



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

Master of Science

**Load Balancing in 5G mmWave small
cells for Integrated Access and Backhaul**

TRAN QUANG HUY

**Department of Electrical, Electronics and
Computer Engineering
University of Ulsan**

Load Balancing in 5G mmWave small cells for Integrated Access and Backhaul

A Thesis Submitted in Partial Fulfillment of the Requirements
for the degree of

Master of Science

by

TRAN QUANG HUY

Department of Electrical, Electronics and

Computer Engineering

University of Ulsan

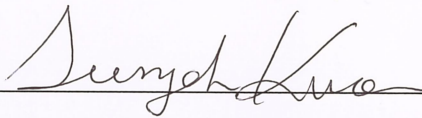
November 2023

**Load Balancing in 5G mmWave small cells for Integrated
Access and Backhaul**

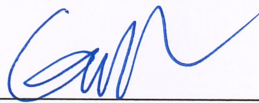
Approved by Supervisory Committee:



Prof. Sunghwan Kim, Chair



Prof. Sungoh Kwon, Supervisor



Prof. Hee-youll Kwak

Department of Electrical, Electronics and Computer Engineering

University of Ulsan

Date: November 2023

Vita

Tran Quang Huy received his Bachelor of Sciences (B.S.) in Mathematics and Computer Science from University of Sciences, Vietnam National University in Ho Chi Minh City, Vietnam in 2017. From 2017 to 2019, he was working as an IT Engineer at Aufinia Limited Company, in Ho Chi Minh City, Vietnam. From 2019 to 2020, he was employed in the E-Commerce company Shoppe. In March 2021, he started studying for a Master's degree in Electrical, Electronics, and Computer Engineering at University of Ulsan (UOU), South Korea, under the supervision of Professor Sungoh Kwon. His research interests include backhaul mmWave communication optimization and machine learning.

Acknowledgements

I would like to express my gratitude to my primary supervisor, Sungoh Kwon, who guided me throughout this project. With his kindness, continuous support, motivation, and patience, I can complete my M.S. degree. His immense knowledge and plentiful experience have encouraged me in all the time of my academic research and daily life. I also thankfully appreciate the other members of the committee of M.S. thesis defense, Professor Kim Sung-hwan, and Professor Kwak Hee-yeol for their visit and their valuable comments, which helped me greatly to improve the quality of the thesis.

I am also grateful to the members of the Communication and Networks Laboratory (ComNet) at University of Ulsan (UoU) for their friendship, enthusiastic help, and cheerfulness during my study. Especially, I would like to thank Dr.Minh-Thang Nguyen and Dr.Tho-Minh Duong for their technical support in my study. I would like to thank other members of the ComNet laboratory for the discussion and the Vietnamese community at UoU for their sharing.

Last but not least, I deeply show my appreciation to my friends and family for their support. Without them, it would not have been possible for me to achieve this milestone.

TRAN QUANG HUY

University of Ulsan,

Republic of Korea

June 2023

Abstract

Load Balancing in 5G mmWave small cells for Integrated Access and Backhaul

by

TRAN QUANG HUY

Supervised By: Professor Sungoh Kwon

Submitted in Partial Fulfillment of the Requirements for the degree of
Master of Science

Millimeter wave (mmWave) communication is a promising technique for keeping up with the exponential growth of next-generation networks, such as 5G, 5G advanced, and beyond future networks. To utilize the mmWave spectrum, these networks need a dense deployment of base stations (BSs) but it leads to high costs in wired infrastructure. Therefore, IAB has emerged as a cost-effective and flexible solution because it uses the multi-hop wireless backhaul to the core network via macro BS. However, multi-hop transmission and spectrum sharing between access and backhaul links cause an unbalanced

load across BSs in the IAB network.

In this thesis, we investigate the unbalanced load problem in the IAB network and propose a graph-based scheduling algorithm that increases the network throughput and achieves load distribution effectively. In the proposed algorithm, the weight of the graph is represented as the cost of the link between two BSs and is defined to reflect the cell load and link capacity. By utilizing the weight, we apply maximum weight-matching in graph optimization to estimate the forthcoming load and adjust the network transmission schedule, thereby achieving load balancing within the network. To validate the effectiveness of our proposed algorithm, we compare it with other algorithms in various environments through simulations. The results indicate that the proposed algorithm not only distributes the load across small cells more evenly but also low packet loss, increases network throughput, and improves the quality of experience at UEs.

Keywords: 5G mmWave, Integrated access and backhaul, Load balancing, Scheduling, Graph Theory

Contents

Vita	iii
Acknowledgements	iv
Abstract	vi
Contents	viii
List of Tables	x
List of Figures	xi
1 Introduction	1
1.1 Current States, Evolutions of 5G System and IAB Network	1
1.2 Scheduling and Related Works	4
1.3 Research Necessities	6
1.4 Thesis Organization and Contributions	9
2 IAB network and System model	10
2.1 IAB Network and In-band Communication	10

2.2	System model	11
3	Problem Formulation	14
4	Proposed Algorithm	18
4.1	The Directional Bipartite Weight Graph	19
4.2	Maximum Weighted Scheduling	20
5	Simulation and Result Analysis	26
5.1	Simulation Parameters	27
5.2	Impact of static UE on network performance	28
5.3	Impact of Network Load Variations	32
5.4	Impact of User Mobility Model	35
5.5	Impact of Heterogeneous Network	37
6	Conclusion	39
	References	41

List of Tables

5.1 Simulation parameter	26
------------------------------------	----

List of Figures

1.1	Expansion of 5G with IAB.	4
1.2	Illustration of the proposed load balancing with 1 donor and 3 IAB nodes with 1 user	7
2.1	An example of IAB network topology	11
4.1	Modeled the IAB network as a network graph	19
4.2	The IAB network modeled using a directional bipartite weight graph	21
4.3	The maximum weighted scheduling process	24
4.4	The schedule for resource allocation: $S = \{0 \rightarrow 1; 2 \rightarrow 3; 2 \rightarrow U_2\}$	25
5.1	Queue status in the network	29
5.2	Standard deviation	30
5.3	Performance under 3GPP constraints	31
5.4	Performance based on the number of UEs	33
5.5	Performance based on UE data rate requirements	34
5.6	Random freeway mobility model	35
5.7	Performance in the UEs mobility	36
5.8	Performance based under heterogeneous network	37

Chapter 1

Introduction

1.1 Current States, Evolutions of 5G System and IAB Network

Nowadays, the next generation of mobile broadband, 5G, has been commercially deployed in its early stages and has delivered significantly higher throughput than its predecessors, namely the fourth generation (4G) or Long-Term Evolution (LTE), and its variants, such as LTE-Advanced. The promised data rates can reach up to 10 Gbps [1], [2]. However, due to the COVID-19 pandemic, mobile data traffic with higher data rates and quality of service (QoS) has exponentially increased. The total global mobile data traffic is expected to reach 325 exabytes per month in 2028, more than six times the amount in 2021 [3], [4]. Consequently, mobile networks need to prepare for massive traffic growth over the upcoming years, driven by the increasing number of devices and advancements in information and communications technologies (ICT) applications.

To meet the increasing demand for data traffic, the introduction of network densifi-

cation with small cells has become crucial in fifth-generation (5G) networks, advanced 5G, and 6G [5], [6], [7], [8], [9]. Another key component driving 5G and 6G is millimeter wave (mmWave) communication, which offers substantial bandwidth and faster transmission speeds [10], [11]. However, a significant obstacle is the cost associated with establishing the necessary infrastructure for connecting the extensive number of base stations (BSs) to the core network via wired connections. This expense becomes particularly evident when considering the growing density of BSs. In order to address this challenge and fulfill the requirements, network densification in 5G and 6G wireless networks involves deploying numerous base stations of different types, resulting in more resource blocks per unit area.

To support backhauling in the network, various wired and wireless technologies have been developed [12]. Traditionally, wired backhaul has been predominantly used in urban and suburban areas with high data traffic, while wireless backhaul, such as satellite [13], microwave radio [14], or free space optical [15] links, have been employed in locations where deploying wired connections proves costly. It is worth noting that fiber optic technology offers high data rates, supporting several Gbps for long-haul transmissions [16]. However, the deployment of fiber optic networks entails significant capital expenditure costs, primarily attributed to trenching and digging, which not only increase installation times but also contribute to the overall expenses. Additionally, in certain cases, obtaining permissions for trenching may not be granted, particularly in metropolitan areas.

In ultra-dense networks with many BSs, wireless backhaul is a promising solution when compared with the other backhaul alternatives because of its financial efficiency, as well as its ability to provide better flexibility and mobility [17]. Due to this, mmWave-based wireless backhauling is considered an attractive replacement for fiber optic, which provides similar data rates with lower cost and simpler deployment effort. Hence, the 3rd

Generation Partnership Project (3GPP) promotes a new multi-hop wireless architecture, named Integrated Access and Backhaul (IAB), where part of the radio resources can be utilized for both access and backhaul network [18], [19]. This architecture has attracted considerable interest in both industry and academia for 5G, 5G advanced, and 6G network [20], [21], [22], [23]. The IAB approach involves deploying multiple relay BSs, referred to as IAB nodes, within the coverage area of a macro BS known as the IAB donor. These IAB nodes form a wireless backhaul network that facilitates packet forwarding between the IAB donor and user equipment (UE) [18]. The primary objective of IAB is to replace expensive backhaul technologies with a more flexible wireless backhaul solution that utilizes the existing spectrum. Furthermore, this approach enables the provision of normal cellular services within the same node, adding to its versatility. Figure.1.1 illustrates a use case of 5G expansion for dense networks using IAB.

In the context of 5G New Radio (NR), Integrated Access and Backhaul (IAB) is expected to achieve greater success, primarily due to the availability of massive bandwidth in the mmWave spectrum. The limited propagation range of mmWave access further reinforces the need for densified deployments, where IAB serves as a key enabler. Additionally, NR incorporates advanced beamforming capabilities and multiple-input-multiple-output (MIMO) techniques, which help mitigate cross-link interference between access and backhaul links. However, the multi-hop nature of the IAB network poses challenges in terms of traffic management and control [24], [25]. When only a small portion of the access and backhaul is scheduled for transmission most of the time, it can lead to various issues such as high latency, low throughput, and congestion in BSs. These problems significantly hinder the advantages offered by the mmWave spectrum. Therefore, efficient scheduling and resource allocation strategies are crucial to achieve high QoS at the user equipment (UE)

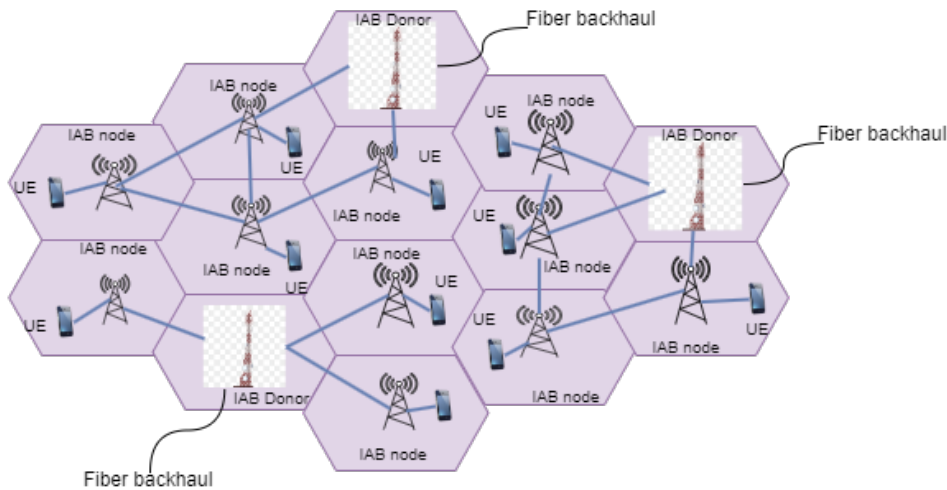


Figure 1.1: Expansion of 5G with IAB.

and maximize the overall performance of the IAB network [17], [24].

