UNIVERSITY OF CINCINNATI

Date: 4-Nov-2010

I, Weihuang Fu, hereby submit this original work as part of the requirements for the degree of:
Doctor of Philosophy

in Computer Science & Engineering

It is entitled:

Analytical Model for Capacity and Delay Optimization in Wireless Mesh Networks

Student Signature: Weihuang Fu

This work and its defense approved by:

Committee Chair: Dharma Agrawal, DSc
Dharma Agrawal, DSc

Chia Han, PhD
Chia Han, PhD

Prabir Bhattacharya, PhD
Prabir Bhattacharya, PhD

Kenneth Berman, PhD
Kenneth Berman, PhD

Wen Ben Jone, PhD
Wen Ben Jone, PhD
Analytical Model for Capacity and Delay Optimization in Wireless Mesh Networks

by

Weihuang Fu

B.S. (Zhejiang University of Technology, China) 2002
M.S. (Zhejiang University of Technology, China) 2005

A dissertation submitted in partial satisfaction of the requirements for the degree of
Doctor of Philosophy
in
Computer Science and Engineering
in the
School of Computing Sciences and Informatics
of the
College of Engineering and Applied Science
of the
UNIVERSITY OF CINCINNATI, OHIO

Committee:
Professor Dharma P. Agrawal, Chair
Professor Kenneth A. Berman
Professor Prabir Bhattacharya
Professor Chia-Yung Han
Professor Wen-Ben Jone

2010
Abstract

Motivated by ubiquitous communication, both wireless network theory and technology have vigorously developed in the past decades that could support broadband wireless access (BWA), and the current trend continues to replace wired network backbone. Conventional network access is served by network infrastructure, which is deployed at fixed locations and acts as “bridge”, i.e., gateway, between wired backbone and mobile clients (MCs) with equipped wired-interface and air-interface. Infrastructures have to be placed at the locations where cables available, including network and power cables, which poses strong constraints on deployment locations, and high cost in cable deployment and maintenance.

Wireless mesh networks (WMNs) are comprised of multi-radio mesh routers (MRs), which interconnect each other using wireless links to form a mesh backbone. This also forms a multi-cell architecture to provide network service for MCs, where Internet gateways (IGWs) are special MRs having wired connection to the Internet. The deployment of MRs is flexible, cost-efficient, self-organizing, etc. Mobile MRs even form a mobile mesh backbone. Due to its advantages, WMN could be one of the promising case of the next generation Internet. However, developing such a network also needs to address many fundamental issues inherited from two-tier network architecture, wireless multi-hop transmission, multi-cell structure, etc.

In this dissertation, we analytically model a two-tier WMN and derive the asymptotic bounds of network capacity and delay, which are essential and tightly related factors in developing a WMN to support delay-sensitive applications such as voice over IP (VoIP), video conference, etc. This dissertation performs the analysis on a WMN backbone formed by self-organizing ad hoc MRs and shows how the net-
work capacity is dominated by the network delay constraints, and the numbers of MCs, MRs and IGWs. We find that the network delay scales to either the number of MRs or the number of IGWs, and dominating factors depends on the type of routing strategy. Some of our results are also applicable to ad hoc networks, which can be equivalent to special case of a self-organizing WMN.

We optimize the backbone capacity by introducing two types of channel assignment schemes in managing spectrum resource and mitigating backbone interference. One is a centralized channel assignment scheme, which is suitable to WMNs deployed by Internet service providers (ISPs), and the other is a distributed channel assignment scheme, which applies to static or mobile self-organizing WMN. In addition, we propose a clustering based fractional frequency reuse for multi-cell coverage of WMNs, which offers resource allocation higher flexibility and better fairness with additional spatial dimension. Our work is to analyze and solve the fundamental problems in developing WMNs for ubiquitous and pervasive access. The results in our dissertation can serve as the guideline in research and design of practical WMNs. We conclude with dissertation with some discussion on future area of research.
Keywords— Channel Assignment, Combinatorics, IEEE 802.16 WirelessMAN, Information Theory, Interference Mitigation, Mobile Ad hoc Networks (MANETs), Multi-cell, Multi-channel and Multi-radio, Multi-hop, Queueing Theory, Resource Management, Stochastic Processes, Wireless Mesh Networks (WMNs).
To my parents and my wife,

for their love, endless support and encouragement.
Acknowledgments

I would like to express my sincere gratitude to my thesis advisor Dr. Dharma P. Agrawal for his valuable suggestions and insightful guidance on my research work. He provided me with the enough freedom to explore the field of my interest and also gave me valuable suggestions and helps whenever I need to figure out the issues. I also thank my committee members Dr. Wen-Ben Jone, Dr. Prabir Bhattacharya, Dr. Kenneth A. Berman, and Dr. Chia-Yung Han for taking their valuable time for being on my thesis committee.

My heartfelt gratitude is extended to my parents and my wife. They played a major role in shaping my career and always ensured that I channelized my effort in the right direction. I am thankful to Haitang Wang, Yi Cheng, Demin Wang, Junfang Wang, Bing He, Jung Hyun Jun, Asitha Bandaranayake, Haiflong Li, Yang Chi, Talmia Oliveira, Amitabh Mishra, Vaibhav Pandit, Ahmad Mostafa, Kuheli Louha, Dushan Don, Abhinav Prakash, Nishan Weragama, and many other fellow graduate students and research colleagues for their support and help. Last but not least, I would like to thank University of Cincinnati for granting me University Graduate Scholarship and providing me a platform to excel professionally.
# Contents

List of Figures ........................................... x

1 Introduction ............................................. 1
  1.1 Two-tier Wireless Mesh Networks ....................... 1
  1.2 Motivations ........................................ 5
  1.3 Organization ....................................... 8

2 Delay and Capacity in WMNs ......................... 10
  2.1 Introduction ....................................... 10
  2.2 Problem Formulation ................................ 14
  2.3 Analytical Modeling on Network Delay and Capacity ... 17
    2.3.1 Network Connectivity .......................... 17
    2.3.2 Number of Routing Hops ....................... 19
    2.3.3 Traffic on Cells ............................... 23
    2.3.4 Cell Queue Model .............................. 27
    2.3.5 Network Delay and Capacity ................... 30
  2.4 Discussion and Implications ........................ 36
  2.5 Conclusion ......................................... 39

3 Centralized Channel Assignment in WMN Backbone .... 40
  3.1 Introduction ....................................... 40
  3.2 Related Work ....................................... 42
  3.3 Network Model ..................................... 44
    3.3.1 Network Architecture .......................... 44
    3.3.2 Interference Model ............................. 45
    3.3.3 Traffic Model .................................. 46
# List of Figures

1.1 Coverage of APs ........................................... 2  
1.2 Cellular Network .......................................... 3  
1.3 Wireless Mesh Network .................................... 4  

2.1 Grid Cells and Connectivity ............................... 18  
2.2 Clustered Routing ......................................... 22  
2.3 Queue Reduction Model for Cell ......................... 27  
2.4 Cell Connectivity .......................................... 31  

3.1 Flow-based Conflict Graph ............................... 54  
3.2 Channel Aggregation ....................................... 57  
3.3 Number of Required Channels ............................ 63  
3.4 Average Degree ........................................... 64  
3.5 Channel Aggregation in Grid Topology .................. 64  
3.6 Channel Aggregation in Hexagon Topology ............... 65  
3.7 Channel Aggregation in Random Topology ............... 65  

4.1 Distributed Channel Assignment and Queue Management in WMNs 71  
4.2 Radio Management ......................................... 75  
4.3 Queue Management ......................................... 78  
4.4 Probability of Connectivity ............................... 79  
4.5 Switching Rate ........................................... 80  
4.6 Delay at MR ................................................ 80  

5.1 Multi-cell Coverage ....................................... 89  
5.2 FFR Schemes .............................................. 93  
5.3 Example of Resource Allocation .......................... 100  

x
Chapter 1

Introduction

1.1 Two-tier Wireless Mesh Networks

Demand for ubiquitous and pervasive communication has led to vigorous development of wireless network associated technology and theory. In the past decade, conventional network access has been adequately supported by the underlying network infrastructure. As an extension at the edge of the Internet, access points (APs) of wireless local area networks (WLANs) [1] are deployed at home or office to serve a group of terminals, such as laptops, PDA, etc., within its coverage distance that varies from tens to hundreds meters. The convenience of network access using wireless connection has resulted in booming deployment of APs at hotspots such as airport, coffee shop, etc. To provide compatible access protocols, IEEE 802.11 family of standards [2] have been defined for physical (PHY) and media access (MAC) layers that support carrier sense multiple access (CSMA) based access and providing tens to hundreds Mbps data rate per AP.

However, deployment of APs is still far away from supporting ubiquitous communication. APs have to be placed only at the places where accessing Internet cable
could be easily feasible. So, most of the APs are available only indoors. If we look at deployment of APs in a geographic area, APs are shown to be randomly scattered in a region $\mathcal{D}$, where the network access for mesh clients (MCs) is only available within the circles around the corresponding APs, as shown in Figure 1.1. Any users outside of the circular area still cannot obtain network services. In addition, mobile users, i.e., MCs, are not supported by APs because there is no handover mechanism designed at the MAC layer to migrate the connection of MC from one AP to another. The connection to AP is easy to break after MC moves out of the coverage.

With an objective to provide seamless and pervasive communication for a group of MCs, multi-cell cellular system [3] have been developed for worldwide base stations (BSs) deployment. The coverage of BSs overlaps, so MCs can access to any nearby BS for the network connection, as illustrated in Figure 1.2. BSs are connected with wired cables or fiber optical links to the core network, where several servers are located for the control purpose. Cellular system supports handover
of MCs from one cell to another by exchanging connection information and re-routing data packets through the core network. A successful cellular system such as global system for mobile communications (GSM) is being widely used for voice service. 3G cellular network is deployed to replace GSM with wider bandwidth and additional support of data service. 4G cellular network (e.g., IEEE802.16m [4], IEEE802.16j [5], LTE advanced [6]) is at the research stage with an objective of supporting better data rate (hundreds Mbps) and higher mobility (above 160 km/hr). However, to deploy a multi-cell network, a lot has to be invested on wired cables and BS deployment. As there are many constraints on the selection of BS’s locations, placement of BSs are still limited to the areas with higher population. Since infrastructure has to be interconnected by wired cables or optical fibers, the multi-cell network mentioned above is restricted at the surface of the terrain and due to cost constraints, only a fraction of terrain has been covered.

To provide outdoor connectivity with high bandwidth, wireless mesh network (WMN) [7] has been introduced which comprises of multi-radio mesh routers (MRs),
which interconnect to each other with wireless links to form a mesh backbone. Such a multi-cell architecture provides high-speed network services for MCs, as illustrated in Figure 1.3. In WMN backbone, Internet gateways (IGWs) are special MRs that have wired connection to the Internet, which serves as the gateways/bridges between the WMN and the Internet. The deployment of MRs is flexible, cost-efficient, self-organizing, etc. Mobile MRs even further provide a moving mesh backbone. Therefore, potential applications of WMNs are much wider and more flexible, which are useful everywhere such as terrain, water surface, under water, in the sky, universal space, etc.

Thus, WMN provides a flexible low cost Internet access for MCs in a wide region, by having a backbone consisting of MRs and IGWs. In such a network, higher throughput and lower power consumption are achieved by having smaller MR cell coverage. On the other hand, MCs can be provided longer battery life by using shorter transmission distance and hence benefit from lower transmit power. MRs
interconnect to each other and route aggregated traffic in a multi-hop fashion, constitutes the backbone of a WMN. The MRs in a WMN can be deployed either by Internet service providers (ISPs) or users. ISP can deploy MRs with various optimization schemes in mind. User deployed MRs are in an ad hoc fashion and self-organizing, forming a self-organizing WMN [8]. MRs may be equipped with multiple radios to support multiple simultaneous transmissions and offset the interference. The analysis on capacity, delay, and their mutual effects are basic parameters for the development of a practical WMN.

A number of academic WMN testbeds, such as UC-WMN, Roofnet [9], CUWiN [10], etc, have been set up to demonstrate feasibility of WMNs, where UC-WMN is deployed by our CDMC laboratory, i.e., Center for Distributed and Mobile Computing at the University of Cincinnati. Some commercial WMNs, such as Meraki Networks in San Francisco [11], have been deployed which actually aim at gaining some market quota in the near future. However, all these WMNs encounter a challenging problem: the network performance is far below users’ expectation. The question that how to achieve desired network performance has become a crucial issue in WMNs.

1.2 Motivations

Due to associated advantages, WMN could constitute one of the alternatives for the next generation Internet. To deploy such a network, the capacity and the expected delay of packets in the network must be known, which is the fundamental problems for any types of networks. As described in previous sections, a WMN is a two-tier wireless network. The network elements include MRs, MCs, IGWs, and
the links among them. Investigating capacity and delay issue on a two-tier network architecture is very complicated. It involves the attributes of each network elements, the links between them, and the impact from two tiers. To investigate the capacity and delay in a two-tier WMN:

Firstly, the number of MRs, MCs, and IGWs will have great impact to the capacity and delay. Also, the geometric deployment of these network elements deterministically affect the performance. We apply asymptotic bounds to the problem to show the dominating effects of the number of MRs, MCs, and IGWs.

