IEEE Communications Magazine*, vol. 51, no. 9, pp. 86–93, 2013.

[49] Y. Hu, Y. Xu, Y. Liu, and H. Yu, "Robust energy-efficiency power allocation in multicell hetnets with channel uncertainties," in *2018 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 426–430, IEEE, 2018.

[50] N. Faruk, K. Ruttik, E. Mutafungwa, and R. Jäntti, "Energy savings through self-backhauling for future heterogeneous networks," *Energy*, vol. 115, pp. 711–721, 2016.

[51] N. Wang, E. Hossain, and V. K. Bhargava, "Backhauling 5G small cells: A radio resource management perspective," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 41–49, 2015.

[52] U. Siddique, H. Tabassum, and E. Hossain, "Downlink spectrum allocation for in-band and out-band wireless backhauling of full-duplex small cells," *IEEE Transactions on Communications*, vol. 65, no. 8, pp. 3538–3554, 2017.

[53] M. Shariat, E. Pateromichelakis, A. u. Quddus, and R. Tafazolli, "Joint tdd backhaul and access optimization in dense small-cell networks," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 11, pp. 5288–5299, 2015.

[54] A. Khodmi, S. Benrejeb, and Z. Choukair, "Iterative water filling power allocation and relay selection based on two-hop relay in 5g/heterogeneous ultra dense network," in *2018 Seventh International Conference on Communications and Networking (ComNet)*, pp. 1–6, IEEE, 2018.

[55] W. Hao, M. Zeng, Z. Chu, S. Yang, and G. Sun, "Energy-efficient resource allocation for mmwave massive mimo hetnets with wireless backhaul," *IEEE Access*, vol. 6, pp. 2457–2471, 2018.

[56] A. Mahmood, M. L. M. Kiah, M. R. Z'aba, A. N. Qureshi, M. S. S. Kassim, Z. H. A. Hasan, J. Kakarla, I. S. Amiri, and S. R. Azzuhri, "Capacity and frequency optimization of wireless backhaul network using traffic forecasting," *IEEE Access*, vol. 8, pp. 23264–23276, 2020.

[57] T. Han and N. Ansari, "Green-energy aware and latency aware user associations in heterogeneous cellular networks," in *2013 IEEE Global Communications Conference (GLOBECOM)*, pp. 4946–4951, 2013.

[58] H. Ma, H. Zhang, X. Wang, and J. Cheng, "Backhaul-aware user association and resource allocation for massive mimo-enabled hetnets," *IEEE Communications Letters*, vol. 21, no. 12, pp. 2710–2713, 2017.

[59] P. Huskov, T. Maksymyuk, I. Kahalo, and M. Klymash, "Smart backhauling subsystem for 5G heterogeneous network," in *The Experience of Designing and Application of CAD Systems in Microelectronics*, pp. 481–483, 2015.

[60] Q. Zhang, W. Ma, Z. Feng, and Z. Han, "Backhaul capacity aware interference mitigation framework in 6g cellular internet of things," *IEEE Internet of Things Journal*, 2021.

[61] W. Pu, X. Li, J. Yuan, and X. Yang, "Resource allocation for millimeter wave self-backhaul network using markov approximation," *IEEE Access*, vol. 7, pp. 61283–61295, 2019.

[62] H. Dai, Y. Huang, J. Wang, and L. Yang, "Resource optimization in heterogeneous cloud radio access networks," *IEEE communications letters*, vol. 22, no. 3, pp. 494–497, 2017.

[63] Y. Xu, G. Gui, H. Gacanin, and F. Adachi, "A survey on resource allocation for 5g heterogeneous networks: Current research, future trends and challenges," *IEEE Communications Surveys & Tutorials*, 2021.

[64] M. Z. Chowdhury, M. Shahjalal, S. Ahmed, and Y. M. Jang, "6g wireless communication systems: Applications, requirements, technologies, challenges, and research directions," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 957–975, 2020.

[65] J. McMenamy, A. Narbudowicz, K. Niotaki, and I. Macaluso, "Hop-constrained mmwave backhaul: Maximising the network flow," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 596–600, 2019.

[66] A. Fouda, A. S. Ibrahim, I. Guvenc, and M. Ghosh, "Uav-based in-band integrated access and backhaul for 5G communications," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, pp. 1–5, 2018.

[67] T. Zhang, Y. Wang, Y. Liu, W. Xu, and A. Nallanathan, "Cache-enabling uav communications: Network deployment and resource allocation," *IEEE Transactions on Wireless Communications*, vol. 19, no. 11, pp. 7470–7483, 2020.

[68] R. Okumura and A. Hirata, "Automatic planning of 300-ghz-band wireless backhaul link deployment in metropolitan area," in *2020 International Symposium on Antennas and Propagation (ISAP)*, pp. 541–542, IEEE, 2021.