1.2 Scheduling and Related Works

Scheduling in wireless networks refers to the process of allocating and managing communication resources, such as bandwidth, or packet, among multiple users or devices within the network. It involves determining when and how users can transmit and receive data over the shared wireless medium. Wireless networks face various challenges due to limited resources, interference, and the shared nature of the wireless medium. Scheduling plays a crucial role in optimizing the utilization of these resources, improving network performance, and ensuring fair access for all users. It aims to maximize the overall system capacity, minimize delays, enhance throughput, and provide QoS guarantees. To achieve such a goal, all kinds of scheduling algorithms have been proposed to align with different network scenarios and purposes.

The IAB architecture is a relatively new research area, and thus, only a few works have

been recently studied. A joint resource allocation between backhaul and access links in time division duplexing (TDD), decomposing it into power and per-link sub-channel allocation was studied [26]. Considering small-cell grouping and resource slicing, a novel algorithm was proposed for rate balancing as well as improving network performance. The authors in [27] provided primitive research on the feasibility of IAB networks as part of a cellular network solution using mmWave communication. They demonstrated that cellular network systems employing IAB and mmWave transmissions require fewer fiber connections to achieve baseline performance, promising high cost-efficiency and ease of deployment for 5G cellular network systems. This investigation was extended in [28], evaluating the performance gains achieved through the use of IAB and non-IAB under different UE rates using joint fixed resource allocation and cost deployment. In [29], the authors explored the bandwidth partition scenarios between the access and backhaul side of the network to derive the downlink rate coverage in one-hop transmission. They studied two strategies (fixed partition and proportional partition) and proved that the proportional partition strategy provided a more load-balanced coverage rate than the fixed partition strategy. Their research was extended by considering a blockage model in the model and proposed an analysis to balance the rate in the network [30]. The authors in [31] concentrated solely on the access link in scheduling, proposing the minimum throughput algorithm needed to satisfy the QoS at UE. However, their works only considered one hop and ignored the backhaul, potentially leading to congestion due to overload in the relay BS.

Several studies have investigated resource allocation in multi-hop transmission within the context of the IAB network. In one study, the author assumed zero interference between any pair of links and proposed optimal scheduling to maximize fairness in link capacity [32]. They extended their research by considering the Quality of Experience (QoE) at UE

in the model, where UE satisfaction is the satisfied UE with the data rate required. An edge coloring scheduling in the backhaul link was proposed to maximize UE satisfaction in the IAB network. Utility maximization in multi-hop transmission has been addressed in [33], [34], [35] and [36]. For example, one study explored joint path selection and scheduling problems as a network utility maximization subject to queue stability and satisfy delay requirements at UE [33]. The work in [34] examined joint power allocation and interference constraints to maximize the data rate at the UE. Additionally, resource allocation based on different constraints is studied to achieve maximum efficiency [35]. On the other hand, the authors in [36] investigated queue flow in a spanning tree graph, proposing a backpressure-based approach to achieve queue stability and introducing a linear algorithm for maximum flow [36]. However, these works generally focused on fixed flows and ignored awareness of both network congestion and the link capacity that reduces the UEs' satisfaction.

1.3 Research Necessities

The introduction of the IAB network in 3GPP Release 16 has sparked significant interest in various research studies. However, the adoption of a multi-hop topology in IAB networks has presented several challenges in terms of scheduling and resource allocation. Given incoming traffic from the IAB donor to UEs through multiple IAB nodes, the use of IAB networks has drastically changed the cellular network model from direct connections between base stations and users to indirect connections via relay stations. This shift highlights the importance of examining the behavior of IAB networks in light of this network model transformation. Additionally, it is crucial to investigate efficient scheduling mechanisms that enable effective connectivity between stations and users within the IAB

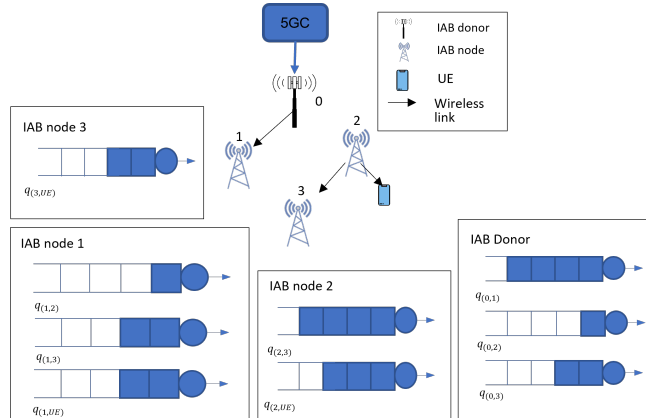


Figure 1.2: Illustration of the proposed load balancing with 1 donor and 3 IAB nodes with 1 user

networks.

To overcome the inefficiency of scheduling, we study load balancing through scheduling in the IAB network. Load balancing is the process of distributing network traffic or workload evenly across multiple BSs to achieve higher efficiency for utilizing network resources and ensure efficient delivery of data to wireless devices. Figure 1.2 shows an illustration of our example of balancing a load with one IAB donor and three IAB nodes from scheduling in a single time slot. Each node maintains a queue length for each destination, whether in access or backhaul. For example, the IAB donor 0 has three queues in backhaul to other IAB nodes: $q_{(0,1)}$, $q_{(0,2)}$, $q_{(0,3)}$. IAB node 1 has two queues in backhaul to IAB node 2 and 3 ($q_{(1,2)}$, $q_{(1,3)}$), and one queue serving to UEs $q_{(1,UE)}$. IAB node 2 has one queue in backhaul to IAB node 3 $q_{(2,3)}$, and one queue serving to UEs $q_{(2,UE)}$. IAB node 3 has only one queue to serve UEs $q_{(3,UE)}$. During the unit time slot, both the IAB donor and IAB node 2 exhibit the highest load compared to the other nodes. As a result, the IAB donor selects a scheduling set where it transmits to IAB node 1 in the backhaul, IAB node 2 transmits to the UE in access, and IAB node 2 further transmits to IAB node 3 in the

backhaul.

The 3GPP considers two topologies for the IAB network for scheduling: the spanning tree and the directed acyclic graph (DAG). A spanning tree is a subgraph of a connected graph that includes all the vertices of the original graph and forms a tree, while a DAG is a directed graph that does not contain any directed cycles. In this network, we opt to adapt the DAG topology because the spanning tree has limited connections, which restricts the flexibility of the IAB network.

This thesis introduces a novel graph-based load-balancing algorithm designed specifically for the IAB network in mmWave communication. Our approach involves modeling the scheduling problem of the 5G mmWave IAB network using a DAG. In order to capture the costs within the graph theory, we consider both the queue length and the capacity of BSs in the network. Our algorithm considers the anticipated load of each BS and applies a maximum weighted matching scheduling algorithm to target overloaded cells in both the access and backhaul link. The algorithm estimates the forthcoming load of each BS and targets overloaded cells in both access and backhaul. By redistributing the load from overloaded cells to neighboring cells and user equipment (UE), it minimizes the overall load and ensures load balance throughout the network. To evaluate the performance of our proposed algorithm, we conduct simulations in diverse environments and compare the results with several existing approaches. The simulation outcomes clearly demonstrate the effectiveness of our algorithm in achieving load balance. Our algorithm outperforms other schemes in terms of average throughput, packet drop rate, and QoE for the user, thereby confirming its superiority and viability. For consistency, we employ the notation that the BS connected to the core network using a wired connection is referred to as the IAB donor, while the relay BS, which establishes a wireless connection with the IAB donor or

neighboring relay BS, is referred to as the IAB node.

1.4 Thesis Organization and Contributions

The thesis is divided into six chapters which are described in the following:

Chapter I: this current chapter represents a brief introduction to 5G, 5G advanced, 6G, the IAB network, scheduling and load balancing, related works, research questions, and contributions to this study.

Chapter II: we describe the system model from the IAB network.

Chapter III: we present the measurement of the load in the model and formulate the load balancing problem in the IAB network.

Chapter IV: we provide the proposed algorithm to solve the problem formulation.

Chapter V: we conduct the simulation of the proposed algorithm and compare it with previous works in various environments.

Chapter VI: Finally, the conclusion and future work are described in this chapter.

Chapter 2

IAB network and System model

2.1 IAB Network and In-band Communication

We consider a multi-hop downlink mmWave IAB network consisting of one IAB donor and N IAB nodes. The network starts with a graph where the IAB donor node has a wired connection to the core network, and IAB nodes communicate wirelessly with the core network via the IAB donor. The IAB node establishes links with neighboring nodes in the backhaul connection and serves UEs in the access link. Fig. 2.1 shows an example IAB network architecture with one IAB donor 0 and five IAB nodes from 1 to 5. A solid arrow represents the backhaul link from transmitting node to receiving node and dashed arrows represent access links from a node to the UEs.

The IAB network has flexible transmitter-receiver relationships. The transmitter can be the IAB donor or the IAB node, while the receiver can be either the UEs or the recipient IAB node. In-band backhauling communication operates in a half-duplex scheme where the node can only transmit or receive in a given time slot. In a single time slot, if the

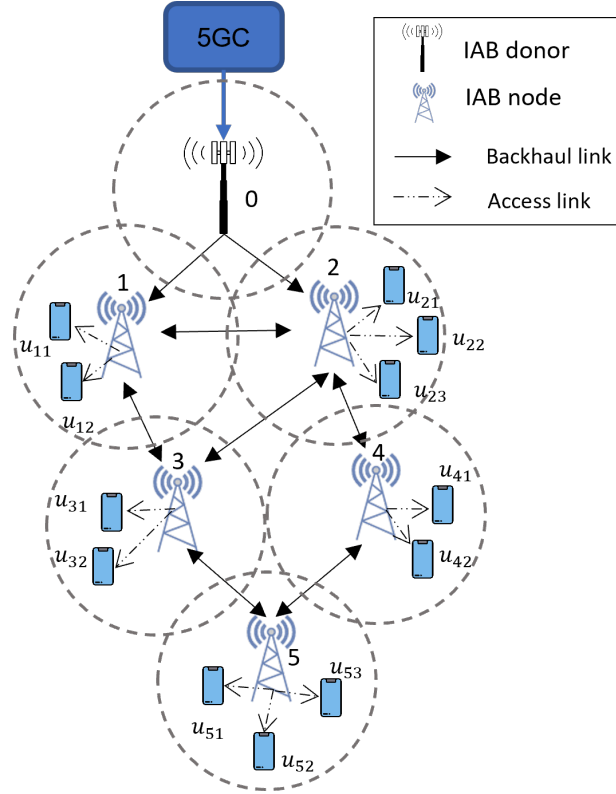


Figure 2.1: An example of IAB network topology

node is decided as a transmitter, that node can transmit data to one neighbor node in the backhaul and to UEs in the access link using frequency-division multiple access (FDMA) [37].