Secondly, the broadcasting nature of wireless signals lead to the share of neighboring wireless links while they operate on same frequency band. The schedule of the transmission for wireless links between MRs is critical for higher throughput and lower latency. Since the transmission in a WMN backbone is featured with multi-hop fashion, the scheduling will happen on links on both intra-routing path and inter-routing path, which lead a complicated optimization problem.

In addition, WMN forms a multi-cell coverage, which has to consider the inter-cell interfere (ICI) problem while providing network access service to MCs. ICI could greatly decrease the cell capacity and drop a large number of packets of the MC located at the edge region of cells while two neighboring cells operate on same frequency band at the same time. Considering the uneven MC locations, the WMN should be able to adjust the frequency bands allocation for higher capacity and lower latency.

This dissertation investigate the corresponding issues and the summarization are as follows:

- Analytical modeling on the network delay and capacity in WMNs. To discover
the dominating factors, we investigate the essential constraints on the capacity of a WMN, which are from the multi-hop transmission in backbone and the inter-cell interference among cells formed by MRs. Chapter 2 analyzes the network delay, capacity and their mutual dependence in a WMN, and show the dominating factors by asymptotic bounds.

- Analyzing channel assignment problem in multi-hop WMN backbone and proposing graphical-based channel assignment scheme in WMNs. Optimal channel assignment is a NP-hard problem [12]. Both radio switching delay and multi-channel hidden problem [13] have to be considered and solved using a robust approach. In Chapter 3, we propose an algorithm, computing the required number of channels for conflict free channel assignment. Given a network with the number of MRs and the traffic profile, the algorithm performs computation with linear programming objective of fair flows and flow-based conflict graph. We also propose a channel assignment approach for a fixed number of channels.

- Analyzing and proposing distributed channel assignment in Chapter 4. The distributed channel assignment can work with mobile MR to perform dynamically on-demand channel assignment. Our approach divides multiple radios into two sets: the receiving radios and the forwarding radios. The receiving radios are fixed at certain channels for a long duration and mainly used to receive packets. The forwarding radios are able to switch dynamically on different channels and used for forwarding the packets. The radios can be adaptively switched from one type to the other as per the traffic changes. Division of radios for different purpose can avoid the channel switching syn-
chronization problem in a dynamic channel assignment, while ensuring the network connectivity and adjusting the radios according to the traffic. In addition, switching of radios between the sets enables the control of MRs throughput, which facilitates traffic balancing in a given network. To further enhance the performance, an appropriate queue management scheme for multi-radio multi-channel WMNs is also proposed.

- Performing analytical work on multi-cell architecture and proposing differentiable FFR scheme. Since MRs form a multi-cell tier to provide service for MCs, the cell capacity should consider the mutually effect of neighboring cells. In chapter 5, we establish an analytical framework for FFR schemes, which also provides a foundation in deriving our scheme. Throughput of randomly deployed MSs in multi-cell networks is estimated with taking effects of frequency band partition and resource allocation of FFR into account. In result, FFR scheme can be evaluated by its cell throughput which is obtained from subchannel capacity. Then, we propose an innovative differentiable FFR (D-FFR) scheme capable of fully utilizing flexibility of OFDMA technology and enables each BS to have differentiable spectrum partition and resource allocations for cells and their zones. Our D-FFR provides a flexibility in ICI mitigation and elasticity in MAC layer quality of service (QoS) support.

1.3 Organization

The reminder of this dissertation is organized as follows: Chapter 2 presents the analytical work on the network delay and the capacity in a two-tier WMN, which
gives the asymptotic bounds. Chapter 3 presents a feasible centralized channel assignment scheme to reach the derived bound for the network backbone capacity, which also serves as the upper bound of distributed channel assignment. Chapter 4 presents a feasible distributed channel assignment scheme for self-organizing WMN backbone, which is also applicable to mobile WMN backbone. Chapter 5 presents a multi-cell resource allocation scheme based on fractional frequency reuse to mitigate ICI and improve cell capacity. Chapter 6 gives the conclusion and some possible future work.
Chapter 2

Delay and Capacity in WMNs

2.1 Introduction

In this chapter, we analyze the network delay, capacity, and their mutual effects on the performance of a WMN. WMN is defined as a multi-hop network that consists of uniformly independent identically distributed (i.i.d.) \( n \) MCs, uniformly i.i.d. \( N_R \) MRs, arbitrary deployed \( N_G \) IGWs. We notice that the routing strategy will have an important impact on the analysis. There are mainly two types of routing strategies in a WMN: ad hoc routing and clustered routing. Ad hoc routing is between a pair of source and destination MRs, using a shortest path. IGWs in a WMN can form multiple clusters where each IGW acts as a cluster head. Clustered routing is between an IGW and an MR that every MR forwards packets to the cluster head, and vice versa.

The definitions of network delay and capacity are as follows:

**Definition 2.1.** Network capacity is said to be feasible if a bit rate \( \Lambda \), can be provided to every MC in a network.
**Definition 2.2.** Network delay $T_D$, is defined as the average time interval from a packet arriving in a WMN backbone to its departure from the backbone, considering both queueing and transmission delays.

Previous work in the literature has focused on the asymptotic capacity, i.e., per-node maximum throughput, of a multi-hop network. The asymptotic capacity\(^1\) in ad hoc networks has been derived by the analytical work in [15], with assumptions that nodes are uniformly distributed and each node randomly chooses a destination. With ad hoc routing in a well-connected network that satisfies critical transmission distance [16], the per-node capacity is given by $\Theta(W \sqrt{n \log n})$, where $W$ denotes data rate of each wireless link and $n$ denotes the number of nodes in the network. With optimization of the geographic locations of nodes, the per-node capacity can be improved by a factor $\Theta(\sqrt{\log n})$ and approaches $\Theta(W \sqrt{n})$.

The investigation in [17] points out that the order of the number of the infrastructure, denoted by $N_G$, makes a different impact on the network capacity. If $N_G = \omega(n)$, the capacity can increase linearly with $N_G$. Otherwise, improvement in the capacity by $N_G$ is not so significant. The capacity in a two-tier network architecture like a WMN is observed in [18]. Similar results are obtained. When the number of arbitrary deployed IGWs is $\omega(\sqrt{N_R})$, where $N_R$ denotes the number of MRs, asymptotic capacity increases linearly with the number of IGWs. Since $N_R$ may be drastically smaller than $n$, the required IGWs in a WMN can be just few. However, these results and models are not directly applicable to a WMN due to its special attributes as the traffic in a WMN cannot be assumed to be uniformly distributed. Previous model may not be applicable or lead to unexpected results due

\(^1\)Knuth’s notation [14] is used to denote the order of numbers.
to modifications in this assumption. Furthermore, the performance of MRs with multiple radios will be much different from a single radio situation, which needs to consider both the network connectivity and the channel assignment problems.

Network delay is an important parameter for delay-sensitive applications [3], such as voice over IP (VoIP), video conference, etc. Nevertheless, the network delay in a WMN has not been well investigated. Although multi-hop transmission has analyzed in [19], the packet delay caused by errors during the transmission process, the transmission delay caused by interference has not been taken into account. The analysis in [20] models the queueing network delay in an ad hoc network as a G/G/1 queueing network. The network discussed in [20] is divided into multiple square areas, each of which has a node communicating with other nodes in neighboring squares. For solving this problem, a symmetrical attribute is assumed for the peer-to-peer traffic in the network. The multiple queues with regular topology network is reduced to a single queue system in [21], where the link transmission delay is assumed to be unity and the single queue is modeled as a G/D/1 queue. The work is primarily based on the connectivity graph. Asymmetric traffic and geometric attributes in a WMN have not been taken into account at all.

Network delay and capacity are tightly related metrics. A few works in the literature have considered this issue. The analysis in [22] enhanced this. However, only substantial transmission delay has been considered in [22] while queueing delay has been ignored, which could be a major factor when the traffic is heavy. The result in [22] is based on the assumption of uniformly distributed traffic in an ad hoc network. Intuitively, delay increases with throughput, but not linearly. When per-node throughput in a network approaches its capacity, the delay may go to infinity. The inter-related effect of the network delay and the capacity in a multi-hop network
still needs further research.

The main results in this chapter are summarized as follows:

- We perform the network delay and capacity analysis on self-organizing ad hoc MRs, which is uniformly i.i.d. Our model and analysis identify the features of a WMN with geometric attribute, traffic pattern, channel assignment, etc. The queueing delay for randomly deployed MRs is bounded by our proposed queue reduction model, which is shown to be feasible.

- Our results show that the network capacity in a WMN using \textit{ad hoc routing} is dominated by the number of MCs, the number of MRs, and the network delay constraint, which can be expressed as:

\begin{equation}
\Lambda'(n, N_R, T_D) = \Theta\left(\frac{W}{n} \sqrt{\frac{N_R}{\log N_R}}\right) - J'(n, N_R, T_D),
\end{equation}

where \( J'(n, N_R, T_D) = \Theta\left(\frac{L_{N_R}}{n \log N_R T_D}\right) \). When the network delay constraint is \( \Omega\left(\sqrt{\frac{N_R}{\log N_R}}\right) \), the network capacity is dominated by the numbers of MCs and MRs, which is \( \Theta\left(\frac{W}{n} \sqrt{\frac{N_R}{\log N_R}}\right) \).

- With \textit{clustered routing} in a WMN, the network capacity is dominated by the number of MCs, the number of IGWs, and the network delay constraint, which is given by:

\begin{equation}
\Lambda''(n, N_G, T_D) = \Theta\left(\frac{W}{n} N_G\right) - J''(n, N_G, T_D),
\end{equation}

where \( J''(n, N_G, T_D) = O\left(\frac{L N_G}{n T_D}\right) \). When the network delay constraint is \( \Omega(1) \), the network capacity is dominated by the numbers of MCs and IGWs, and is indicated by \( \Theta\left(\frac{W}{n} N_G\right) \).
• On the other hand, the network delay with \textit{ad hoc routing} is shown to be dominated by the number of MRs, which is expressed by:

\[ T_D' = \Theta\left(\sqrt{\frac{N_R}{\log N_R}}\right). \] (2.3)

The analytical results also show that the network delay with \textit{clustered routing} is given by:

\[ T_D'' = \Omega\left(\frac{L}{W}\right) \wedge O\left(\frac{L}{W} \sqrt{\frac{N_R}{N_G \log N_R}}\right). \] (2.4)

• Ad hoc network is shown to be a special case of our results. The network capacity and the delay are given by taking the number of MRs as \( n \).

• After comparing ad hoc routing with clustered routing in a WMN, we obtain an interesting result. When \( N_G = O\left(\sqrt{\frac{N_R}{\log N_R}}\right) \), ad hoc routing in a WMN will have a higher network capacity with a factor \( \Theta\left(\frac{1}{N_G} \sqrt{\frac{N_R}{\log N_R}}\right) \), but clustered routing has lower network delay with a factor \( \Omega\left(\sqrt{\frac{\log N_R}{N_R}}\right) \wedge O\left(\sqrt{\frac{1}{N_G}}\right) \).

The rest of this chapter is organized as follows: the network model is developed in Section 2.2. Section 2.3 analyzes the network delay and the capacity while considering various conditions. A detailed discussion of the derived results is given in Section 2.4. Finally, the chapter is concluded in Section 2.5.

2.2 Problem Formulation

To observe and analyze the problem, we first state the assumptions in this section. We consider such a model of WMN that \( N_R \) MRs are uniformly i.i.d. deployed in a unit square \( \mathcal{D} \) in \( \mathbb{R}^2 \) to provide network connection service for uniformly i.i.d.
n MCs in the same square. $N_G$ IGWs are arbitrary deployed in $D$. MCs are only allowed to connect to MRs or IGWs with one hop wireless link for network access. MRs are able to aggregate the traffic from MCs in its service region and forward it to a next hop MR or IGW. IGWs are interconnected through wired connections which has enough bandwidth for the traffic between any two IGWs. MRs and IGWs constitute the backbone of a WMN. The maximum data rate for backbone link between two MRs is $W$ bps. It assumes that the links between MCs and MRs, called access links, use different spectrum which does not interfere with the backbone links interconnecting MRs and IGWs. Backbone link $W$ can be divided into $c$ orthogonal channels, where each channel’s maximum data rate is $W/c$. Each MR is equipped with $m$ radios for communication in backbone, which can be switched to any channel by channel assignment.