[69] M. Polese, M. Giordani, T. Zugno, A. Roy, S. Goyal, D. Castor, and M. Zorzi, "Integrated access and backhaul in 5G mmwave networks: Potential and challenges," *IEEE Communications Magazine*, vol. 58, no. 3, pp. 62–68, 2020.

[70] H. Dahrouj, A. Douik, F. Rayal, T. Y. Al-Naffouri, and M. Alouini, "Cost-effective hybrid rf/fso backhaul solution for next generation wireless systems," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 98–104, 2015.

[71] Y. Zhang, J. Liu, M. Sheng, Y. Shi, and J. Li, "Leveraging the coupling of radio access network and mmwave backhaul network: Modeling and optimization," *IEEE Transactions on Vehicular Technology*, 2021.

[72] M. N. Kulkarni, A. Ghosh, and J. G. Andrews, "Max-min rates in self-backhauled millimeter wave cellular networks," *arXiv preprint arXiv:1805.01040*, 2018.

[73] L. Zhang and N. Ansari, "Optimizing the deployment and throughput of dbss for uplink communications," *IEEE Open Journal of Vehicular Technology*, vol. 1, pp. 18–28, 2019.

[74] S. Saadat, D. Chen, and T. Jiang, "Multipath multihop mmwave backhaul in ultra-dense small-cell network," *Digital Communications and Networks*, vol. 4, no. 2, pp. 111–117, 2018.

[75] A. I. Nasr and Y. Fahmy, "Millimeter-wave wireless backhauling for 5G small cells: Scalability of mesh over star topologies," in *2017 IEEE 18th Interna-*

tional Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM), pp. 1–6, IEEE, 2017.

[76] A. Łukowa, V. Venkatasubramanian, E. Visotsky, and M. Cudak, "On the coverage extension of 5g millimeter wave deployments using integrated access and backhaul," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1–7, IEEE.

[77] W. Qu, G. Li, and Y. Zhao, "On the coverage problem in device-to-device relay networks," *IEEE Communications Letters*, vol. 23, no. 11, pp. 2139–2143, 2019.

[78] I. Singh and N. P. Singh, "Coverage and capacity analysis of relay-based device-to-device communications underlaid cellular networks," *Engineering Science and Technology, an International Journal*, vol. 21, no. 5, pp. 834–842, 2018.

[79] A. Sharma, R. K. Ganti, and J. K. Milleth, "Joint backhaul-access analysis of full duplex self-backhauling heterogeneous networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1727–1740, 2017.

[80] A. S. Khan, G. Chen, Y. Rahulamathavan, G. Zheng, B. Assadhan, and S. Lambotharan, "Trusted uav network coverage using blockchain, machine learning, and auction mechanisms," *IEEE Access*, vol. 8, pp. 118219–118234, 2020.

[81] M. Kamel, W. Hamouda, and A. Youssef, "Uplink coverage and capacity analysis of mmtc in ultra-dense networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 1, pp. 746–759, 2019.

[82] S. K. Zaidi, S. F. Hasan, X. Gui, N. Siddique, and S. Ahmad, "Exploiting uav as noma based relay for coverage extension," in *2019 2nd International Conference on Computer Applications & Information Security (ICCAIS)*, pp. 1–5, IEEE, 2019.

[83] Y. Jo, H. Kim, J. Lim, and D. Hong, "Self-optimization of coverage and system throughput in 5g heterogeneous ultra-dense networks," *IEEE Wireless Communications Letters*, vol. 9, no. 3, pp. 285–288, 2019.

[84] M. G. Khoshkholgh, K. Navaie, H. Yanikomeroglu, V. C. M. Leung, and K. G. Shin, "Coverage performance of aerial-terrestrial hetnets," in *2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring)*, pp. 1–5, 2019.

[85] S. M. Azimi-Abarghouyi, B. Makki, M. Haenggi, M. Nasiri-Kenari, and T. Svensson, "Coverage analysis of finite cellular networks: A stochastic geometry approach," in *2018 Iran Workshop on Communication and Information Theory (IWCIT)*, pp. 1–5, 2018.

[86] R. Borralho, A. Mohamed, A. U. Quddus, P. Vieira, and R. Tafazolli, "A survey on coverage enhancement in cellular networks: Challenges and solutions for future deployments," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 1302–1341, 2021.

[87] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz (Release 16)," Technical Report (TR) 38.901, 3rd Generation Partnership Project (3GPP), 12 2019. Version 16.1.0.

[88] Gurobi Optimization, LLC, "Gurobi Optimizer Reference Manual," 2021.

[89] "Ieee standard for information technology–telecommunications and information exchange between systems local and metropolitan area networks–specific requirements part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications amendment 2: Enhanced throughput for operation in license-exempt bands above 45 ghz," *IEEE Std 802.11ay-2021 (Amendment to IEEE Std 802.11-2020 as amendment by IEEE Std 802.11ax-2021)*, pp. 1–768, 2021.