2.2 System model

Downlink IAB network G is composed of four elements $(\mathcal{N}, \mathcal{U}, E, \mathcal{Q})$. \mathcal{N} is the set of nodes, including IAB donor 0 and N IAB nodes numbered 1 to N . \mathcal{U} comprises all UEs in the network, such that $\mathcal{U} = \{U_1, U_2, \dots, U_N\}$, where U_i is a set of UEs for IAB node i . E is the set of directed edges, where (i, j) defines a wireless link from node $i \in \mathcal{N}$ to node

$j \in \mathcal{N} \setminus \{0, i\}$. $\mathcal{Q} = \{\{Q_{(i,j)}\}, Q_{(i,U_i)}\}$ is the set of queues in which $Q_{(i,j)}$ is the queue of edge (i, j) for the backhaul link and $Q_{(i,U_i)}$ is the queue for the access link of node i . Additionally, each queue has a buffer for packet queuing with a maximum queue length of q_{\max} . If the length of the queue reaches the maximum queue length q_{\max} , any incoming packets are discarded, resulting in dropped packets.

The IAB node is equipped with an antenna array for directional beamforming, which enables high array gain and transmission of constant power density over a bandwidth B . All nodes have M RF chains and conduct beamforming to downlink simultaneously up to M . In mmWave frequencies, beamforming significantly reduces interference between concurrent links owing to the short communication range and directional properties. Therefore, we assume that there is no interference between concurrent links, and the signal-to-noise ratio (SNR) is defined as explained in [38]

$$\eta_{(i,j)} = \frac{p_{(i,j)}}{N_0}, \quad (2.1)$$

where $p_{(i,j)}$ is the power received at node j from node i and N_0 is noise power. Consequently, link capacity is divided into two components (the backhaul link and the access link) with restrictions that are not exceeded by the number of supporting RF chains [39]. To define the link capacity on both links, we consider bandwidth allocation for access link and backhaul link based on the demand by UEs or a neighbor node. Bandwidth allocation for the backhaul link is

$$B_{(i,j)} = B \frac{q_{(i,j)}}{q_{(i,j)} + q_{(i,U_i)}}, \quad (2.2)$$

and bandwidth allocation for the access link is

$$B_{(i,U_i)} = B \frac{q_{(i,U_i)}}{q_{(i,j)} + q_{(i,U_i)}}. \quad (2.3)$$

The link capacity from node i to node j for the backhaul link is determined as

$$c_{(i,j)} = B_{(i,j)} \log_2(1 + \eta_{(i,j)}), i \in \mathcal{N}, j \in \mathcal{N} \setminus \{0, i\}, \quad (2.4)$$

and for the access link, it is

$$c_{(i,U_i)} = \sum_{u \in U_i} B_{(i,U_i)} \log_2(1 + \eta_{(i,U_i)}), i \in \mathcal{N} \setminus \{0\}. \quad (2.5)$$

The link capacity at node i is calculated with

$$c_i = c_{(i,j)} + c_{(i,U_i)}. \quad (2.6)$$

In a time slot t , node i transmits packets to node j in the backhaul link with the number of delivered packets, which is defined as

$$r_{(i,j)}(t) = \frac{c_{(i,j)}(t) \Delta t}{p}, \quad (2.7)$$

where Δt is the time slot duration and p is the packet size as specified in [40], [41].

Similarly, the number of delivered packets to UEs at node i in the access link is

$$r_{(i,U_i)}(t) = \frac{c_{(i,U_i)}(t) \Delta t}{p}. \quad (2.8)$$

The number of packets transmitted from node i to node j in the backhaul link is

$$\mu_{(i,j)}(t) = \min(q_{(i,j)}(t), r_{(i,j)}(t)), \quad (2.9)$$

and for the access link, the number of packets transmitted to UEs at node i is expressed as

$$\mu_{(i,U_i)}(t) = \min(q_{(i,u)}(t), r_{(i,U_i)}(t)). \quad (2.10)$$

The total number of packets transmitted by node i is

$$\mu_i(t) = \mu_{(i,j)}(t) + \mu_{(i,U_i)}(t). \quad (2.11)$$

Chapter 3

Problem Formulation

Our objective in this paper is to balance the unevenly distributed load of the whole network. In the IAB network, the mobility and data demand from UEs and the randomness of the network topology lead to some nodes having a high load while others have a small load. The result is network congestion at some high-load nodes and low quality of service (QoS) for the UEs. To optimize performance of the network through load balancing, we use the sum of queue length for both backhaul and access links as a measure of node load. For a given time slot t , $q_i(t)$ is the load of node i in time slot t , and the average load of all nodes in the network at the time slot t is defined as

$$\bar{q}(t) = \frac{1}{N} \sum_{i \in \mathbb{N}} q_i(t),$$

with standard deviation of the load on the network in time slot t expressed as

$$\sigma_q(t) = \sqrt{\frac{\sum_{i \in \mathcal{N}} |q_i(t) - \bar{q}(t)|^2}{N}}.$$

In each time slot t , the donor sets up a schedule set, S , for resource allocation in the IAB network. For schedule set S , one node acts as transmitter or receiver (excluding the

donor, which only works as a transmitter or does nothing). If one node is the transmitter, it can transmit not only to one neighbor node but also to UEs under that node's coverage. Hence, when node i is scheduled to transmit to node j , queue length $q_{(i,j)}$ in time slot $t + 1$ is defined as

$$q_{(i,j)}(t + 1) = \max(q_{(i,j)}(t) - \mu_{(i,j)}(t), 0), \quad (3.1)$$

where $\mu_{(i,j)}(t)$ is the number of packets transmitted from node i to node j at time t . Similarly, for the access link being served, queue length $q_{(i,U_i)}$ for node i is computed as

$$q_{(i,U_i)}(t + 1) = \max(q_{(i,U_i)}(t) - \mu_{(i,U_i)}(t), 0), \quad (3.2)$$

where $\mu_{(i,U_i)}(t)$ is the number of packets transmitted to UEs at node i at time t . Then, the load on node i in time slot $t + 1$ is

$$q_i(t + 1) = q_{(i,j)}(t) + q_{(i,U_i)}(t). \quad (3.3)$$

To balance an unevenly distributed load for the whole network, we formulate the problem of load balancing as reducing the standard deviation

$$\begin{aligned} \min_S \sigma_q(t + 1), \\ \text{s.t. } 0 \leq q_i(t + 1) \leq q_{\max}, \forall i \in \mathcal{N} \end{aligned} \quad (3.4)$$

where q_{\max} is the maximum queue length for each node.

The associations from the IAB donor to the IAB node, or from the IAB node to neighboring IAB nodes and UEs in the IAB network can be modeled using a bipartite graph. The bipartite graph G_b consists of two sets of nodes (a source set \mathcal{N} and a destination set V) with the set of edges E_b connecting the source set to the destination set. A semi-matching in G_b is a set of edges, $M \subseteq E_b$, such that no two edges in M share a common

node. In other words, if a node on an edge is chosen for semi-matching, that node is not chosen again. The weight for a semi-matching is the sum of weights $w(e)$ for edges e in semi-matching

$$\sum w(e), e \in M,$$

where weight $w(e)$ is a cost metric of edge e . Finding a semi-matching M that has the highest possible total weight maximizes the fairness of the chosen edge in the graph, which means minimum variance of the loads [42]. Hence, to solve the IAB network load problem, we need to utilize semi-matching from the bipartite graph.

In the IAB network, we modeled graph G as a bipartite graph, $G_b = (\mathcal{N}, V, E_b)$, where \mathcal{N} is the set of nodes and V is the set of nodes such that $V = \mathcal{N} \setminus \{0\}$. Each edge (i, j) in edge set E_b is a directional connection from node $i \in \mathcal{N}$ to node $j \in \mathcal{N} \setminus \{0\}$. The weight for a connection from node i to node j in G_b is defined for each time slot t as

$$\omega_{(i,j)}(t) = \begin{cases} \min(q_{(i,j)}(t), r_{(i,j)}(t)), & \text{for } i \neq j \\ 0, & \text{for } i = j \end{cases} \quad (3.5)$$

$$(3.6)$$

and the weight from node i to UEs under that node's coverage in each time slot t is

$$\omega_{(i,U_i)}(t) = \min(q_{(i,U_i)}(t), r_{(i,U_i)}(t)). \quad (3.7)$$

When scheduling for resource allocation in the IAB network, each node considers both access and backhaul links. To emphasize the selection of a node's role, we define weight $W_{(i,j)}(t)$ of each edge by adding the weight for the access link and backhaul link

$$W_{(i,j)}(t) = \omega_{(i,U_i)}(t) + \omega_{(i,j)}(t). \quad (3.8)$$

Following the semi-matching problem, we define semi-matching as a set of edges $S \subset E_b$ such that no two edges in S share a common node. Edge $(i, j) \in S$ is referred to as

a matching edge, whereas edge $(i, j) \notin S$ is a non-matching edge. Under the bipartite graph for the IAB network, we find a set of matching edges, which have the largest total maximum weight. Therefore, the problem of reducing the standard deviation in (3.4) can be reformulated as finding a semi-matching $S \subseteq E_b$ that has the maximal total weight from a graph G_b .

$$\max \sum_{(i,j) \in S, S \subseteq E_b} W_{(i,j)},$$

Chapter 4

Proposed Algorithm

This section presents an algorithm for load balancing in 5G IAB small-cell mmWave networks, which adapts to the network load status via maximum weight matching. Starting with a downlink IAB network graph as shown in Fig. 4.1a, we model the IAB network as network graph G including a set of nodes, \mathcal{N} , a family of UEs, \mathcal{U} , a set of directed edge, E , and a set of queues, \mathcal{Q} . The proposed algorithm calculates the directional link capacity from (2.4) and (2.5). The directional bipartite weight graph (Algorithm 2) models the bipartite graph from network graph G and computes the weight of each edge based on the load distribution of the whole network. Then, maximum weighted scheduling (Algorithm 3) constructs the resource allocation from the directional bipartite weight graph based on maximum weighted matching. Algorithm 1 summarizes the main algorithms and the details of the other two algorithms are described in the following subsections.