The unit square can be divided into multiple smaller virtual square cells [23] which are denoted as $V = \{V_1, \ldots, V_p\}$. Without loss of generality (w.l.o.g.), we assume the number of cells $|V|$ is a square number. The edge length of a cell $V_i$ is denoted by $a = \sqrt{\frac{1}{|V|}}$. The packet arrival rate at cell $i$ is denoted by $A_i$ and the departure rate is denoted by $S_i$. The packet delay at cell $i$ is denoted by $D_i$.

Transmission may occur between two adjacent MRs whose distance is within their transmission distance, denoted by $r_T$. The receiver needs to operate one of the radios on the channel used for the transmission time period. Two transmitting pairs operating on the same channel may interfere with each other if they are very close. The signals from one of the pairs will be taken as the noise by the receiver of the other pair. Since the wireless signals exponentially attenuate with distance, if there is enough distance between two simultaneous transmitting MRs, the interference signals is negligible. Suppose MR $u$ transmits packets over a channel to an MR $v$,
then the transmitting packet is successfully received by MR $v$ if $|X(w) - X(v)| \geq (1 + \Delta)|X(u) - X(v)|$ for every other MR simultaneously transmitting over the same channel, where $w$ denotes the interfering MR and $\Delta$ is the parameter to determine the interference region. This is affected by various factors such as the threshold of signal to noise ratio (SNR), signal distance attenuating factors, etc. So, given transmission distance, $r_T$, we have interference distance $r_I = (1 + \Delta)r_T$.

We identify a traffic flow $i$ in the network by $\langle s_i, d_i \rangle$, where $s_i$ denotes the MR from which flow $i$ starts and $d_i$ denotes the MR at which flow $i$ ends. The route path of flow $i$ is denoted by $P_i := \langle v^0_i, \ldots, v^h_i \rangle$, where $v^j_i$ is the MR at the $j$th hop. In the express of route path, $s_i$ is the first MR $v^0_i$ and $d_i$ is the last MR $v^h_i$. The packet arrival rate at MR $k$ is denoted by $\lambda_k$. Note that the arrival packets at MR $k$ may come from multiple flows. The departure rate at MR $k$ is denoted by $\mu_k$. For packet delayed at MR $k$ includes queueing and transmission delays. Queueing delay is the time interval of the packet buffered at MR before transmission service and transmission delay is the time for packet to be sent to a next hop MR. We assume the bandwidth of wired connections between IGWs are wide enough, the transmission delay from an IGW to another is small enough to be ignored.

Some useful Lemmas are listed here. Proofs of Lemmas 1 and 2 are given in [24].

**Lemma 2.1.** Chernoff Upper Tail Bound: if $\xi$ is the sum of independent indicator random variables, then:

$$P[\xi \geq (1 + \delta)E[\xi]] \leq e^{-\frac{\delta^2 E[\xi]}{2 + \delta}}, \quad \delta > 0.$$  \hfill (2.5)

**Lemma 2.2.** Chernoff Lower Tail Bound: if $\xi$ is the sum of independent indicator
random variables, then:

\[ P[\xi \leq (1 - \delta)E[\xi]] \leq e^{-\frac{\delta^2 E[\xi]}{2}}, \quad \delta \in (0, 1]. \] 

(2.6)

2.3 Analytical Modeling on Network Delay and Capacity

MRs constitute a backbone of a WMN. MCs connect to a nearby MR for network access. Since the MRs are randomly deployed rather than pre-planed, we first need to observe the connectivity of MRs.

2.3.1 Network Connectivity

Unit square \( \mathcal{D} \) can be divided into \( \mathcal{V} \) virtual cells, as shown in Figure 2.1. The random points in the figure denote randomly deployed MRs. In formed virtual cells, packets are assumed to be transmitted from a cell to a neighboring cell. As shown in Fig 2.1, the MRs in the cell highlighted with gray shadow, can have communication links with MRs in neighboring cells, shown by lines between points in Figure 2.1. Given \( N_R \) MRs, with appropriately selected length of a cell, denoted by \( a \), the network is able to form a connected backbone network with high probability (w.h.p.), where a connected network means every MR in the network can reach any other MRs by single or multi-hop transmission.

Theorem 2.1. If \( a = \Theta(\sqrt{\log \frac{N_R}{N_R}}) \) and the transmission distance of MR \( r_T = \sqrt{5}a \), the backbone network formed by \( N_R \) uniformly i.i.d. MRs is connected w.h.p.
Given the transmission distance of MC $r_C = a = \Theta\left(\sqrt{\frac{\log N_R}{N_R}}\right)$, the whole two-tier WMN is connected w.h.p.

**Proof.** Taking $a = c_1 \sqrt{\frac{\log N_R}{N_R}}$, where $c_1$ is a positive constant. According to Chernoff lower tail bound, the number of MRs in a cell $V_i$ is bounded as:

$$P[H_i^R \leq c_2 \log N_R] \leq e^{-\frac{(1-c_2/c_1)^2\log N_R}{2}},$$

where $c_2$ is a positive constant. When $N_R$ approaches infinity, the right hand side (RHS) of the equation approaches zero. When $a \geq c_1 \sqrt{\frac{\log N_R}{N_R}}$, the number of MRs in a cell is more than $c_2 \log N_R$ w.h.p.

For any two adjacent cells $V_i$ and $V_j$, they can be contained by a circle with radius $\frac{\sqrt{5}a}{2}$, which means the maximum distance between two points in two adjacent cells is no more than $\sqrt{5}a$. So, given a circle centered at a point in $V_i$ with radius
\[ \sqrt{5a}, \text{ it contains all adjacent cells of } V_i. \] In other words, any MR in cell \( V_i \) can have communication links with MRs in adjacent cells, as the example shown in Fig 2.1. Since each cell has no less than \( c_2 \log N_R \) MRs \( w.h.p. \), the whole backbone network comprised by MRs is connected.

As every cell can connect to its neighboring cell, the network is obviously connected. We also assume the MCs’s transmission distance \( r_C \) is equal to \( a \), which is \[ r_C = \Theta\left(\sqrt{\frac{\log N_R}{N_R}}\right). \] MCs in a virtual cell \( V_i \) will connect to MRs in the same cell in one hop transmission. Since every cell has no less than \( c_2 \log N_R \) MRs, MC in the network can find an MR for network access connection \( w.h.p. \). The traffic from an MC to another MC goes through the backbone network and route on grid cells, continuously relayed from one cell to another.

Therefore, such a two tier WMN is connected \( w.h.p. \).

2.3.2 Number of Routing Hops

The number of routing hops is essential to the network capacity and delay. The advantages of multi-hop transmission are lower transmission power, higher data rate, lower device complexity, etc. However, while an MR relays the packets, it costs additional spectrum resource, which may decrease the network capacity. In addition, queueing delay also introduced at the relaying nodes, which may prolong the network delay.

The effects of multi-hop transmission are mainly determined by the routing strategy. There are two kinds of routing strategies in a WMN: Ad hoc routing and clustered routing. With ad hoc routing, the routing from one MR to another is searched for the shortest path in the backbone. IGWs in a WMN can form multiple clusters,
where a cluster consists of an IGW and multiple MRs around it. In clustered routing, the packets at MRs are forwarded towards its cluster head (IGW), and vice versa.

**Lemma 2.3.** The number of routing hops with ad hoc routing is expressed by:

$$h' = \Theta(\sqrt{|V|}) = \Theta(\sqrt{\frac{N_R}{\log N_R}}).$$  (2.8)

**Proof.** Traffic with ad hoc routing is between a pair of MRs in the network. When $N_R$ is relatively large, each MR has the same probability to be a source or a destination. A cell can be indexed by two integers $(i, j)$, where $i$ denotes the sequence on x-axis and $j$ denotes the sequence on y-axis. Then, an identical $(i, j)$ uniquely identifies a cell in the network.

Randomly picking up two MRs on grid $(m, n)$ and $(i, j)$ as the source MR and the destination MR, respectively. The number of hops for the routes on the grids is at least $h = |m - i| + |j - n|$. Considering possibilities of all such pairs, the lower bound of the average number of hops $h_L$ is computed as follows:

$$E[h_L] = E[|m - i|] + E[|j - n|] = 2E[E[|m - i||i]],$$  (2.9)

where $E[|m - i||i]$ can be calculated by:

$$E[|m - i||i] = \frac{1}{\sqrt{N_R}}(\sum_{m=1}^{i-1}(i - m) + \sum_{m=i}^{\sqrt{N_R}}(m - i))$$

$$= \frac{i^2}{\sqrt{N_R}} - (\frac{1}{\sqrt{N_R}} + 1)i + \frac{\sqrt{N_R} + 1}{2}. \quad (2.10)$$

$E[i^2]$ is computed by following equation:

$$E[i^2] = \frac{(\sqrt{N_R} + 1)(2\sqrt{N_R} + 1)}{6}.$$

(2.11)
Substituting $E[i^2]$ from Eq. (2.11) into Eq. (2.10), we obtain:

$$E[E[m - i|i]] = \frac{1}{3} (\sqrt{N_R} - \frac{1}{\sqrt{N_R}}).$$  \hspace{1cm} (2.12)

Thus, the lower bound of average number of hops is:

$$\bar{h}_L = E[h_L] = \frac{4}{3} (\sqrt{|\mathcal{V}|} - \frac{1}{\sqrt{|\mathcal{V}|}}).$$  \hspace{1cm} (2.13)

A feasible routing solution serves as the upper bound for the average number of hops, denoted by $\bar{h}_U$. Traffic can be distributed and balanced on the routes available within the minimum rectangular area constructed by the source and the destination MRs. The number of hops with such a feasible routing method is equal to the number of hops of the shortest path. Thus, the upper bound of the average number of hops is at the same order of the lower bound. The lemma holds.

With clustered routing, the number of routing hops is directly related to the number of IGWs in the network. We show the clustered routing with an example in Figure 2.2. Assume that there are four IGWs and they equally divide the network into four clusters. The squares in the figure denote virtual cells. Each IGW located at the center cell of its cluster. The cell relays the traffic towards or from the cluster head (IGW). The number in the small square is to denote the number of hops from the cell to the IGW. For example, all cells around the IGW has the number ‘1’ in the square, which means they can transmit packets to the IGW in one hop. The most far away cells in the figure are four hops away from the IGW. The gray level of square also helps show the nearness of the cell to the IGW, where deeper gray color means that the cell is closer to the IGW.
Lemma 2.4. The number of routing hops with clustered routing is given by:

\[ h'' = O\left(\sqrt{\frac{|V|}{N_G}}\right) = O\left(\sqrt{\frac{N_R}{N_G \log N_R}}\right). \]  

(2.14)

Proof. IGW is assumed to be a cluster head of MRs within a square consisting of \(M^2\) cells. Given \(N_G\) IGWs, \(M\) satisfies:

\[ \frac{1}{a\lfloor \sqrt{N_G} \rfloor} \leq M \leq \frac{1}{a\lceil \sqrt{N_G} \rceil}, \]

(2.15)

where \(\lfloor x \rfloor\) is to take the maximum integer no greater than \(x\) and \(\lceil x \rceil\) means to take the minimum integer no less than \(x\).

The number of routing hops in clustered routing \(h''\) is no more than \(M\). So, we have:

\[ h'' = O\left(\frac{1}{a\lfloor \sqrt{N_G} \rfloor}\right) = O\left(\sqrt{\frac{|V|}{N_G}}\right) = O\left(\sqrt{\frac{N_R}{N_G \log N_R}}\right). \]

(2.16)
From Lemmas 2.3 and 2.4, we see that the average number of routing hops increases with the number of MRs in the network, either by ad hoc routing or clustered routing. While comparing $h'$ with $h''$, clustered routing can reduce the average number of hops by at least a factor of $\Theta\left(\frac{1}{\sqrt{N_G}}\right)$, with functionality of IGWs.

### 2.3.3 Traffic on Cells

The MRs in a cell relay not only the traffic of its associated MCs but also the traffic from other cells. While the traffic load increases, the queueing delay is increased nonlinearly. It is important to avoid traffic congestion at certain cells. Routing strategy highly affects the traffic loads at cells. Appropriate routing strategy can distribute traffic on different cells more evenly, so the network delay can be effectively decreased and the network capacity can be enhanced. As there are two major types of routing strategies in a WMN, we investigate their traffic load on the cells.