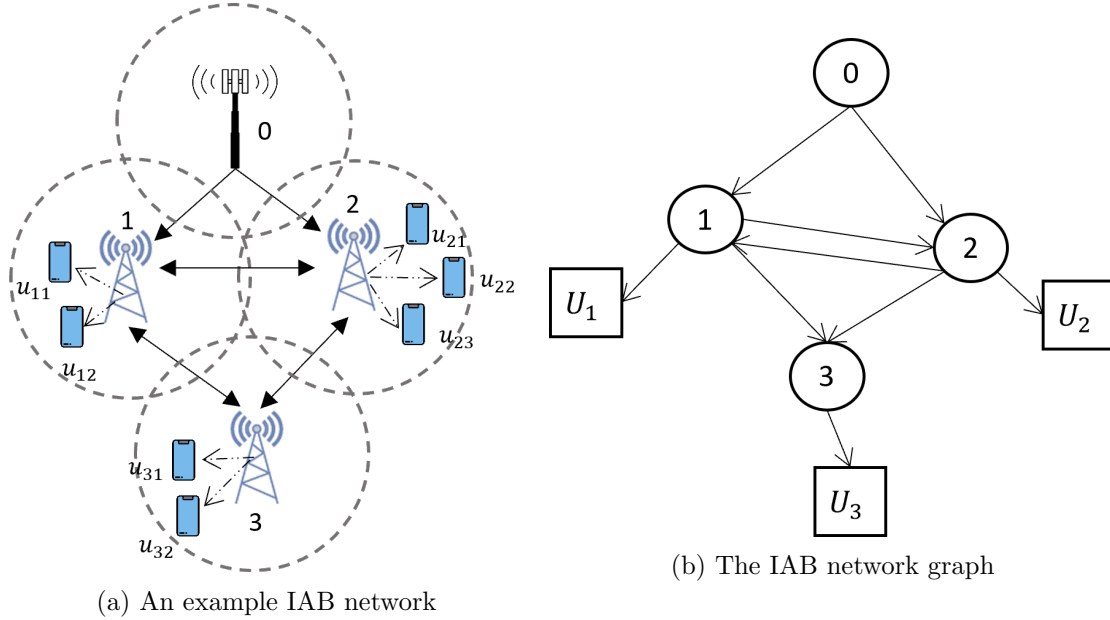


Figure 4.1: Modeled the IAB network as a network graph

Algorithm 1 Overall Algorithm

- 1: Input the IAB network as graph G including $(\mathcal{N}, \mathcal{U}, E, \mathcal{Q})$
 - 2: Calculate link capacity of node i, j where $i \in \mathcal{N}, j \in \mathcal{N} \setminus \{0, i\} \cup \mathcal{U}$
 - 3: Directional bipartite weight graph (Algorithm 2)
 - 4: Resource allocation from directional bipartite weight graph using maximum weight matching (Algorithm 3)
-

4.1 The Directional Bipartite Weight Graph

For a directional bipartite weight graph, the algorithm models the bipartite graph from network graph G and calculates the weight ω based on queue lengths and the link capacity for each pair of nodes in backhaul link or from a node to UEs on the access link. The weight serves as the metric for measuring the IAB network load. Algorithm 2 summarizes the directional bipartite weight graph.

First, the Algorithm 2 initializes a set of edge E_b , which is the same as E and is used to create a directional bipartite graph. At each node $i \in \mathcal{N} \setminus \{0\}$, the algorithm adds the edge

Algorithm 2 Directional Bipartite Weight Graph

```
1:  $E_b \leftarrow E$ 
2: for each node  $i \in \mathcal{N} \setminus \{0\}$  do
3:    $E_b \leftarrow E_b \cup (i, i)$ 
4: end for
5: for each edge  $(i, j) \in E_b$  do
6:   if  $i \neq j$  then
7:      $\omega_{(i,j)}(t) \leftarrow \min(q_{(i,j)}(t), r_{(i,j)}(t))$ 
8:   else
9:      $\omega_{(i,i)}(t) \leftarrow 0$ 
10:     $\omega_{(i,U_i)}(t) \leftarrow \min(q_{(i,U_i)}(t), r_{(i,U_i)}(t))$ 
11:   end if
12: end for
```

(i, i) to set E_b . Then, the algorithm employs a loop for each edge $(i, j) \in E_b$ and assigns the weight to the access link and the backhaul link based on the specific conditions. If node i is not equal to node j , the weight $\omega_{(i,j)}$ for the backhaul link is computed using (3.5). On the other hand, the algorithm calculates the weight $\omega_{(i,j)}$ for the backhaul link by using (3.6) and the weight $\omega_{(i,U_i)}$ for access link to UE by using (3.7).

Fig.4.2a shows a bipartite graph derived from Fig.4.1b, where nodes on the left side establish directional links with nodes or UEs on the right side. All the left-side nodes are the transmitters, while all the right-side nodes and UEs are the receivers. Then, Fig. 4.2b illustrates a directional bipartite weight graph in which each directional link has weight ω to specify the load from the node on the left side to the node or UEs on the right side.

4.2 Maximum Weighted Scheduling

The maximum weighted scheduling algorithm (Algorithm 3) is designed to determine resource allocation from the directional bipartite weight graph (Algorithm 2) via maximum total weighted matching.

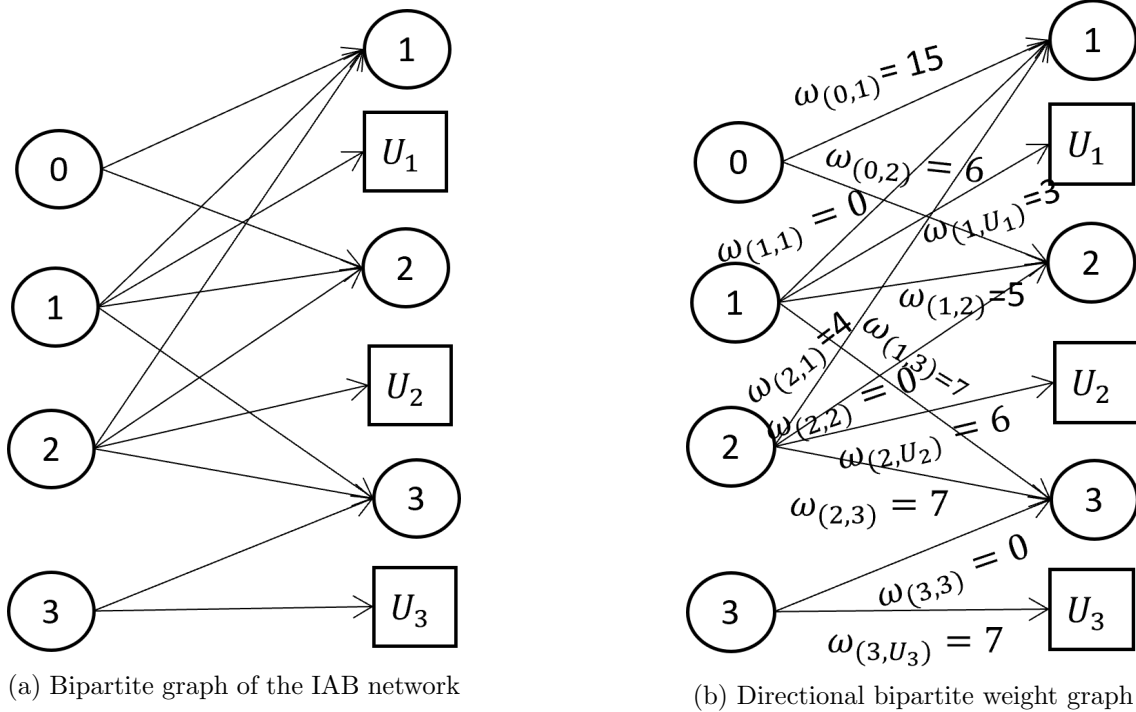


Figure 4.2: The IAB network modeled using a directional bipartite weight graph

First, the Algorithm initializes an empty weight matching schedule set S and a candidate set of edge E' that can be added to S . Before deciding the node's responsibility as transmitter or receiver for resource allocation in the IAB network, we calculate a new weight, $W_{(i,j)}(t)$, for each edge (i, j) in the set of edges E' by adding the weights of the access link and backhaul link used in (3.8). For example, from Fig.4.2b, weight $W_{(1,3)}$ is $\omega_{(1,3)} + \omega_{(1,U_1)} = 7 + 3 = 10$. Then, we employ a while loop on E' in which the algorithm stops if E' is an empty set.

From line 7 to line 13, the algorithm checks each potential direction $(m, l) \in E'$, where (m, l) represents a candidate edge to add to the matching schedule set S . In other words, the algorithm selects the highest $W_{(m,l)}(t)$ and adds edge (m, l) to S as a matching for resource allocation. Then, all the direction links originating from m and l are removed in

the set E'

$$E' = E' \setminus (m, \cdot) \setminus (l, \cdot), \quad (4.1)$$

where (m, \cdot) and (l, \cdot) denote all directional links from nodes m and l , respectively. The algorithm continues to find an (m, l) that pairs with the highest weight until E' is empty. The result is the schedule set S to allocate resources using maximum total weighted matching.

Fig.4.3 illustrates the maximum weighted scheduling algorithm for the directional bipartite weight graph in Fig.4.2b. In the first iteration in Fig.4.3a, the algorithm finds highest weight $W_{(0,1)}$ (solid line arrow) among those weights and then adds $(0, 1)$ to S . All directional links from IAB donor 0 and IAB node 1 (the dashed arrow) are considered for removal. After the first iteration, there are three remaining directional links (the dotted arrow), including IAB node 2 to IAB virtual node 2 and IAB node 3 with serving UEs. Another is IAB node 3 serving UEs under node 3's coverage, as shown in Fig.4.3b. Fig.4.3c shows the next iteration, the algorithm chooses directional links from IAB node 2 to IAB node 3 and to UEs under IAB node 2's coverage because of the highest weight, and then adds $(2, 3)$ with $(2, U_2)$ to S . Then, the links from IAB node 2 to virtual node 2 and from IAB node 3 to UEs are removed. Finally, there is no directional link to find and the algorithm stops with the resource allocation $S = \{(0, 1), (2, U_2), (2, 3)\}$ in Fig.4.3d. Fig. 4.4 demonstrates the final resource allocation including three transmitter-receiver pairs. One transmitter-receiver pair is from the donor node 0 to IAB node 1 and one transmitter-receiver pair is from IAB node 2 to IAB node 3 with UEs beneath the IAB node 2.

Algorithm 3 Maximum Weighted Scheduling (MWS)

```
1:  $S \leftarrow \emptyset$ 
2:  $E' \leftarrow E_b$ 
3: for  $(i, j) \in E'$  do
4:    $W_{(i,j)}(t) \leftarrow \omega_{(i,j)}(t) + \omega_{(i,U_i)}(t)$ 
5: end for
6: while  $E' \neq \emptyset$  do
7:   for  $(m, l) \in E'$  do
8:     Choose the highest  $W_{(m,l)}(t)$ 
9:      $S \leftarrow S \cup (m, l)$ 
10:    Remove all direction links from  $m$  and  $l$  as in (4.1)
11:   end for
12: end while
13: Return  $S$ 
```