**Lemma 2.5.** Cell traffic load with ad hoc routing is expressed by:

$$A'_j = \Theta(n \lambda(n) \sqrt{\log \frac{N_R}{N_D}}).$$  \hspace{1cm} (2.17)

**Proof.** For a flow path $P_i$, the probability that an intermediate MR $v^k_i$ is in a given cell $V_j$ is indicated by $1^k_i(j)$, which is equal to 1 if $v^k_i$ is in cell $V_j$ and equal to 0 otherwise. Then, the total number of intermediate MRs falling in cell $V_j$ is computed by:

$$N'_j = \sum_{i=1}^{n} \sum_{k=1}^{h_i} 1^k_i(j).$$  \hspace{1cm} (2.18)
If a flow has an intermediate MR in cell \( V_j \), it means the flow passing through the cell. So, the expected number of flows passing through cell \( V_j \) is given by:

\[
E[N'_j] = E[\sum_{i=1}^{n} \sum_{k=1}^{h_i} 1^k_i(j)] \\
= E_h[\sum_{i=1}^{n} E[\sum_{k=1}^{h_i} 1^k_i(j)|h_i]].
\]

(2.19)

Note that \( 1^k_i(j) \) and \( 1^m_i(j) \) \((i \neq l)\) are i.i.d. On the other hand, \( 1^k_i(j) \) and \( 1^m_i(j) \) \((k \neq m)\) are identical distributed but dependent. Since a path only passes a cell once, we have \( P[1^k_i(j) = 1, 1^m_i(j) = 1] = 0 \).

According to the symmetry of the cells, we have \( P[1^k_i(j) = 1] = \frac{1}{|V|} \). So, we obtain:

\[
E[N'_j] = nE_h[\frac{h_i}{|V|}].
\]

(2.20)

The expected number of flows passing through cell \( V_j \) is:

\[
E[N'_j] = \Theta\left(\frac{n}{\sqrt{|V|}}\right) = \Theta(an).
\]

(2.21)

Based on Chernoff upper tail bound, we have:

\[
P[N'_j \geq (1 + c_3)an] \leq e^{-\frac{c_3^2an}{2+3}},
\]

(2.22)

where \( c_3 \) is a positive constant. When \( n \) goes to infinity, the RHS of the equation approaches zero. So, we have:

\[
N'_j = \Theta(an).
\]

(2.23)

The traffic load at cell \( V_j \) is sequentially derived, which is expressed by:

\[
A'_j = \Theta(an\lambda(n)) = \Theta(n\lambda(n)\sqrt{\frac{\log N_R}{N_R}}).
\]

(2.24)
The cell traffic is more complex than that in ad hoc routing as we take the edge effect of clusters inside the unit square into consideration. We give Lemmas 2.6 and 2.7 for the number of cells at the $k$th hop of a cluster and the number of MCs in cell, respectively. Then, we compute cell traffic in Lemma 2.8.

**Lemma 2.6.** Given a cluster formed by $M^2$ (assuming $M$ is an odd number w.l.o.g.) small square cells, the number of cells that is $k$ hops away from the center is:

$$N_{hk} = \begin{cases} 4k, & k = 1, \ldots, \frac{M-1}{2} \\ 4(M-k), & k = \frac{M+1}{2}, \ldots, M-1. \end{cases}$$

(2.25)

**Proof.** Identify the cells in a cluster by $(i, j)$, where $i$ and $j$ are integers, and $i$ denotes the location in x-axis and $j$ denotes the location on y-axis. The scope of $i$ or $j$ is the integers from $-\frac{M-1}{2}$ to $\frac{M-1}{2}$. The center cell having cluster head (IGW) is denoted by $(0, 0)$. The minimum number of hops from a cell $(i, j)$ in cluster to the head is $k = |i| + |j|$. The scope of $k$ is an integer in $[1, M-1]$. We have expression that $|j| = k - |i|$. According to the scope of $j$, we have

$$0 \leq k - |i| \leq \frac{M-1}{2}. \quad (2.26)$$

Combining with the scope of $i$, we obtain:

$$\max\{k - \frac{M-1}{2}, 0\} \leq |i| \leq \min\{k, \frac{M-1}{2}\}. \quad (2.27)$$

For different $k$, we have

$$\begin{cases} 0 \leq |i| \leq k, & k \leq \frac{M-1}{2} \\ k - \frac{M-1}{2} \leq |i| \leq \frac{M-1}{2}, & k > \frac{M-1}{2}. \end{cases} \quad (2.28)$$
When $k \leq \frac{M-1}{2}$, there are $2k + 1$ different integers for $i$ to satisfy the inequality. For each $i$, its corresponding value of $j$ satisfying $|i| + |j| = k$ is two, except for $|i| = k$. So, the number of cells at $k$ ($k \leq \frac{M-1}{2}$) hops away is $2((2k+1) - 2) + 2 = 4k$.

When $k > \frac{M-1}{2}$, there are $2(M - k)$ integers for $i$ in the scope. For each $i$, its corresponding value of $j$ satisfying $|i| + |j| = k$ is two, so we can obtain the total number of cell at $k$ ($k > \frac{M-1}{2}$) hops away, which is $4(M - k)$.

Lemma 2.7. The number of MCs in a cell $V_j$ is $\Theta(\frac{n \log N_R}{N_R^2})$.

Proof. Let $n^j_C$ denotes the number of MCs in a cell $V_j$. According to Chernoff lower tail bound, we have:

$$P[n^j_C \leq (1 - c_4)a^2n] \leq e^{-\frac{c_4^2a^2n}{2}},$$

where $c_4$ is a positive constant no more than one. Since $N_R = O(n)$, when $n$ approaches infinity, the RHS of the equation goes to zero. So, the number of MCs in a cell is $\Omega(\frac{n}{|V|})$. On the other hand, $\sum_{j=1}^{|V|} n^j_C = n$, so the number of MCs in a cell is no higher than the order of $\Theta(\frac{n}{|V|})$. Thus, the result is yielded that $n^j_C = \Theta(\frac{n \log N_R}{N_R^2})$.

Lemma 2.8. Cell traffic $A''_{h_j}$, with clustered routing, is given by:

$$A''_{h_j} = \begin{cases} 
\Theta(\frac{n \lambda(n) \log N_R 2M^2-(2j-1)^2-1}{2j}), & j \leq \frac{M-1}{2} \\
\Theta(\frac{n \lambda(n) \log N_R M-j+1}{2}), & j > \frac{M-1}{2}.
\end{cases}$$

Proof. The traffic in a cluster is assumed, w.l.o.g, to be towards the IGW. For a cell $j$ hops away from the IGW, it needs to relay its traffic and traffic of $k$ ($k > j$) hops away cells. According to the symmetry, the number of the $j$th hop cell is $\frac{\sum_{k=j}^{M-1} N_{h_k}}{N_{h_j}}$.

So, we have:
Substituting Eq. (2.25) in Lemma 2.6, we obtain:

\[ A''_{h,j} = \Theta \left( \frac{n\lambda(n) \sum_{k=j}^{M-1} N_{h_k}}{|\mathcal{N}| N_{h_j}} \right). \]  \hspace{1cm} (2.31)

which then yields the result.

2.3.4 Cell Queue Model

There is a number of MRs in a virtual cell. MR in a neighboring cell can select any one of the MRs for the traffic relay. We show that multiple queues of MRs in a virtual cell can be reduced to a single queue and the average packet delay of the single queue is the lower bound of packet delay in a cell. As shown in Figure 2.3(a), the dotted square denotes a cell and there are several queues belonging to the MRs inside the cell. Every MR maintains a queue to receive the traffic from the MRs in neighboring cells, denoted by arrow line going into the cell, and relays the traffic to MRs in neighboring cells, denoted by arrow line going out of the cell. We use
a single queue to make reduction for the multiple queues in a cell as illustrated in Figure 2.3(b). The arrival traffic of the single queue is the summation of the arrival traffic of other MRs in the cell. It will be shown by Theorem 2.1 that the queueing delay of the reduced single queue is the lower bound of multiple queues and this is an achievable lower bound.

Theorem 2.1. A single queue can model the achievable lower bound of packet queueing delay at a cell.

Proof. Assume the maximum packet departure rate of cell $V_i$ is denoted by $R_i$. The departure rate of queue $k$ of MR in cell $V_i$ is denoted by $\mu_{i,k}$. The length of queue MR $k$ in cell $V_i$ at time $t$ is denoted by $q_{i,k}(t)$. Then, we have:

$$q_{i,k}(t + \Delta t) = q_{i,k}(t) + A_{i,k}(\Delta t)|_t - B_{i,k}(\Delta t)|_t,$$  \hspace{1cm} (2.33)

where $A_{i,k}(\Delta t)|_t$ denotes the number of arrival packets for time duration $\Delta t$ after time point $t$ and $B_{i,k}(\Delta t)|_t$ denotes the number of departure packets in this duration.

The total number of packets in a cell with traffic distributed into multiple queues is expressed by:

$$\tilde{Q}_i(t + \Delta t) = \sum_{k=1}^{N_i^R} q_{i,k}(t + \Delta t)$$

$$= \sum_{k=1}^{N_i^R} (q_{i,k}(t) + A_{i,k}(\Delta t)|_t - B_{i,k}(\Delta t)|_t)$$

$$= \tilde{Q}_i(t) + A_i(\Delta t)|_t - \sum_{k=1}^{N_i^R} B_{i,k}(\Delta t)|_t.$$  \hspace{1cm} (2.34)

On the other hand, we denote the length of the single queue of cell $V_i$ by $Q_i$. The length of single queue is given by:

$$Q_i(t + \Delta t) = Q_i(t) + A_i(\Delta t)|_t - B_i(\Delta t)|_t,$$  \hspace{1cm} (2.35)
where \( A_i(Δt) \) denotes the number of arrival packets and \( B_i(Δt) \) denotes the number of departure packets of cell \( V_i \) during time duration \( Δt \). The maximum number of packets can be departed from cell \( V_i \) during \( Δt \) is denoted by \( I(Δt) \). So, the departure packet with single queue reduction is given by:

\[
B_i(Δt) = \begin{cases} 
I(Δt), & Q_i(t) + A_i(Δt) > I(Δt) \\
Q_i(t) + A_i(Δt), & \text{otherwise}
\end{cases}
\]

(2.36)

We prove the delay of single queue reduction is the lower bound of multiple queue delay by mathematical induction.

**Basic case:** At \( t = 0 \), we have \( Q_i(0) = \tilde{Q}_i(0) \).

**Induction hypothesis:** Assume the theorem holds at time \( t \) such that \( Q_i(t) \leq \tilde{Q}_i(t) \).

**Induction Step:** If \( Q_i(t) + A_i(Δt) \) \( ≥ \) \( I(Δt) \):

\[
Q_i(t + Δt) = Q_i(t) + A_i(Δt) - I(Δt) \\
\leq Q_i(t) + A_i(Δt) - \sum_{k=1}^{N_i^p} B_{i,k}(Δt) \\
\leq \tilde{Q}_i(t) + A_i(Δt) - \sum_{k=1}^{N_i^p} B_{i,k}(Δt) \\
= \tilde{Q}_i(t + Δt);
\]

else

\[
Q_i(t + Δt) = 0 \leq \tilde{Q}_i(t + Δt).
\]

(2.37)

Therefore, the theorem holds for \( Q_i(t) \leq \tilde{Q}_i(t) \). A single queue model of cell can serve as the lower bound for packet delay at a cell.
Any MRs can connect to other MRs in an adjacent cell within one hop. Since we can perform such a single queue at one of the MRs in a cell, it is also an achievable bound.

Having a single queue in a cell, the network connectivity of multiple queueing model shown in Figure 2.4(a) can be changed to that illustrated in Figure 2.4(b), where Figure 2.4 is a part of network from Figure 2.1. After using single queue model, the network connectivity is still ensured and every cell has link to connect to every neighboring cell. In addition, such a model can greatly reduce the complexity of link scheduling. With multiple queue model, the number of links of a cell to its neighboring cells is $\Theta(\log N_R)$. When using single queue model, it is $\Theta(1)$.

### 2.3.5 Network Delay and Capacity

Since we have a single queue model for each cell, we observe the hop delay with the assumption that the cell traffic goes to a single queue. The transmission delay from a cell to another is related to the number of cells in the interference region. Lemma 2.9 shows that the number of cells in the interference region is bounded.

**Lemma 2.9.** The number of cells that interfere with each other, is no more than $5\pi(1 + \Delta)^2$.

**Proof.** Given MR $u$ in $\mathcal{D}$, the interference region is a disk centered at $u$ with radius $r_I = (1 + \Delta)r_T$, denoted by $C(u, r_I)$. The number of cells in an interference region, denoted by $N_V$, is given by:

$$N_V \leq \frac{\pi r_I^2}{a^2} = 5\pi(1 + \Delta)^2. \quad (2.39)$$
When the number of channels, $c$, is greater than $N_V$, MRs in the interference region can transmit with conflict free-scheduling. The channel assignment can be achieved by modified graph coloring method [25].

With channel assignment, the achievable throughput of a cell is $\frac{W}{N_V}$ bps. Assume the length of a packet is $L$. Then, the transmission delay of a packet is $\frac{LN_V}{W}$. The packet departure rate at a cell $k$ can achieve:

$$\mu_k = \frac{W}{LN_V}. \quad (2.40)$$

**Theorem 2.2.** Given a self-organizing WMN using ad hoc routing, the per-MC capacity is given by:

$$\Lambda'(n, N_R, T_D) = \Lambda'(n, N_R) - J'(n, N_R, T_D), \quad (2.41)$$

where $\Lambda'(n, N_R) = \Theta\left(\frac{W}{n} \sqrt{\frac{N_R}{\log N_R}}\right)$ and $J(n, N_R, T_D) = \Theta\left(\frac{LN_R}{n \log N_R T_D}\right)$.