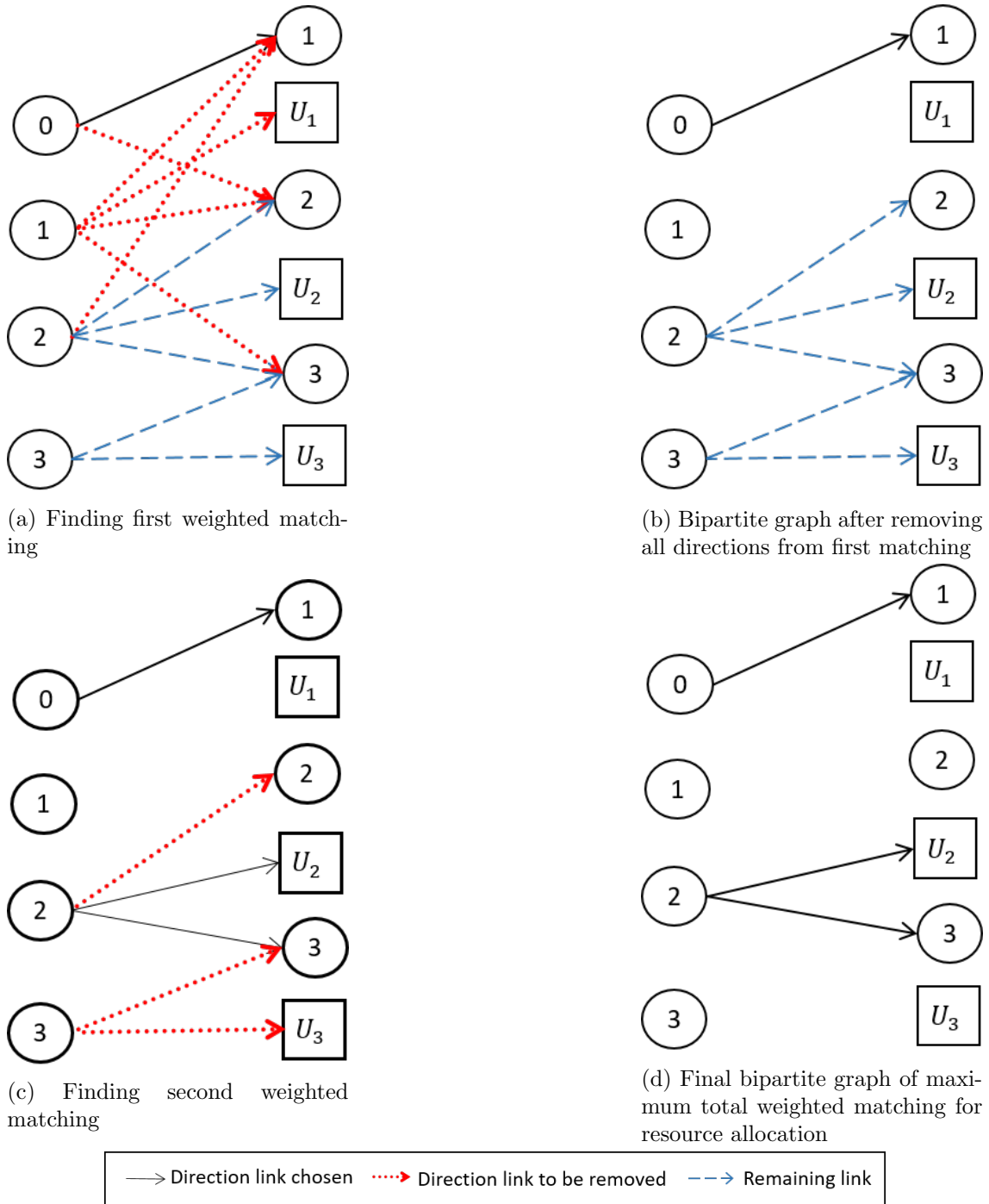


Figure 4.3: The maximum weighted scheduling process

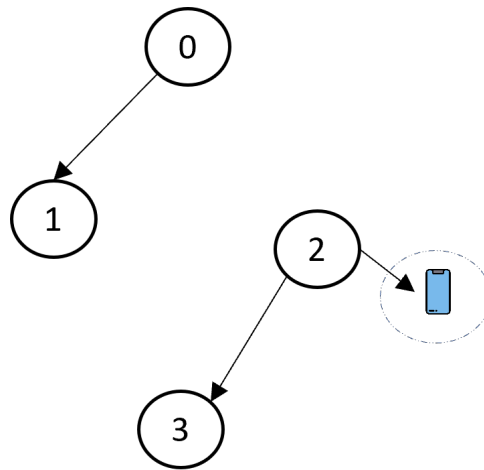


Figure 4.4: The schedule for resource allocation: $S = \{0 \rightarrow 1; 2 \rightarrow 3; 2 \rightarrow U_2\}$

Chapter 5

Simulation and Result Analysis

In this chapter, we describe the simulation parameters of the network with various environments. After that, we discuss the result between our proposed algorithm and other algorithms in each environment.

Table 5.1: Simulation parameter

Parameter	Value
Number of nodes (include IAB donor node)	6
Carrier frequency	28 GHz
Bandwidth	1 GHz
Transmission power	30 dBm
Directivity gain	30 dB
Number of RF chains	8
Node's buffer maximum	100
Random blockage	0.2 - 0.4

5.1 Simulation Parameters

To evaluate the performance of the proposed algorithm, we simulated the mmWave IAB network specified in [37] consisting of an IAB donor and five IAB nodes with transmission power of 30 dBm. The distance between two nodes for backhaul was set in a uniform range between 80 and 200 m, while the distance between nodes and UEs ranged between 0 and 50 m. The simulation operated at a frequency of 28 GHz with 1 GHz bandwidth. The path loss was given by [38]

$$PL(d) \text{ [dB]} = \alpha + \beta 10 \log_{10}(d) + \xi, \quad (5.1)$$

where d is the distance in meters, with α and β are set to 61.4 and 2, respectively. Log-normal shadowing ξ follows the normal distribution of mean 0 and variance 5.8. The link capacity was computed using the Shannon capacity formula in (2.4) and (2.5), assuming that each node is equipped with eight RF chains for beamforming and has a directivity gain of 30 dB. We consider the blockage, where the probability of blockage ranges from 0.2 to 0.4 depending on the distance between a node and a UE since mmWave signals are highly sensitive to blockages, according to [43]

$$P_{block}(d) = 1 - e^{-2\lambda r d},$$

where d is the distance between the node and the UE. Table 5.1 shows the simulation parameters.

We explored the performance of the proposed MWS algorithm in three different environments. The first environment involved 30 static UEs deployed uniformly according to the 3GPP IAB report. The second environment involves load variations in the network. Finally, we considered an environment where 30 UEs moved around in the network using

the freeway mobility model (FMM) [44]. These environments allowed us to evaluate the performance of our algorithm under different network conditions and mobility patterns of UEs. We compared the MWS algorithm with the MaxMinThroughput algorithm in the self-backhaul mmWave, maximizing the minimum throughput for UEs [31]. The other comparison is the Backpressure algorithm, which maximizes the backlog queues [36]. For comparison metrics, we used the standard deviation, packets dropped, average throughput, and quality of experience of UEs.

5.2 Impact of static UE on network performance

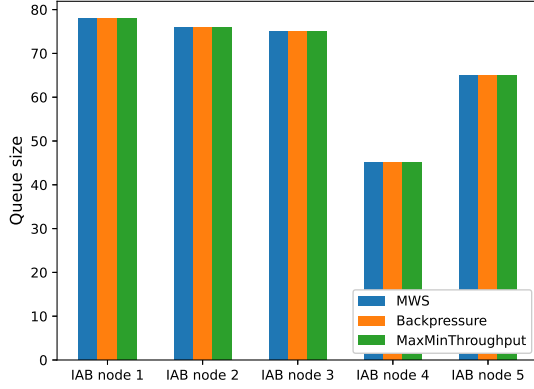
To identify the impact of the proposed algorithm and previous algorithms in the static environment, we simulated the network operation in 10,000 time slots with UEs requiring a data rate of 2 Mbps. We investigated the performance in two aspects: load balancing and QoE for UEs. For load balancing, we used queue status in a specific time slot and standard deviation to express load distribution in the network. The standard deviation σ_q is defined as

$$\sigma_q = \sqrt{\frac{\sum_{i \in \mathcal{N}} |q_i - \bar{q}|^2}{N}}.$$

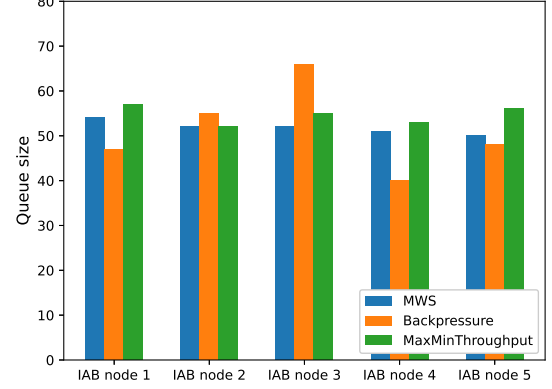
where q_i is the load of node i , \bar{q} is average load in the network and N is total number of nodes. The lower σ_q value shows that all nodes are evenly distributed, while the higher σ_q indicates all nodes are imbalanced. For assessing the QoE of the UEs, we considered the number of packets dropped over 3,000 packets being sent per time slot, which is derived

$$P_{drop} \triangleq \sum_{i \in \mathcal{N}} \phi_i(t),$$

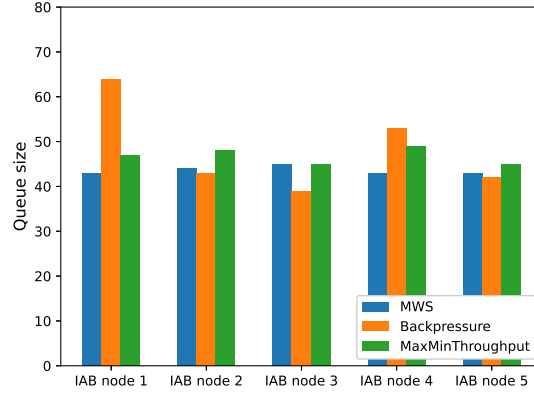
where $\phi_i(t)$ is the number of packets dropped at node i in time slot t . Other metrics to assess were the average throughput during the simulation and the percentage of satisfied



(a) $t = 1$



(b) $t = 7000$



(c) $t = 10000$

Figure 5.1: Queue status in the network

UEs, which is the number of UEs getting required data from the serving node over the total number of UEs. The average throughput is computed as

$$R \triangleq \frac{\sum_{i \in \mathcal{U}} \sum_{\tau \in [t, t+T)} \gamma_i(\tau)}{T},$$

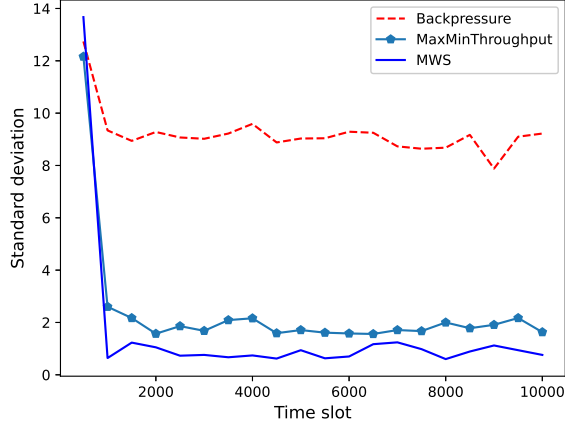


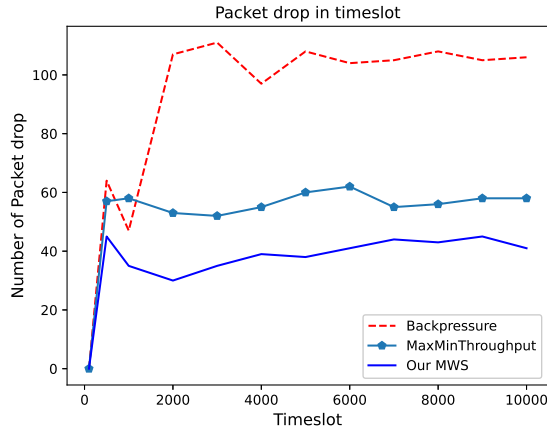
Figure 5.2: Standard deviation

where $\gamma_i(\tau)$ represents the number of bits received at time τ by the i -th UE. The percentage of satisfied UEs at node i is calculated as

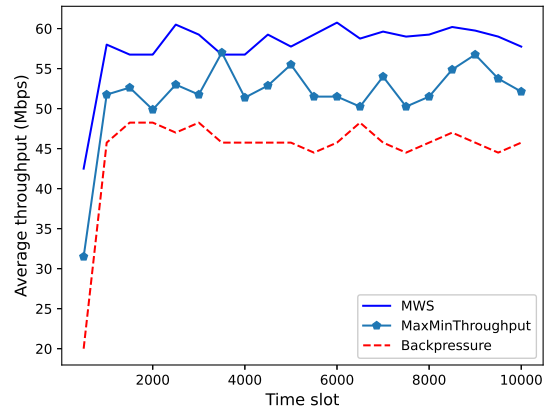
$$\theta_{pct} \triangleq \frac{\theta_i}{|U_i|} \cdot 100,$$

where θ_i and $|U_i|$ are the number of satisfied UEs and the total number of UEs at node i , respectively.