**Proof.** The delay at a cell is approximated by:

$$D'_k = \frac{1}{\mu_k - \lambda_k}, \quad (2.42)$$

where $\lambda_k$ is packet arrival rate of the queue.
The network delay is then given by:

\[ T_D = h' D'_k = \frac{h'}{\mu_k - \lambda_k}. \]  

(2.43)

From above equation, we have following derivation:

\[ \lambda_k = \mu_k - \frac{h'}{T_D} \]

\[ \lambda_C = \frac{1}{c_5n a} (\mu_k - \frac{h'}{T_D}). \]  

(2.44)

Since the network capacity \( \Lambda' = \Theta(\max\{\lambda_C L\}) \), we can obtain per-MC capacity with expression:

\[ \Lambda'(n, N_R, T_D) = \Theta\left(\frac{L}{n a} (\mu_k - \frac{h'}{T_D})\right) \]

\[ = \Theta\left(\frac{L}{n} \left(\sqrt{\frac{N_R}{\log N_R}} \left(\frac{W}{L \sqrt{N_R}} - c_6 \sqrt{\frac{N_R}{\log N_R T_D}}\right)\right)\right). \]  

(2.45)

While delay constraint is not a dominating factor, i.e., \( T_D = \Omega(\sqrt{\log N_R}) \), the capacity can be expressed by:

\[ \Lambda'(n, N_R) = \Theta\left(\frac{W}{n} \sqrt{\frac{N_R}{\log N_R}}\right). \]  

(2.46)

While given delay constraint \( T_D \), the capacity degradation is given by:

\[ J'(N_R, T_D) = \Theta\left(\frac{L N_R}{n \log N_R T_D}\right). \]  

(2.47)

The network capacity is then computed by Eqs. (2.46) and (2.47), which is given by:

\[ \Lambda'(n, N_R, T_D) = \Lambda(n, N_R) - J(n, N_R, T_D). \]  

(2.48)
Corollary 2.1. In a WMN with ad hoc routing, if $T'_D = \Omega(\sqrt{\frac{N_R}{\log N_R}})$, per-MC capacity is dominated by the number of MCs, $n$, and the number of MRs, $N_R$, which is $\Lambda'(n, N_R, T'_D) = \Lambda'(n, N_R)$.

Proof. With constraints on the delay, per-MC capacity is decreased with $J'(n, N_R, T'_D)$. If $T'_D = \Omega(\sqrt{\frac{N_R}{\log N_R}})$, we have an interesting result that $J'(n, N_R, T'_D)$ has an order that is not higher than $\Lambda'(n, N_R)$. Thus, $\Lambda'(n, N_R, T'_D) = \Lambda'(n, N_R)$ holds. \qed

When the number of MRs approaches the number of MCs, the model is applicable to ad hoc networks.

Corollary 2.2. In an ad hoc network, if $T_D = \Omega(\sqrt{\frac{n}{\log n}})$, per-MC capacity is dominated by the number of MCs, which is $\Lambda(n) = \Theta(\frac{W}{\sqrt{n \log n}})$.

Proof. We can see from Theorem 2.2 that per-MC capacity, $\Lambda(n, N_R)$, with ad hoc routing is mainly affected by the numbers of MCs and MRs. While take the number of MRs equal to $n$, there exists a one-to-one mapping between MR and MC. Since MRs have the same geometric distribution as that of MCs, our results can be used for the random ad hoc network. If the number of MRs approaches the number of MCs $n$, it goes to $\Theta(\frac{W}{\sqrt{n \log n}})$, which is the asymptotic capacity of a random multiple hop wireless network. \qed

Theorem 2.3. Given a self-organizing WMN using ad hoc routing, the network delay is expressed by:

$$T'_D = \Theta\left(\sqrt{\frac{N_R}{\log N_R}}\right). \quad (2.49)$$

Proof. From Eq. (2.44), we can have the network delay computed by:

$$T'_D = \frac{c_7 N_R}{c_8 \lambda C} - \frac{n \lambda C}{c_6 n \lambda C}. \quad (2.50)$$
When $\lambda_C = O\left(\frac{W}{n} \sqrt{\frac{N_R}{\log N_R}}\right)$, the network delay is given by:

$$T_D' = \Theta\left(\sqrt{\frac{N_R}{\log N_R}}\right).$$

(2.51)

Since the capacity of network with ad hoc routing is $O\left(\frac{W}{n} \sqrt{\frac{N_R}{\log N_R}}\right)$, given by the proof in Theorem 2.2, the theorem holds.

**Theorem 2.4.** Given a self-organizing WMN using clustered routing, the per-MC capacity is given by:

$$\Lambda''(n, N_G; T_D) = \Theta\left(\frac{W}{n} N_G\right) - J''(n, N_G; T_D),$$

(2.52)

where $J''(n, N_G; T_D) = O\left(\frac{L_{N_G}}{n T_D}\right)$.

**Proof.** The network delay is bounded by:

$$T_D \geq \frac{1}{\mu^{(1)} - \lambda^{(1)}},$$

(2.53)

and

$$T_D \leq \frac{h''}{\mu^{(1)} - \lambda^{(1)}},$$

(2.54)

where $\mu^{(1)}$ and $\lambda^{(1)}$ denote the packet departure rate and packet arrival rate of MR that one hop away from the IGW, respectively.

On one hand, from Eq. (2.53), we have:

$$\lambda^{(1)} \leq \mu^{(1)} - \frac{1}{T_D}.$$  (2.55)

Substituting $\lambda^{(1)}$ by last hop clustering traffic, we obtain:

$$c_7 n \lambda_C a^2 M^2 - \frac{1}{4} \leq \mu^{(1)} - \frac{1}{T_D},$$

$$\lambda_C \leq \frac{1}{c_7 n \frac{c_a}{N_G} - a^2 \left(\frac{W}{L N_V} - \frac{1}{T_D}\right)},$$

$$\lambda_C \leq \frac{1}{c_7 n \frac{c_a}{N_G} - c_{fg} \frac{\log N_R}{N_R} \left(\frac{W}{L N_V} - \frac{1}{T_D}\right)}.$$  (2.56)
The number of cells in a network is $|\mathcal{V}| = \frac{1}{\sigma^2} = \Theta\left(\frac{N_R}{\log N_R}\right)$, so the number of IGWs in the network satisfies:

$$N_G = O\left(\frac{N_R}{\log N_R}\right). \quad (2.57)$$

If network delay is not a dominating factor, i.e., $T_D = \Omega(1)$, we can obtain the following equation from Eqs. (2.56) and (2.57):

$$\Lambda''(n, N_G) = O\left(\frac{W}{n} N_G\right). \quad (2.58)$$

From Eq. (2.54), we obtain:

$$\lambda^{(1)} \geq \mu^{(1)} - \frac{k''}{T_D} \quad (2.59)$$

$$\lambda_G \geq \frac{1}{c_7n} \left[ \frac{4}{N_G} - c_9 \frac{\log N_R}{N_R} \right] \left( \frac{W}{LN_V} - c_{10} \frac{\sqrt{N_R}}{T_D} \right) \sqrt{N_G \log N_R}$$

When $T_D = \Omega\left(\sqrt{\frac{N_G \log N_R}{N_R}}\right)$, we get:

$$\Lambda''(n, N_G) = \Omega\left(\frac{W}{n} N_G\right). \quad (2.60)$$

Since $N_G = O\left(\frac{N_R}{\log N_R}\right)$, we can take $T_D = \Omega(1)$ for the conditions of both Eqs. (2.58) and (2.60). Then, we can obtain:

$$\Lambda''(n, N_G) = \Theta\left(\frac{W}{n} N_G\right). \quad (2.61)$$

The impact of delay is shown by the equation:

$$J''(n, N_G, T_D) = O\left(\frac{LN_G}{nT_D}\right). \quad (2.62)$$
Corollary 2.3. If $T_D = \Omega\left(\frac{L}{W}\right)$, per-MC capacity is $\Theta\left(\frac{W}{n} N_G\right)$.

Proof. When $T_D = \Omega\left(\frac{L}{W}\right)$, $J''(n, N_G, T_D) = O\left(\frac{W}{n} N_G\right)$. Thus, the corollary is yielded. □

Theorem 2.5. Given a self-organizing WMN utilizing clustered routing, the network delay is expressed as:

$$T_D'' = \Omega\left(\frac{L}{W}\right) \land O\left(\frac{L}{W} \sqrt{\frac{N_R}{N_G \log N_R}}\right). \quad (2.63)$$

Proof. Extended from Eqs. (2.53) and (2.54), the network delay is shown by:

$$\frac{c_{10} N_G}{c_9 \frac{N_G W}{L} - n \lambda_C} \leq T_D'' \leq \frac{c_{11} \sqrt{\frac{N_G N_R}{\log N_R}}}{c_9 \frac{N_G W}{L} - n \lambda_C}. \quad (2.64)$$

When $\lambda_C = O\left(\frac{W}{n} N_G\right)$, the network delay is then sequentially given by:

$$T_D'' = \Omega\left(\frac{L}{W}\right) \land O\left(\frac{L}{W} \sqrt{\frac{N_R}{N_G \log N_R}}\right). \quad (2.65)$$

□

2.4 Discussion and Implications

In a self-organizing WMN, MR is able to discover neighboring MRs and create wireless links with them. The backbone is formed by these MRs in an ad hoc fashion. Due to higher flexibility and lower deployment cost of a WMN, it has attracted more researchers to conduct theoretical investigation and practical consideration for the next generation wireless network. The analytical process and derived results in this chapter lead to several interesting and useful conclusions.
Our queue reduction model shows that an appropriate topology control is needed in a self-organizing WMN. We can visualize the MR as a group of formed virtual cells. While all available MRs in a cell participate in packet forwarding, the average packet delay may be prolonged if the packets cannot be evenly distributed among the MRs in that cell. It is relatively hard to distribute packets evenly to all MRs in a virtual cell and the performance could not be better than that in a single queue model, as shown earlier. So, it is better to forward packets only to a selected MR in a virtual cell. With MR selection, the backbone network topology is also changed. Many wireless links can be removed, which can effectively reduce the complexity of the channel assignment in a WMN. Without any queue reduction, the number of potential wireless links to neighboring cell is $\Theta(\log N_R)$. If all links are kept, the channel assignment has to consider allocation for all these links. With single queue model, only one MR is selected in a cell, which reduces the number of wireless links connecting to neighboring cells to $\Theta(1)$. Then, a much faster channel assignment can be performed.

The routing strategies also play an essential role on the network capacity and the delay. Ad hoc routing and clustered routing are affected by different attributes of network delay and the capacity. With ad hoc routing, the network delay is $\Theta(\sqrt{\frac{N_R}{\log N_R}})$. The network delay increases with the number of MRs. While the number of MRs increases, the size of virtual cells will be decreased. To route a given distance, packet delivery needs a larger number of hops, which introduces more queueing and transmission delays. By increasing the number of MRs, the required transmission distance to maintain a connected network is reduced. Consequently, the interference region is also decreased. So, the number of simultaneous transmission can be increased, which has the potential to enhance the network capacity. When the
network delay is allowed to be $\Omega(\sqrt{\frac{N_R}{\log N_R}})$, the per hop delay is of the order $\Omega(1)$. Then, the delay becomes a non-dominating factor for the network capacity and network capacity of $\Theta\left(\frac{W n}{\sqrt{N_R \log N_R}}\right)$ is achieved. If the delay constraint in the network is $o\left(\frac{N_R}{\log N_R}\right)$, the allowed per hop delay will be quickly reduced with an increase in the number of MRs, which leads to the delay-constrained capacity goes to zero. Thus, delay constraint $o\left(\sqrt{\frac{N_R}{\log N_R}}\right)$ is not suitable for a large scale WMN with ad hoc routing.

With clustered routing, the network delay becomes $\Omega\left(\frac{L}{W}\right) \land O\left(\frac{L}{W} \sqrt{\frac{N_R}{N_G \log N_R}}\right)$. Increasing the number of IGWs can effectively reduce the delay caused by an increase in the number of MRs. If $N_G = \Theta\left(\sqrt{\frac{N_R}{\log N_R}}\right)$, the number of hops from any MR to an IGW will be bounded by a constant and the network delay will be $\Theta(1)$. When the network delay is $\Omega(1)$, network capacity with clustered routing is $\Theta\left(\frac{W n}{N_G}\right)$, which is nearly linear with the number of IGWs. Since the number of virtual cell is $O\left(\frac{N_R}{\log N_R}\right)$, the number of IGWs is also $O\left(\frac{N_R}{\log N_R}\right)$. With ad hoc routing, the network capacity with clustered routing is lower or equal to the order of that.

Comparing ad hoc routing with clustered routing in a WMN, we find that: if $N_G = O\left(\sqrt{\frac{N_R}{\log N_R}}\right)$, ad hoc routing in a WMN will have higher network capacity with a factor $\Theta\left(\frac{1}{N_G} \sqrt{\frac{N_R}{\log N_R}}\right)$ but clustered routing has lower network delay with a factor $\Omega\left(\sqrt{\frac{\log N_R}{N_R}}\right) \land O\left(\frac{1}{N_G}\right)$. The results indicate that ad hoc routing is more preferred by bandwidth-sensitive applications if $N_G = O\left(\sqrt{\frac{N_R}{\log N_R}}\right)$, and clustered routing is a better option for delay-sensitive applications.
2.5 Conclusion

In this chapter, we have derived the network delay and the capacity for a two-tier self-organizing WMN. The analysis of network is based on forming virtual cells, and multi-hop transmission is used to relay in neighboring cells. Queue reduction is used to model multiple queues in a cell. The results in this chapter provide the range of network delay and capacity with different types of routing strategies. A comparison of results indicate that different routing strategies can be selected according to the scale of a WMN. Furthermore, some of our results are equally applicable to the analysis of ad hoc networks.
Chapter 3

Centralized Channel Assignment in WMN Backbone

3.1 Introduction

Network capacity could be seriously constrained the interference happened during multi-hop relay because of channel contention and collision. Multi-radios could improve the network capacity by operating radios at non-overlapping channels, separating MR radios over different frequency bands and eliminating the probability of channel collisions. In addition, operating multiple radios in distinct channels enables simultaneous transmission and/or reception. Orthogonal channels have to be carefully allocated. Otherwise, network partition may result that the radios of neighboring MRs may operate in different channels without any common channel for connection establishment between them. The channel assignment also needs to consider allocation in a way to accommodate the network traffic, so as to balance the traffic over different channels.

Finding an optimal channel assignment is NP-hard [12]. Channel assignment is-
Sue can be formulated as a graph coloring problem or its variations [13][26]. To have a better modeling of the assignment problem with consideration of traffic demands, Mishra et al. [27] further modeled it as an weighted edge-coloring problem. Solving a given graph coloring problem can use one of known edge coloring algorithm. However, while using graph coloring algorithm, it actually assumes the number of colors/channels is sufficient, which does not address the constraint of channel numbers. In practical situation, it is possible that the number of channels very is limited. Then, the implement of coloring algorithm cannot be accomplished if the density of nodes is high, meaning that the available channels cannot guarantee adjacent edges to be assigned with different colors.

On the other hand, from the point view of traffic in the network, WMN has distinguished traffic pattern as compared to an ad hoc network. The traffic in a WMN is predominantly directed between the MR and the IGW due to the rich Internet services, different from peer-to-peer traffic dominating an ad hoc network. The MR whose location is closer to the IGW usually has much higher traffic load than that of the MR at the network boundary because all the Internet traffic from the MR away from the IGW have to be forwarded to the MR located closing to the IGW. Sequentially, the traffic load becomes significantly different. The special traffic pattern in WMN architecture leads to a challenging issue in channel assignment.

In this chapter, we discuss the channel assignment and take the number of orthogonal channels into account. Our method computes the minimum number of channels needed for a conflict-free channel assignment [25]. Given WMN network graph $G$ and the corresponding traffic profile (i.e., traffic demands), allocation of traffic flows over the MRs and their links can be modeled as a linear programming problem, which targets that each MR has the same proportional traffic that will be
successfully forwarded to the IGW. With objective to allocate fair flows, we form a flow-based conflict graph for further channel assignment. The scheme needs to compute the minimum number of required channels with which a flow-based graph can be colored in a way that adjacent edges does not have a same color. In the case that the number of available channels is not sufficient to perform edge coloring, our algorithm can do a priority-based channel assignment. While the shortage of the channel number happens, channel aggregation allows a set of radios with light traffic to share a common channel even though these radios are within each other’s interference region. The selection of radios to perform channel aggregation is based on the traffic load information, which is estimated while forming the flow-based conflict graph. This procedure indicates that the radio having less traffic is more likely to be selected to share a channel. Of course, a channel should be able to be reassigned to a radios if the traffic profile in the network is significantly changed.

The rest of this chapter is organized as follows: Related work is in Section 3.2. Network model is given in Section 3.3. Section 3.4 establishes a linear programming problem for a given network and discusses the channel assignment algorithm on the traffic-based conflict graph. Section 3.5 evaluates the proposed algorithm by extensive simulation. This chapter is concluded in Section 3.6.

### 3.2 Related Work

The main purpose of channel assignment is to assign multiple channels to the multiple radios for higher throughput benefit in the network. In the past few years, a number of channel assignment schemes have been proposed and can be implemented in several ways. Linear programming was used [28] to model and address...
the throughput maximization problem, which is based on using a non-trivial ideal scheduling mechanism. Instead of linear programming, conflict graph \[29\] was proposed to find the conflict free links in a wireless network. Graph coloring was used \[30\] in modeling routes and scheduling flows in a network.

There are three major approaches to do channel assignment: static, dynamic, and hybrid. Channels will be assigned to radios for a long duration in a static approach \[12\][28]. As the assignment will last for a long time period, static assignment is good for the situation that network and traffic are statistically stable. On the other hand, static channel assignment lacks flexibility in channel usage and needs effort to adjust a whole network allocation while varying channel conditions and network traffic. In the dynamic approach \[31\], a radio is allowed to switch and operate on multiple different channels, which can be adaptive to the channel interference conditions. Efficient synchronization is required in dynamic channel assignment and fast channel switch is also required. Hybrid approach \[13\][32] uses both static and dynamic approaches, the radios are divided into two sets: one set is statically assigned in certain designed channels while the other set of channels is dynamically switchable in deploying channels.

Channel assignment usually takes network traffic into account for better channel utility. The past observations show that MRs located closer to an IGW would be able to transmit a larger volume of traffic to the Internet for their users than that of MRs away from the IGW. Thus, an objective to optimization in channel assignment should enhance per MR throughput instead of the whole one.
3.3 Network Model

3.3.1 Network Architecture

We consider a WMN backbone with the Internet-connection, which consists of MRs and IGWs. IGW is the gateway that provides the Internet connection for the whole wireless backbone formed by MRs and their links. Each MR can transmit packets to one of the IGWs by multi-hop fash. The bandwidth of the wired connection at the IGW is regarded as a very big number of times of the wireless links. This means that the WMN throughput is only constrained by the wireless links. Each MR is equipped with multiple radios, capable of switching and operating on multiple different channels. MR can transmit and receive data packets from neighboring MRs simultaneously by employing multiple radios. In multi-radio multi-channel environment, the establishment of a wireless link needs two compatible radios at two separate MRs are within their transmission ranges and operating on a common channel. Otherwise, the two MRs are disconnected even if in the vicinity of each other. Wireless backbone of a WMN can be further divided into multiple local domains [33] [34]. The division could depend on the relation among the MRs and the IGWs in the network. A local domain is led by an IGW and a number of MRs that are attached to the IGW with multi-hop paths. For clarification, we define such a local domain as an undirected network connectivity graph $G(V, E)$, where $V$ represents a set of nodes (i.e., multiple MRs and an IGW) and $E$ denotes a set of undirected edges. An edge between two nodes represents the geometric distance relationship, meaning if they are within the transmission range of each other. An edge between two MRs is not necessary to mean a link but they have the potential for establishing a wireless link.
In other words, a communicating link can be established by employing a channel assigned to the edge. The number of radios at MR \( u \in V \) is denoted by \( N_{\text{radio}}(u) \) and the set of radios operated by MR is given by \( \text{Radio}(u) = \{1, 2, \ldots, N_{\text{radio}}(u)\} \). The set of orthogonal channels in the network is expressed by \( CH = \{1, 2, \ldots, N_{\text{ch}}\} \). We assume that every radio has the same transmission range, which is denoted by \( RT \).

### 3.3.2 Interference Model

While two MRs (i.e., \( u \in V \) and \( v \in V \)) are within their transmission range of each other, a common channel \( c \in CH \) has to be established for two compatible radios so as to support the wireless link between them. Due to broadcast nature of omnidirectional antenna that are implemented by most current radios, wireless signals will interfere with each other if they are concurrently employing a same channel within the interference range. The interference could cause serious collision to the packet transmission and in sequence significantly degrade the link throughput. If two wireless links employ two different orthogonal channels, the MR pairs will allow simultaneously transmission or reception of packets even they are located within each other’s interference range. Considering the throughput of a channel \( k \in CH \) is \( W^k \) shared by the nodes within the interference range, the maximal throughput in the interference range is \( W = \sum_{k=1}^{N_{\text{ch}}} W^k \).

We denote the interference range by \( R_I \), which is \( q (q \geq 1) \) times of transmission range \( RT: R_I = q \times RT \). To spatially share a common channel \( c \in CH \) for transmission, IEEE 802.11 CSMA mechanism can be implemented. Before data transmission the radio using IEEE 802.11 CSMA senses if the accessing channel is
idle or not. While the channel is idle, the radio transmits packet to its neighboring MR. Otherwise, the transmission is backoff until the channel is available. Given two MRs \( u \in V \) and \( v \in V \) that are neighboring, i.e., \( d(u, v) \leq R_T \), where \( d(u, v) \) is the physical distance between them, the CSMA mechanism allows \( u \in V \) and \( v \in V \) to establish a link over edge \( uv \in E \) and avoid co-channel interference. There are no other transmission allowed in the interference region while the transmission of packets happens between \( u \) and \( v \). This is the interference model used by this chapter in generating a flow-based conflict graph.

### 3.3.3 Traffic Model

Multi-hop transmission enables WMN to serve relative larger area as compared to that of a WLAN. The WMN traffic is firstly originated at MCs and then is collected at MRs which again deliver the data packets to the destination. The traffic could be heterogeneously scattered over a network. Thus, hot spots may be formed, where the traffic is heavy. MC may exhibit mobility by changing its location continuously. Based on mobility, MC is further grouped into two types: static MC and mobile MC. Static MC stays or deployed at certain location for a long duration. We can denote a static MC traffic demand as variable \( x \). Considering \( N_s \), the number of static mesh clients attached to one MR \( u \ (u \in V) \), the originated traffic at a static MC is represented by \( \varphi_s(u) = \sum_{k=1}^{N_s} x_k \). On the contrary, when MC moves from one MR to another, the traffic of the mobile MC will be migrated from an earlier MR to the newly attached MR accordingly. If there are \( N_m \) number of mobile MCs attached to an MR, the originated traffic of mobile MCs can be represented by \( \varphi_m(u) = \sum_{k=1}^{N_m} x_k \). The total aggregated traffic in a MR \( u \) is \( \varphi(u) = \varphi_s(u) + \varphi_m(u) \).
This shows that the volume of the aggregated traffic in MR $u$ is determined by the static MCs and the mobile MCs. If there is $\rho(u)$ relaying traffic in MR $u$, the total traffic flow is denoted as $f(u) = \varphi(u) + \rho(u)$, indicating the time-varying traffic in MR $u$. Due to variance in the traffic demand on MR $u$, it requires that channel assignment to adjust according to the variation in the total traffic demand, i.e., $f(u)$.

There are two types of traffic in $u \in V$ in traffic direction: Internet traffic and local traffic. The Internet traffic is directed between the MCs and the IGW. In contrast to the Internet traffic, the local traffic is a type of peer-to-peer traffic. It is between two MRs without traveling through the IGW. Due to rich Internet services recently, we consider the percentage of the local traffic is very small as compared to that of Internet traffic [34][35]. $f(u)$ is used to represent the Internet traffic.

### 3.4 Proportional Flow and Channel Assignment

Flow-based channel assignment is detailed in this section, which includes three stages:

- **Stage 1**: Proportional flow formulation: performing the link scheduling for traffic flow over wireless links. While doing the scheduling, it estimates the ration of the scheduled traffic to the traffic demands at every MRs and achieves proportional fairness at every MR. It formulates the problem of finding the fair flow in the network as a linear programming.

- **Stage 2**: Conflict-free channel assignment: the flow-based graph is obtained from the flow graph in this step. Then, we use graph coloring algorithm to get the required number of channels. The coloring process is equivalent to find
the conflict free assignment in the flow-based conflict graph. So the result obtained from coloring applies to the conflict free channel assignment problem.

- **Stage 3: Channel aggregation:** in a random network, it is possible that the nodes are located somewhere with high density. The scheme needs a channel aggregation algorithm to deal the situation that the available number of channels is less than the minimum number for a conflict free channel assignment. We implement a channel aggregation algorithm to reduce a conflict free channel assignment to an assignment with channel sharing at certain region.