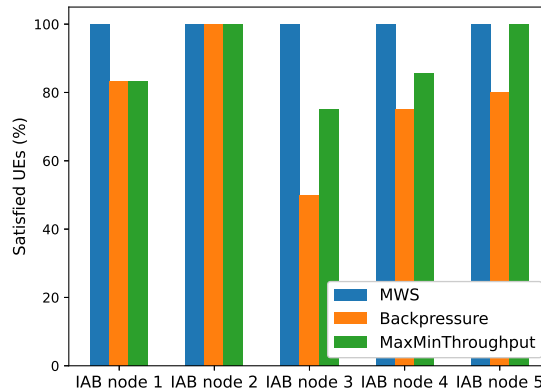
Fig. 5.1 indicates that the load among the nodes becomes more evenly distributed over time. Starting with the same initial load in all three algorithms, as shown in Fig. 5.1a, the overloaded nodes released some of their load to the underloaded neighboring nodes in the backhaul link and to UEs in the access link under our MWS algorithm. As a consequence, the load across the nodes in the network became more balanced. The figure shows that the gap between the maximum node's queue length and the minimum node's queue length, from the starting time to the ending time, was reduced from 30 to 5, respectively. This reduction leads to lower variance in the loads among nodes in the network. As a result, Fig. 5.2 shows that the proposed algorithm achieved the lowest load variance compared



(a) Packets dropped per time slot



(b) Average throughput by simulation time



(c) Percentage of UEs satisfied

Figure 5.3: Performance under 3GPP constraints

to other algorithms, resulting in a more balanced load distribution in each time slot.

While balancing loads, the QoE of UEs may be affected. In the Backpressure and MaxMinThroughput algorithms, overloaded nodes were unable to allocate the required resources to neighboring nodes and to UEs due to a shortage of resources, while underloaded nodes remained underutilized. On the other hand, the MWS algorithm distributed

the load from overloaded nodes to underloaded nodes and to UEs. This allowed the nodes to fully allocate the required resources on both access and backhaul links, resulting in a low number of dropped packets, as shown in Fig. 5.3a. Furthermore, the proposed algorithm increased network throughput and achieved a higher percentage of satisfied UEs, as depicted in Figs. 5.3b and 5.3c, respectively. As a result, not only did the network throughput increase, but UE satisfaction also improved compared to other algorithms.

5.3 Impact of Network Load Variations

In this subsection, we studied the load variations of the network under two conditions: various numbers of UEs and different data rate requirements of UEs. To investigate how the network load affects the IAB network, the number of UEs varied from 10 to 60 while keeping the data rate of each UE constant at 2 Mbps. For different data rate requirements of UEs, 20 UEs were uniformly deployed in the network and varied their rates from 0.5 to 5 Mbps. We simulated 10,000 time slots under both conditions. For each condition, 10 simulations were considered with a 95% confidence interval.

Fig. 5.4 illustrates the performance of the three algorithms as the number of UEs varied. Fig. 5.4a shows that the proposed algorithm had a better standard deviation balance than the other algorithms under all network load conditions. Figs. 5.4b and 5.4c show the average throughput and the percentage of satisfied UEs, respectively, as the number of UEs increased. In scenarios with a small number of UEs, the MWS algorithm effectively balanced the load with little impact on throughput. However, as traffic increased, our proposed algorithm attained higher throughput and a higher percentage of satisfied UEs than the other algorithms. Moreover, the proposed system achieved a 11% higher satis-

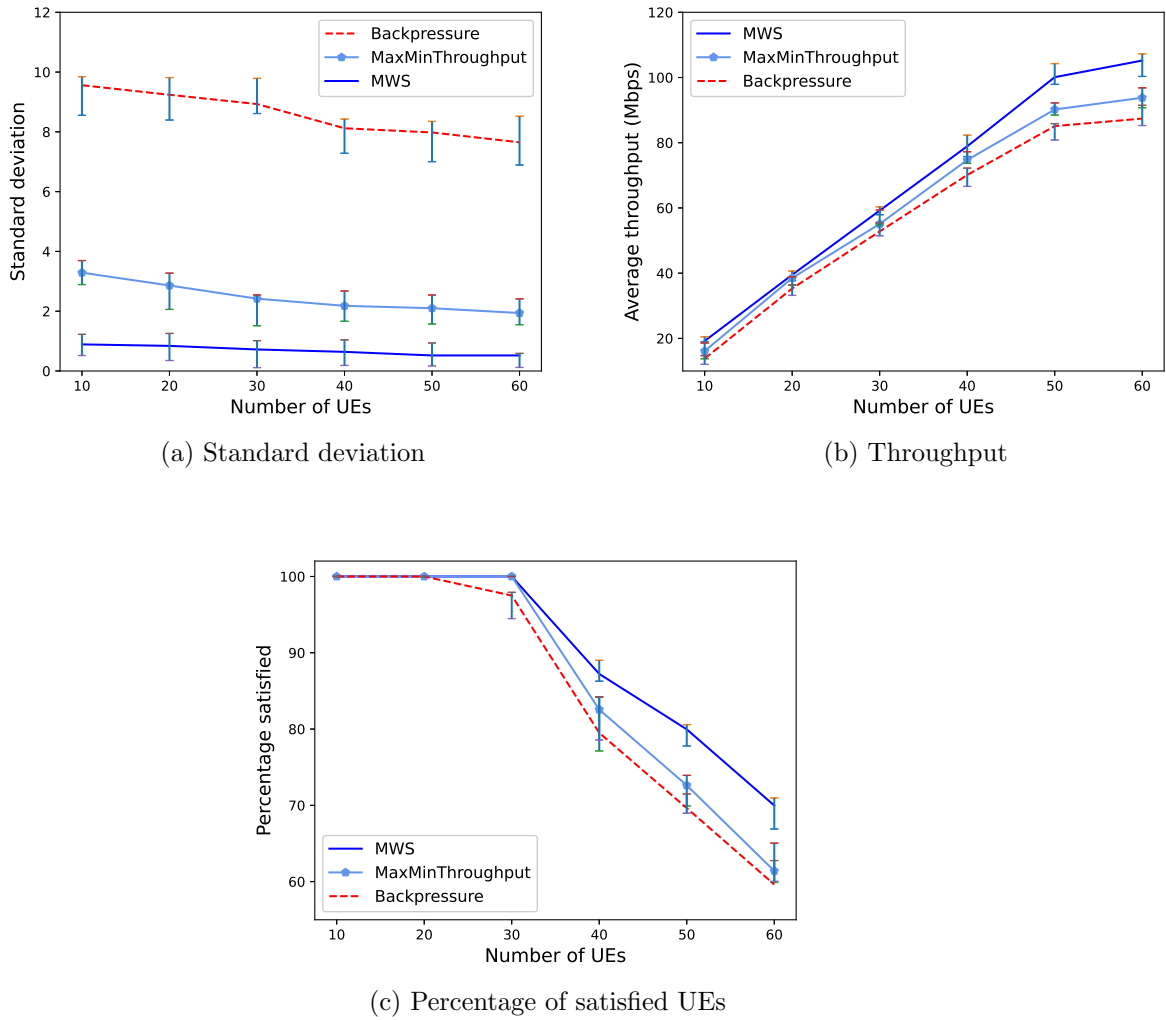


Figure 5.4: Performance based on the number of UEs

fraction rate in QoE user satisfaction and a smaller narrow confidence interval in standard deviation and average throughput than previous algorithms.

We also investigated how the performance of the three algorithms was affected by different data rate requirements for the UEs. As shown in Fig.5.5a, the proposed algorithm effectively balanced the load compared to the other algorithms. Fig. 5.5b depicts the