- **Our network has a mechanism that adaptively re-assigns and adjusts the channel assignment if the traffic demands at the MRs change significantly.**

3.4.1 **Proportional Flow**

At the beginning of channel assignment, the IGW is in the charge of discovering the network topology. It can use the network information to construct a connectivity graph $G$. At the same time, the IGW also collects the traffic demands of MRs in its management region. We define $\lambda$ ($0 < \lambda \leq 1$) as a traffic proportion indicator, which means each MR will transmit $\lambda$ proportion of its total aggregate traffic. The objective of traffic scheduling in channel assignment is to maximize this indicator $\lambda$. Given any two neighboring MRs, for example, $u$ and $v$ ($u, v \in V$), use $s_{uv}^{k,t}$ as a binary variable, where $t$ represents a time slot used for scheduling of transmission over channel $k$ ($k \in CH$). $s_{uv}^{k,t} = 1$ means link $uv$ is active for packet transmission at time slot $t$ and employs channel $k$. Otherwise, $s_{uv}^{k,t} = 0$. To achieve optimization, the first stage of channel assignment can be further formulated as a Linear Programming
problem subjected to the following constraints.

**MR-radio constraint:** Radios are used for data transmission and/or reception. Thus, the total number of active links for a MR at a given time slot cannot exceed the total number of radios at the MR. While expressing this relationship in mathematics, we have the following equation: Eq. (3.1):

\[
\sum_{k} \sum_{v \in N(u)} (s_{uv}^{k,t} + s_{vu}^{k,t}) \leq N_{Radio}(u), \quad \forall u \in V,
\]  

(3.1)

where \( N(u) \) represents the set of neighbors of MR \( u \) and \( N_{Radio}(u) \) denotes the number of radio at MR \( u \).

**MR-Interference Constraint:** Given an interference distance, which is \( R_I = q \times R_T \), the number of simultaneous flows at the interference region is constrained by \( c(q) \), where \( c(q) \) is a constant that depends on \( q \) and denotes the maximum number of simultaneous transmission in interference region. Given two neighboring MRs, \( u \in V \) and \( v \in V \), Eq. (3.2) gives the MR-interference allocation constraint:

\[
s_{uv}^{k,t} + \sum_{ij \in I^k(uv)} s_{ij}^{k,t} \leq c(q), \quad \forall uv \in E,
\]  

(3.2)

where \( ij \in E \) denotes the link of two neighboring MRs \( i \) and \( j \) within the interference range of \( u \) or \( v \), and \( I^k(uv) \) is the set of links within the combined interference range of \( u \) and \( v \).

We can find out the value of \( c(q) \) by following way. The union area of two circles centered at \( u \) and \( v \) with radius \( R_I \), denoted by \( C_u \) and \( C_v \), respectively, is the interference region formed by link \( uv \). For simplicity purpose, instead of using two circle areas (\( C_u \) and \( C_v \)), the interference region is approximated by a circular area centered at the middle of link \( uv \) with a radius equal to \( R_I + 0.5R_T \), denoted by
$C_{uv}$, which serves as a loose upper bound for the interference region. The region affected by signal transmission can be denoted by a circle with radius $0.5R_I$ and such a circle is denoted by $C_{R_I}$. As to assign conflict free channels needs at least $R_I$ distance apart for every link, the number of simultaneous transmissions in $C_{uv}$ can be obtained by computing the maximum number of such circles that can fill in the $C_{uv}$ without overlapping. The value of $c(q)$ is the maximum number of $C_{R_I}$ that can be filled in the circle $C_{uv}$ without non-overlapping. Given ratio $q$, the approach to compute the value of $c(q)$ is further given in [36].

The integer variable constraints such as $s_{uv}^{k,t}$ let the liner programming become an integer linear programming problem which is intractable in general. We need to relax the linear programming problem by following way. We can derive the link flows at corresponding channel $k$ by using Eq. (3.3) over a time period $T$, where $W^k$ is the bandwidth of channel $k$.

$$f_{uv}^k = \frac{W^k}{T} \sum_{0 \leq t < T} s_{uv}^{k,t}, \quad \forall uv \in E, k \in CH. \quad (3.3)$$

Then, the MR-radio constraint and MR-interference constraint are expressed by:

$$\sum_{v \in N(u)} (f_{uv} + f_{vu}) \leq N_{radio}(u)W^k, \quad \forall u \in V, \quad (3.4)$$

$$f_{uv}^k + \sum_{ij \in P(uv)} f_{ij}^k \leq c(q)W^k, \quad \forall uv \in E. \quad (3.5)$$

Eq. (3.4) provides the bi-directional flow (i.e., $f_{uv}$ and $f_{vu}$) constraint in the network and Eq. (3.5) shows the flow constraint at interference range. The number of radios at MR should be less or equal to the total number of channels as there is no sense to have extra radios. Even if $N_{radio}(u)$ is larger than $N_{CH}$, the maximal
bandwidth is still $W$. So, $N_{\text{radio}}(u)W^k \leq W$. If we add up all flows scheduled over all available channels, Eq. (3.6) is obtained from Eq. (3.5):

$$f_{uv} + \sum_{i,j \in I(uv)} f_{ij} \leq c(q)W, \quad \forall uv \in E.$$ (3.6)

Considering flows in the network confirm flow conservation law that the volume of the incoming traffic equals to that of the outgoing traffic, which is given by the following Eq. (3.7) and Eq. (3.8).

$$\sum_{v \in N(u)} (f_{uv} - f_{vu}) = \lambda g(u), \quad \forall u \in V/IGW.$$ (3.7)

where $V/IGW = V - V_{IGW}$, denoting the MR set that removes $IGW$ from $V$. The traffic at the IGW is:

$$\sum_{v \in N(IGW)} f_{vIGW} = \lambda \sum_{u \in V/IGW} g(u).$$ (3.8)

The maximum volume of flows on an edge is constrained by the capacity of the edge, which is expressed by Eq. (3.9):

$$0 \leq f_{uv} \leq W, \quad \forall uv \in E.$$ (3.9)

where the maximum bandwidth of an edge is $W$.

Taking all constraints together and the objective to maximize proportional parameter $\lambda$, we have the problem formulation here:

$$\max \lambda$$
According to the traffic model, \( f(u) \) denotes the total traffic at node \( u \in V \): 

\[ f(u) = \varphi(u) + \rho(u), \quad u \in V. \]

On the other hand, \( f(u) \) is given by \( \sum_{v \in N(u)} f_{uv} + f_{vu} \). The two equations above indicate the relationship between the traffic demands and the traffic flows. Given the traffic demand at each MR, the proportional fair flows can be obtained by solving the linear programming problem under constraints in Eq. (3.10) with the objective to maximize proportion parameter \( \lambda \). After obtaining the traffic flow and \( G(V, E) \), a flow graph \( G_f(V, E_f) \) can be further generated from \( G(V, E) \) by removing the link not having any scheduled traffic. While traffic is predominant with Internet access, many links in \( G(V, E) \) will be removed while transforming to flow graph \( G_f(V, E_f) \). It is noted that in the process of scheduling traffic flow, we relax the linear programming problem. To have the scheduling to be feasible, we need process channel assignment with the next two stages.

### 3.4.2 Conflict Free Channel Assignment

Flow-based conflict graph \( G_c(V_c, E_c) \) can be further derived from the above flow graph \( G_f(V, E_f) \) that is obtained by solving the linear programming problem. We first depict the procedures of creating \( G_c(V_c, E_c) \) from \( G_f(V_f, E_f) \), which includes three steps:
For each flow link $uv (uv \in E_f)$, a corresponding node is generated in $G_c(V_c, E_c)$ to represent the flow link. For example, flows $f_A$, $f_B$ and $f_C$ in Figure 3.1(a) are the three distinct nodes in Figure 3.1(b), $A$, $B$, and $C$ respectively. Each flow link in $G_f(V, E_f)$ has a corresponding node represented in $G_c(V_c, E_c)$.

- If two flow links in $G_f(V, E_f)$ are within interference region of each other, an edge is accordingly established in $G_c(V_c, E_c)$ between two nodes that represents the two flow links. For example, $f_A$ and $f_C$ are within the interference region of each other, an edge is therefore added to connect node $A$ and node $C$ in $G_c(V_c, E_c)$.

- A weight will be assigned to every node in graph $G_c$. The weight is determined by the corresponding flow volume in flow graph $G_f$. We denote the set of weights in $G_c$ as $w(V_c)$.

Conflict link graph $G_c(V_c, E_c)$ is actually very useful to reflect the complicated interference relationship among flow links. For two distinct flow links $uv$ and $xy$ at nodes $u$, $v$, $x$ and $y$, if any of the distances $d(x, u)$, $d(x, v)$, $d(y, u)$ and $d(y, v)$ is not larger than $R_I$, the links would interfere with each other. The interference relationship can be expressed by $I(uv, xy) = 1$. $I(uv, xy) = 0$, otherwise. Without loss of generality, considering two nodes $A$ and $B$ in $G_c$, and their corresponding flow links are $uv$ and $xy$, respectively, edge $ab (ab \in E_c)$ exists if and only if $I(uv, xy) = 1$, which indicates their interfere. Take the example in Fig 3.1(a), where exists three wireless flow links : $f_A$, $f_B$, and $f_C$. The dotted line in Fig 3.1(a) represents a link without flow, and thus this link is not included in $E_f$. Consequently, only three links in Figure 3.1(a), i.e., $f_A$, $f_B$, and $f_C$, are transferred to nodes in conflict graph as
Figure 3.1: Flow-based Conflict Graph

shown in Figure 3.1(b) namely, A, B, and C respectively. The edge between A and B indicates that $f_A$ and $f_B$ are interfering each other.

Establishing the flow-based conflict graph is to transfer the channel assignment on flow graph $G_f$ to the node coloring problem in conflict graph $G_c$. Conflict free channel assignment is then to perform for assigning different channels to all links within interference region in flow-based conflict graph $G_f$. It assigns available channels to the nodes in $G_c(V_c, E_c)$ in a way such that any two adjacent nodes in $G_c(V_c, E_c)$ are assigned different channels, which is equivalent to do coloring in a graph.

Now, we have a look the graph coloring problem that existing for long time. The graph coloring problem itself is a NP-hard problem [37]. In general, a greedy coloring algorithm can be used for problem solving. The usage and selection of greedy algorithm will depend on the attributions of the problem as the performance of a greedy algorithm is strongly related to the graph. We can use node coloring algorithm to find a near optimal channel assignment. We use following way to perform the node coloring in the conflict graph where each node in $V_c$ is assigned a channel and there is no neighboring node in $G_c$ using the same channel. The details of greedy algorithm are as follows. It starts the coloring from node sequential enumer-
ation of $V_c$: \{v_1, ..., v_{n_c}\}, where $n_c$ denotes the number of nodes in $G_c$. At the same time, an index is used for channel to indicate the usage of the channel, e.g., 0, 1, ...$k$, where $k$ means the channel has been used $k$ times so far. The greedy algorithm visits the nodes in an order and assigns each $v_i$ with the first available channel, having the smallest channel index and not used by any neighbor of $v_i$. While considering $v_i$, the algorithm ignores all neighbors $v_j$ of $v_i$ for $j > i$. Assume that $G_c^i$ is a subgraph of $G_c$ and built from $G_c^{i-1}$ by adding $v_i$ as well as the corresponding edges. The number of channels will be minimized if $v_i$ has minimum degree in $G_c^i$. Therefore, it firstly chooses node $v_{n_c}$ with the smallest degree in $G_c$, and next chooses the node $v_{n_c-1}$ with the smallest degree in $G_c^{n_c-1}$ formed by removing $v_{n_c}$ and the corresponding edges from $G_c$. It continues to choose $v_i$ with the smallest degree in $G_c^i$ and forms $G_c^{i-1}$ by removing $v_i$ as well as the corresponding edges from $G_c^i$ until $v_1 = G_c^1$. Considering each node in sequence \{v_1, ..., v_{n_c}\}, it assigns each node with the first available channel (e.g., the channel that is least used in nodes and not used by any assigned neighbors).

In this node coloring problem, the optimal number of colors $\chi(G_c)\) is bounded by Eq. (3.11) [37]:

$$\text{ch}(G_c) \leq \chi(G_c) \leq \frac{1}{2} + \sqrt{2|E_c| + \frac{1}{4}},$$

(3.11)

where $\text{ch}(G_c)$ denotes the largest clique size in graph $G_c$. The clique size of a graph means the number of nodes in the largest clique of the graph. This bound is also applicable to channel assignment. The lower bound is trivial to prove. The number of optimal channels is greater than or equal to the largest clique size of $G_c$. The upper bound is related to the number of edges $|E_c|$. If we assume that $G_c$ is optimally assigned by channels. All nodes in $G_c$ form $\chi(G_c)$ channel sets that
each channel set consists of nodes using the same channel. As any two neighboring
nodes have no common channel and the graph has full connectivity, at least one
edge exists between any two channel sets. Thus, we can have $|E_c| \geq \left(\frac{2}{\chi(G_c)}\right) = \frac{1}{2}\chi(G_c)(\chi(G_c) - 1)$, which serves as the upper bound.