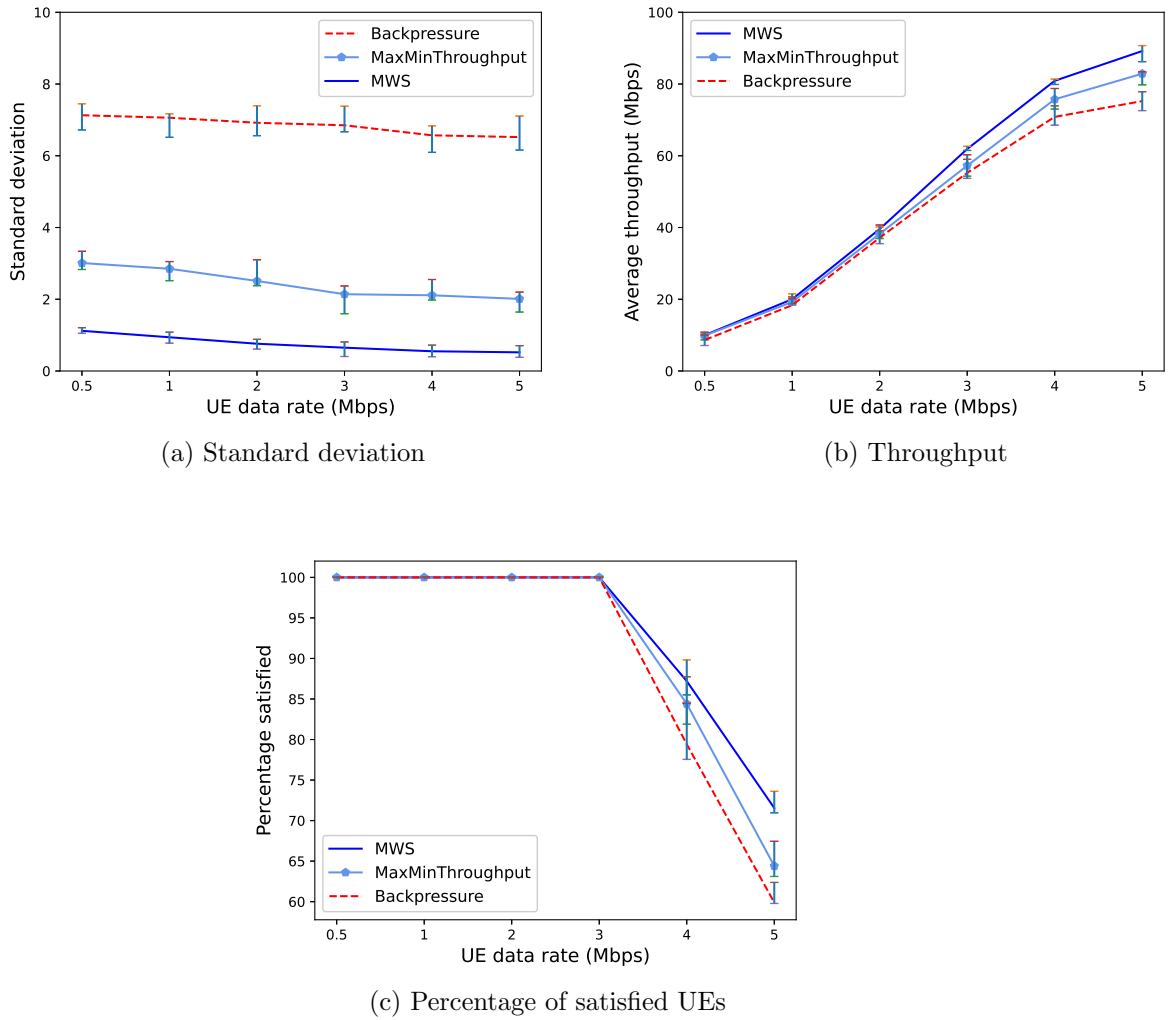


Figure 5.5: Performance based on UE data rate requirements

average throughput in the network. In low- and moderate-traffic load situations where most small cells remain underloaded, all algorithms balanced the load well. Even when the traffic load is high, the proposed algorithm performed better than the others with low variation in confidence intervals and achieved a satisfaction rate more than 14% higher in terms of the percentage of satisfied UEs, as shown in Fig. 5.5c.

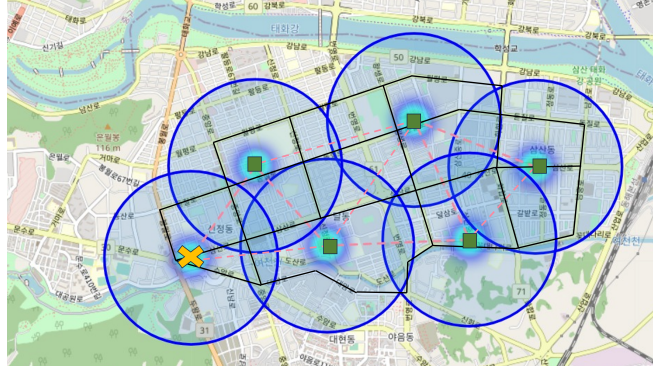
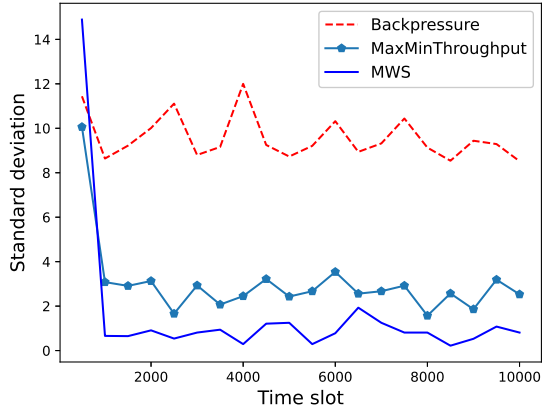


Figure 5.6: Random freeway mobility model

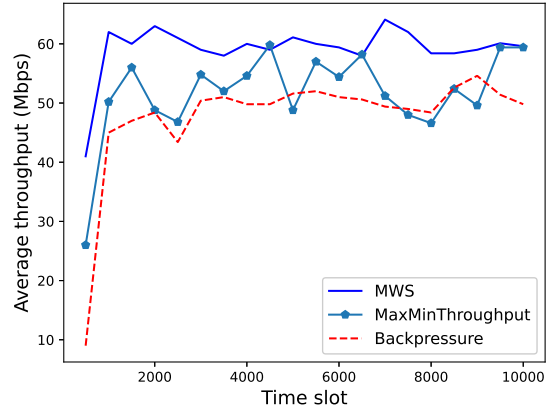
5.4 Impact of User Mobility Model

In practice, mobile users may have different mobility directions and speeds, which can affect the performance of the MWS algorithm. Therefore, to examine the randomness of mobility in a real environment, we consider the freeway mobility model (FMM) to simulate UE mobility [44]. The FMM simulates the movement of users on roads. Fig. 5.6 illustrates a uniform deployment of nodes with circular coverage around each node on the map. The IAB donor is depicted as an X; squares indicate the IAB node. The dashed lines represent links between each node, whereas solid lines represent random mobility in the UEs. 30 UEs are uniformly deployed and randomly move on the road with a uniform speed ranging from 1 m/s to 10 m/s and select random directions at the intersection.

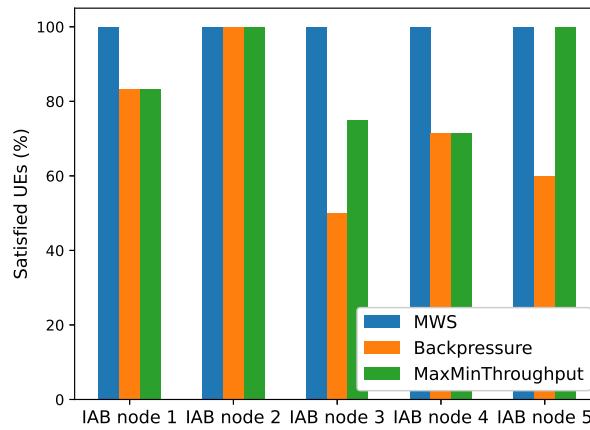
Fig. 5.7 illustrates the performance of the proposed algorithm in relation to user mobility. Due to higher user mobility, frequent changes in UE topology and handovers lead to greater load fluctuations among the nodes, as shown in Fig. 5.7a. However, our MWS algorithm has fewer fluctuations in standard deviation than the others. Furthermore, Fig. 5.7b shows that our proposed algorithm achieved higher average throughput in the network. Similarly, our algorithm ensures higher QoE for UEs compared to the other algorithms,



(a) Standard deviation



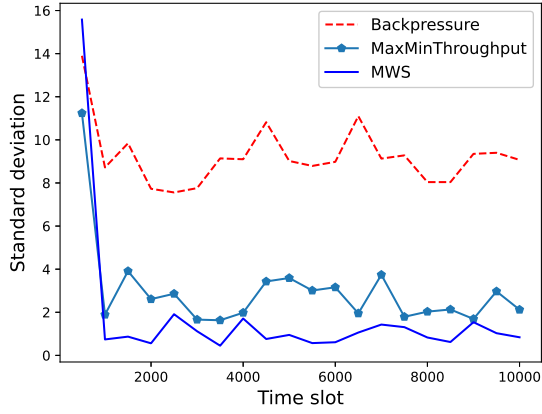
(b) Throughput



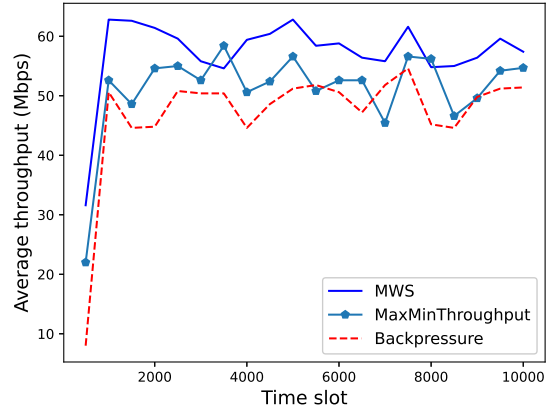
(c) Percentage of satisfied UEs

Figure 5.7: Performance in the UEs mobility

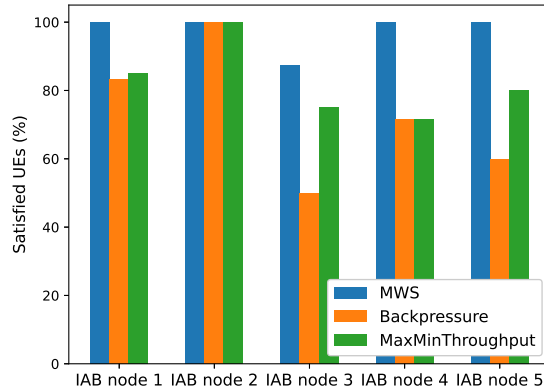
as depicted in Fig. 5.7c. The proposed algorithm achieves 100% user satisfaction, while the MaxMinThroughput and Backpressure algorithms only satisfied QoS requirements of 85.95%, 72.95% of the UEs in the network, respectively.



(a) Standard deviation



(b) Throughput



(c) Percentage of satisfied UEs

Figure 5.8: Performance based under heterogeneous network

5.5 Impact of Heterogeneous Network

To verify the correctness and assess the performance of the proposed algorithm in a heterogeneous network, we simulated environment in Section V-D and separated the transmission power in access and backhaul of each node. The transmission range was a uniform

range between 23 and 26 dBm for serving UEs in access link, while the transmission range in backhaul link ranged between 31 and 35 dBm. Following [45], we consider the Rician fading model in LOS links for access links by modifying (5.1).

$$PL(d) \text{ [dB]} = \alpha + \beta 10 \log_{10}(d) + \xi + \epsilon,$$

where ϵ is Rician fading set as 13 dB.

Fig. 5.8 illustrates the performance of the proposed algorithm in the heterogeneous network. Due to irregular node deployment, higher user mobility and fading, frequent changes in UE topology within the environment result in more load fluctuations among the nodes and variations in the received signal strength of UEs. However, our MWS algorithm outperformed other algorithms in standard deviation and average network throughput, as depicted in Figs. 5.8a and 5.8b, respectively. Similarly, Fig. 5.8c shows that our algorithm ensured higher QoE for UEs, where the proposed algorithm achieved a greater increase of 18.9% and 33.6% than MaxMinThroughput and Backpressure algorithms in user satisfaction, respectively. Figs. 5.8a, 5.8b, and 5.8c clarify that the proposed algorithm is more robust than the others in the heterogeneous network.

Chapter 6

Conclusion

IAB technology offers a flexible solution for reducing deployment costs by facilitating wireless backhaul and streamlining architecture. It is characterized by the extensive reuse of functions and interfaces initially defined for the access link and the backhaul link. Additionally, IAB relaying holds the promise of delivering both increased throughput, QoE, and extended coverage, especially in 5G, 5G advanced, and 6G. While recent literature has explored resource allocation, the unbalanced load in the IAB network remains uncertain.

In this thesis, we proposed a graph-based algorithm for load balancing in small-cell IAB networks. Mobility of UEs and randomness of the network topology led to imbalanced load distributions across the cells, resulting in degraded network performance in throughput and QoS for UEs. We considered the sum of queues on both access and backhaul links as a metric to measure the load and emphasized load balancing in link scheduling by minimizing the load variance. To address the load-balancing problem, we modeled the network as a directional bipartite graph. The algorithm calculated weight as a cost metric in the network graph based on the current load and the link capacity of the BSs.

By finding the semi-matching of the maximum weight, the algorithm constructed flexible scheduling for resource allocation in overloaded nodes to release packets to neighboring nodes and UEs under their coverage. Our work demonstrated that the proposed algorithm effectively balanced loads, reduced dropped packets, achieved high average throughput, and ensured satisfactory QoE for UEs in three environments (static, load variations, mobility, and heterogeneous scenarios) better than other algorithms. Furthermore, the proposed algorithm allocated sufficient resources to all UEs, ensuring that 100% the UEs achieved their required data rate, while MaxMinThroughput and Backpressure only satisfied 85.95%, 72.95% of the UEs.

References

- [1] Mamta Agiwal, Abhishek Roy, and Navrati Saxena. Next generation 5g wireless networks: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 18(3):1617–1655, 2016.
- [2] Muhammad Amjad, Leila Musavian, and Mubashir Husain Rehmani. Effective capacity in wireless networks: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 21(4):3007–3038, 2019.
- [3] Ericsson. Mobile data traffic outlook, Nov 2022.
- [4] Cisco. Global - 2021 forecast highlights - cisco, 2021.
- [5] Jeffrey G. Andrews, Stefano Buzzi, Wan Choi, Stephen V. Hanly, Angel Lozano, Anthony C. K. Soong, and Jianzhong Charlie Zhang. What will 5g be? *IEEE Journal on Selected Areas in Communications*, 32(6):1065–1082, 2014.
- [6] Theodore S. Rappaport, Shu Sun, Rimma Mayzus, Hang Zhao, Yaniv Azar, Kevin Wang, George N. Wong, Jocelyn K. Schulz, Mathew Samimi, and Felix Gutierrez. Millimeter wave mobile communications for 5g cellular: It will work! *IEEE Access*, 1:335–349, 2013.

- [7] Wei Jiang, Bin Han, Mohammad Asif Habibi, and Hans Dieter Schotten. The road towards 6g: A comprehensive survey. *IEEE Open Journal of the Communications Society*, 2:334–366, 2021.
- [8] Marco Giordani and Michele Zorzi. Non-terrestrial networks in the 6g era: Challenges and opportunities. *IEEE Network*, 35(2):244–251, 2021.
- [9] Walid Saad, Mehdi Bennis, and Mingzhe Chen. A vision of 6g wireless systems: Applications, trends, technologies, and open research problems. *IEEE Network*, 34(3):134–142, 2020.
- [10] Sundeep Rangan, Theodore S. Rappaport, and Elza Erkip. Millimeter-wave cellular wireless networks: Potentials and challenges. *Proceedings of the IEEE*, 102(3):366–385, 2014.
- [11] Amitava Ghosh, Timothy A. Thomas, Mark C. Cudak, Rapeepat Ratasuk, Prakash Moorut, Frederick W. Vook, Theodore S. Rappaport, George R. MacCartney, Shu Sun, and Shuai Nie. Millimeter-wave enhanced local area systems: A high-data-rate approach for future wireless networks. *IEEE Journal on Selected Areas in Communications*, 32(6):1152–1163, 2014.
- [12] Hayssam Dahrouj, Ahmed Douik, Frank Rayal, Tareq Y. Al-Naffouri, and Mohamed-Slim Alouini. Cost-effective hybrid rf/fso backhaul solution for next generation wireless systems. *IEEE Wireless Communications*, 22(5):98–104, 2015.
- [13] Xavier Artiga, Jose Nunez-Martinez, Ana Perez-Neira, Gorca Juan Lendrino Vela, Juan Mario Fare Garcia, and Georgios Ziaragkas. Terrestrial-satellite integration in dynamic 5g backhaul networks. In *2016 8th Advanced Satellite Multimedia Sys-*

tems Conference and the 14th Signal Processing for Space Communications Workshop (ASMS/SPSC), pages 1–6, 2016.

- [14] Sérgio Julião, Francisco Martins, André Prata, Sergio Lopes, Joao Duarte, Paulo Jesusy, Nelson Silvey, Arnaldo S. R. Oliveira, and Nuno B. Carvalho. High performance microwave point-to-point link for 5g backhaul with flexible spectrum aggregation. In *2015 IEEE MTT-S International Microwave Symposium*, pages 1–4, 2015.
- [15] Behrooz Makki, Tommy Svensson, Thomas Eriksson, and Mohamed-Slim Alouini. On the performance of rf-fso links with and without hybrid arq. *IEEE Transactions on Wireless Communications*, 15(7):4928–4943, 2016.
- [16] Orawan Tipmongkolsilp, Said Zaghoul, and Admela Jukan. The evolution of cellular backhaul technologies: Current issues and future trends. *IEEE Communications Surveys & Tutorials*, 13(1):97–113, 2011.
- [17] Berke Tezergil and Ertan Onur. Wireless backhaul in 5g and beyond: Issues, challenges and opportunities. *IEEE Communications Surveys & Tutorials*, 24(4):2579–2632, 2022.
- [18] 3GPP. NR: Integrated access and backhaul (iab) radio transmission and reception, technical specification (TS) 38.174. Technical report, 3GPP, Jun. 2020. v.0.1.0.
- [19] Xingqin Lin. An overview of 5g advanced evolution in 3gpp release 18. *IEEE Communications Standards Magazine*, 6(3):77–83, 2022.
- [20] Cedric Dehos, Jose Luis González, Antonio De Domenico, Dimitri Ktésas, and Laurent Dussopt. Millimeter-wave access and backhauling: the solution to the exponential

- data traffic increase in 5g mobile communications systems? *IEEE Communications Magazine*, 52(9):88–95, 2014.
- [21] Yilin Li, Emmanouil Pateromichelakis, Nikola Vucic, Jian Luo, Wen Xu, and Giuseppe Caire. Radio resource management considerations for 5g millimeter wave backhaul and access networks. *IEEE Communications Magazine*, 55(6):86–92, 2017.
- [22] Erik Ekudden. Integrated Access and Backhaul – a new type of wireless backhaul in 5G. Technical report, Ericsson, Stockholm, Sweden, Jun 2020.
- [23] Mark Cudak Amitabha Ghosh. Integrated Access and Backhaul: Why it is essential for mmwave deployments. Technical report, Nokia, Nov 2020.
- [24] Michele Polese, Marco Giordani, Tommaso Zugno, Arnab Roy, Sanjay Goyal, Douglas Castor, and Michele Zorzi. Integrated access and backhaul in 5g mmwave networks: Potential and challenges. *IEEE Communications Magazine*, 58(3):62–68, 2020.
- [25] Yekaterina Sadovaya, Dmitri Moltchanov, Wei Mao, Oner Orhan, Shu-ping Yeh, Hosein Nikopour, Shilpa Talwar, and Sergey Andreev. Integrated access and backhaul in millimeter-wave cellular: Benefits and challenges. *IEEE Communications Magazine*, 60(9):81–86, 2022.
- [26] Mehrdad Shariat, Emmanouil Pateromichelakis, Atta ul Quddus, and Rahim Tafazolli. Joint tdd backhaul and access optimization in dense small-cell networks. *IEEE Transactions on Vehicular Technology*, 64(11):5288–5299, 2015.
- [27] Muhammad Nazmul Islam, Sundar Subramanian, and Ashwin Sampath. Integrated access backhaul in millimeter wave networks. In *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6, 2017.

- [28] Muhammad Nazmul Islam, Navid Abedini, Georg Hampel, Sundar Subramanian, and Junyi Li. Investigation of performance in integrated access and backhaul networks. In *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 597–602, 2018.
- [29] Chiranjib Saha, Mehrnaz Afshang, and Harpreet S. Dhillon. Bandwidth partitioning and downlink analysis in millimeter wave integrated access and backhaul for 5G. *IEEE Transactions on Wireless Communications*, 17(12):8195–8210, 2018.
- [30] Chiranjib Saha and Harpreet S. Dhillon. On load balancing in millimeter wave het-nets with integrated access and backhaul. In *2019 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6, 2019.
- [31] Chao Fang, Charitha Madapatha, Behrooz Makki, and Tommy Svensson. Joint scheduling and throughput maximization in self-backhauled millimeter wave cellular networks. In *2021 17th International Symposium on Wireless Communication Systems (ISWCS)*, pages 1–6, 2021.
- [32] Dingwen Yuan, Hsuan-Yin Lin, Jörg Widmer, and Matthias Hollick. Optimal joint routing and scheduling in millimeter-wave cellular networks. In *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, pages 1205–1213, 2018.
- [33] Trung Kien Vu, Chen-Feng Liu, Mehdi Bennis, Mérouane Debbah, and Matti Latva-aho. Path selection and rate allocation in self-backhauled mmwave networks. In *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6, 2018.

- [34] Yilin Li, Jian Luo, Richard A. Stirling-Gallacher, and Giuseppe Caire. Integrated access and backhaul optimization for millimeter wave heterogeneous networks, 2019.
- [35] Matteo Pagin, Tommaso Zugno, Michele Polese, and Michele Zorzi. Resource management for 5G NR Integrated Access and Backhaul: A semi-centralized approach. *IEEE Transactions on Wireless Communications*, 21(2):753–767, 2022.
- [36] Swaroop Gopalam, Stephen V. Hanly, and Philip Whiting. Distributed and local scheduling algorithms for mmwave integrated access and backhaul. *IEEE/ACM Transactions on Networking*, 30(4):1749–1764, 2022.
- [37] 3GPP. 5G - NR: Study on Integrated Access and Backhaul (IAB), Technical Specification (TS) 38.874. Technical report, 3GPP, Jan. 2019. v.16.0.0.
- [38] Mustafa Riza Akdeniz, Yuanpeng Liu, Mathew K. Samimi, Shu Sun, Sundeep Rangan, Theodore S. Rappaport, and Elza Erkip. Millimeter wave channel modeling and cellular capacity evaluation. *IEEE Journal on Selected Areas in Communications*, 32(6):1164–1179, 2014.
- [39] Shu Sun, Theodore S. Rappaport, Robert W. Heath, Andrew Nix, and Sundeep Rangan. MIMO for millimeter-wave wireless communications: beamforming, spatial multiplexing, or both? *IEEE Communications Magazine*, 52(12):110–121, 2014.
- [40] 3GPP. 5G - NR: Study on Integrated Access and Backhaul (IAB): radio transmission and reception, Technical Specification (TS) 38.174. Technical report, 3GPP, April. 2022. v.17.0.0.
- [41] 3GPP. 5G: Study on scenarios and requirements for next generation, Technical Report (TR) 138.913. Technical report, 3GPP, April. 2022. v.17.0.0.

- [42] Nicholas J.A. Harvey, Richard E. Ladner, László Lovász, and Tami Tamir. Semi-matchings for bipartite graphs and load balancing. *Journal of Algorithms*, 59(1):53–78, 2006.
- [43] Tianyang Bai and Robert W. Heath. Coverage and rate analysis for millimeter-wave cellular networks. *IEEE Transactions on Wireless Communications*, 14(2):1100–1114, 2015.
- [44] F. Bai, Narayanan Sadagopan, and A. Helmy. Important: a framework to systematically analyze the impact of mobility on performance of routing protocols for adhoc networks. In *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No.03CH37428)*, volume 2, pages 825–835 vol.2, 2003.
- [45] Mathew K. Samimi, George R. MacCartney, Shu Sun, and Theodore S. Rappaport. 28 ghz millimeter-wave ultrawideband small-scale fading models in wireless channels. In *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, pages 1–6, 2016.