In our greedy coloring algorithm, the number of channels $\chi_{\text{greedy}}(G_c)$ is less than or equal to the maximization of the smallest node degree in subgraph $G'_c$:

$$\chi_{\text{greedy}}(G_c) \leq \max_{G'_i \subseteq G_c} \{ \min_{v_i \in G'_i} \{ \text{deg}(v_i) \} \} + 1. \quad (3.12)$$

The upper bound of the minimum number of required channels is provided in Eq. (3.12) for the greedy coloring algorithm, which is an approximation solution using $\chi_{\text{greedy}}(G_c)$ channels. The efficiency of the approximation algorithm can be indicated by the approximation ratio, which is given by:

$$\frac{\chi_{\text{greedy}}(G_c)}{\chi(G_c)} \leq \frac{\max_{G'_i \subseteq G_c} \{ \min_{v_i \in G'_i} \{ \text{deg}(v_i) \} \} + 1}{\chi(H)}. \quad (3.13)$$

While $\frac{\chi_{\text{greedy}}(G_c)}{\chi(G_c)}$ is closer to 1, the greedy coloring algorithm is approaching to an optimal solution.

### 3.4.3 Channel Aggregation

While the number of available channels $N_{ch}$ is less than $\chi_{\text{greedy}}(G_c)$, the conflict free channel assignment is not feasible due to the shortage of channels. In this case, if coloring two adjacent nodes with a same color, indicating two neighboring flow links share a common channel, we can have shared channel in the network. While enable the share of channel at two neighboring flow links, the sharing issue can be solved by the MAC protocol to avoid packet collision. While doing a channel sharing, the channel assignment may not be conflict free. Thus, the objective for channel
sharing in channel assignment is to minimize interference. In this section, we introduce a channel aggregation algorithm that assigns channel with low interference.

Let us show an example in Figure 3.2 to illustrate the main idea in channel aggregation. In Figure 3.2(a), there is a flow-based conflict graph obtained from Figure 3.1. Assume that three orthogonal channels with flows \( f_A > f_B > f_C \). After the greedy channel assignment algorithm, we assign channels 1, 2, and 3 to nodes A, B, and C respectively. If we only have two channels, e.g., 1 and 2, we have to replace the node using channel 3 with one of the available channels, 1 or 2. In this case, we consider two questions: (i) which node should be selected for channel aggregation, and (ii) which channel should be used in channel aggregation.

We choose the node which has the minimum weight in the flow-based conflict graph. The reason is that the node having the minimum weight will cause less interference to the channel which it is going to be united to. In the example in Figure 3.2, we choose the node \( f_C \) as it has the minimum flow compared to \( f_A \) and \( f_B \) (i.e., \( f_A > f_B > f_C \)).

After the node is chosen, e.g., \( f_C \), we consider the selection of channel. In order to avoid high interference, our criteria is to choose a channel with less introduced interference. We evaluate the interference on each node and choose the channel.
having less interference introduced by the newly added node. In this example, let us consider channel 1 for node $C$ (i.e., $f_C$), denoting the interference at node $A$ on channel 1 as $Int^1(A)$, where the interference is from the neighboring nodes of $A$ in the conflict graph and the set of neighboring nodes is denoted as $N(A)$. Then, we compute $Int^1(A) = Int^1(C) = f_A + f_C$. The degree of interference is approximated by the volume of the traffic flow. The reason behind this is that more traffic leads to more interference in the network. If we assign $f_C$ with channel 2, the introduced interference would be $f_B + f_C$. Because of $f_B + f_C < f_A + f_C$, we determine to assign $f_C$ by channel 2. Therefore, we color node $B$ and $C$ with a same color in
Figure 3.2(b) and (c), which means flows $B$ and $C$ share the same channel.

Algorithm 3.1: Channel Aggregation Algorithm

**input**: $G_c(V_c, E_c, w(V_c))$: weighted conflict graph

**input**: $CH(V_c)$: channel assignment on nodes

**output**: $CH'(V_c)$: adjusted graph

1. $CH(V_c) := \{CH_1, CH_2, \ldots \}$, $CH_i$: node set on channel $i$

2. **while** $|CH(V_c)| > N_{ch}$ **do**

3. **foreach** $CH_i$ **do**

   4. $w(CH_i) \leftarrow \sum_{v \in CH_i} w(v)$;

   **end**

6. Sort $CH(V_c)$ ascendingly according to $w(CH_i)$;

7. $CH_c \leftarrow CH(V_c).pop();$

8. **foreach** $CH_i \in CH(V_c)$ **do**

9. $U \leftarrow CH_i \cup CH_c;$

10. **foreach** $u \in U$ **do**

11. $Int^i(u) \leftarrow w(u) + \sum_{v \in N(u), CH(v) \in U} w(v);$  

12. **end**

13. $Int^i \leftarrow \max_{u \in U}\{Int^i(u)\};$

14. **end**

15. Find $k$ such as $\min_{1 \leq k \leq |CH(V_c)|} Int^k$;

16. $CH_k \leftarrow CH_k \cup CH_c;$

17. **end**

The channel assignment firstly performs the greedy algorithm no matter if the
number of channels is sufficient or not. If the required channels is higher than the
available number, it then lets some nodes to share the channel until the number of
assigned channels is equal to the available channels. In other words, if $\chi_{\text{greedy}}(G_c)$
is larger than the number $N_{ch}$, it initiates the channel aggregation algorithm as illus-
trated in Algorithm 3.1. In the algorithm, $V(CH_k)$ is the set of nodes assigned with
channel $k$ in $G_c$. While the nodes in $V(CH_l)$ are selected to use the same channel
as the node in $V(CH_k)$, the introduced interference to $V(CH_k)$ is modeled as:

$$I_{nt}^k(u) = f_u^k + \sum_{v \in N(u)} f_v^k, \quad u \in V(CH_k) \cup V(CH_l).$$  \hspace{1cm} (3.14)

The maximum interference introduced on channels $k$ after the nodes in $V(CH_l)$
switch the channel is given by:

$$I_{nt}^k = \max_{u \in V(CH_k) \cup V(CH_l)} \{I_{nt}^k(u)\}. \hspace{1cm} (3.15)$$

Algorithm 3.1 performs an iterative search to find the channel that has the min-
umum $I_{nt}^k(u)$. Once the number of assigned channels ($|CH(V_c)|$) is greater than
the number of available channels $N_{ch}$, i.e., $|CH(V_c)| > N_{ch}$ at line 2, the algorithm
searches the channel set that has the minimum weight (from line 3 to line 6). The
corresponding channel ($CH_c$) is then selected for aggregation, at line 7. Algorithm
3.1 computes the possible maximum interference at every channel set after potential
aggregation and records the channel set $CH_k$ (from line 8 to line 15), which do the
minimization to the maximum interference. At the end of each iteration, the selected
channel set is aggregated to the recorded channel set, $CH_k \leftarrow CH_k \cup CH_c$ (line
16). After aggregating channel, the search continues until the number of channels
($|CH(V_c)|$) in use is equal to the number of available channels $N_{ch}$.
3.4.4 Adaptive Channel Re-assignment

Although the traffic demands exhibit the stability property in terms of long term average, it may vary with hourly or daily basis. For example, it is possible that less traffic in the commercial area at night but higher traffic at residential area at the same time. When the traffic demands change significantly, we perform channel re-assignment in accordance with the changes in traffic demands. To perform re-assignment, it needs to have a new flow graph and a new proportional parameter $\lambda$. The network trigger of channel reassignment can be determined by traffic history in the network area. Alternatively, it can be triggered by in a real-time manner with monitoring the network traffic. We have a simple way here for illustration. The IGW can take the responsibility in monitoring the aggregated traffic on the links adjacent to it. For instance, the IGW monitors the fraction of the traffic in each channel. The total traffic on channel $k$ is expressed below:

$$f^k = \sum_{v \in N(IGW)} f^k_{vIGW}. \quad (3.16)$$

We defined $p^k$ the traffic fraction on channel $k$:

$$p^k = \frac{f^k}{\sum_{i \in CH} f^i}. \quad (3.17)$$

As indicated by Eq. (3.17), if $p^k$ is less than a predefined threshold, the channel re-assignment can be triggered.

3.5 Performance Evaluation

Simulation-based performance evaluation is conducted in this section, with three types of topology: grid, hexagon, and random. The distance between two neighbor-
ing MRs is given by $R_T$ in the grid and hexagonal topology. As we deploy all MRs randomly in the region, we choose the MR that mostly geometrically close to the center-point of the network as the IGW. The Internet traffic on each MR confirms Poisson distribution. We compare the minimum number of required channels in these three topologies and then look at the throughput degradation after performing channel aggregation.

We vary the number of MRs from 10 to 40 in the grid, hexagonal and random topologies. Every MR and IGW are configured with two wireless radios. The minimum number of channels required for a conflict-free channel assignment is shown in Figure 3.3. Given the network as the experimental scenario, the conflict-free channel assignment is hard to achieve if $\chi_{\text{greedy}}(G_c)$ is less than the channel number in Figure 3.3. In the grid and hexagonal topologies, $\chi_{\text{greedy}}(G_c)$ is almost constant irrespective of any increase in the number of MRs. Figure 3.3 also shows that the hexagonal topology needs more channels as it has a higher node connectivity than that of grid topology. On the contrary, in a random topology $\chi_{\text{greedy}}(G_c)$ is gracefully increased with the increase of the number of MRs. On the other hand, the random topology has a larger $\chi_{\text{greedy}}(G_c)$ than the other two regular topologies has.

The average degree of connectivity and flow graphs is shown in Figure 3.4, where the connectivity graph is denoted by $G$ and the flow graph is expressed by $G_f$. With an increasing number of MRs, the node degree in the flow graph keeps almost consistent with factor 2, which indicates the flow graph is able to form a tree structure rooted at the IGW. The tree structure requires fewer number of channels and results in an efficient network structure in terms of the channel assignment. The average node degree in the connectivity graph formed by hexagonal topology is greater than that of the random topology. However, the number of required channels
in hexagonal topology is fewer than that of a random topology. This is because the number of channels is primarily determined by the maximum number of MRs in the interference range. While the number of nodes is fixed, random topology may have more MRs in an interference range due to random factor, which leads to more channels is required at those dense region. Although it is true that the required channels at certain region is fewer, the determination of the number of required channel is the maximum number of required channels in overall network.

If the number of available channels is fewer than $\chi_{\text{greedy}}(G_c)$, we use the proposed channel aggregation. As $\chi_{\text{greedy}}(G_c)$ is 6 in the grid topology, obtained from previous results, we assume the number of available channels to vary from 1 to 6 as shown in Figure 3.5. The number of MR $N$ are chosen from 20, 30, and 40. In Figure 3.5, the throughput degradation appears once the number of channel is greater than four or five. The reason behind this is that the throughput is mainly constrained

![Figure 3.3: Number of Required Channels](image-url)
Figure 3.4: Average Degree

Figure 3.5: Channel Aggregation in Grid Topology
Figure 3.6: Channel Aggregation in Hexagon Topology

Figure 3.7: Channel Aggregation in Random Topology
by the throughput of the IGW instead of the number of channels. If the number of channels is less than 4 or 5, the throughput degradation becomes serious, especially when the number of channels is limited to just one.

Figure 3.6 shows the throughput results with different size of hexagonal topology. As seen from Figure 3.3, the networks deployed with different number of MRs: 20, 30, and 40. $x_{\text{greedy}}(G_e)$ is 7, 8 and 9 respectively. It shows same trend as that in the grid topology, throughput decreases as the number of channel becomes small.

The results of throughput degradation in the random topology is given in Figure 3.7. When the number of channels varies from 10 to 1, the throughput degradation begins at the number five and becomes more serious while further decreasing the channel numbers.

### 3.6 Conclusion

In this chapter, we address a flow-based channel assignment in a channel constrained network, and propose a channel assignment taking the number of available channels into account. To have fair/proportional flows, we use linear programming to model and solve the problem and the channel assignment is further implemented in terms of available channels. The network performance is evaluated with three types of different topologies and the results show how much the number of available channels affect network performance.
Chapter 4

Distributed Channel Assignment in WMN Backbone

4.1 Introduction

This chapter introduces a distributed channel assignment that uses radio partition and queue assignment approach for the MRs in a WMN [38]. MAC protocols such as IEEE 802.11 and IEEE 802.16, divide the frequency bands into multiple channels. Operating the radios of MRs over different channels can substantially enhance the capacity by increasing the number of simultaneous transmissions, and decrease the end-to-end delay by channel diversity along the path. MR equipped with multiple radios, is capable of more effectively utilizing multiple channels to transmit multiple packets simultaneously. The problem to efficiently manage and assign channels to radios, is very challenging. The number of simultaneously operating channels by each MR is restricted by its radio number. The strategy that operates all radios in assigned channels for a long duration is called as a static channel assignment. Although it is easy to implement in a pre-planed WMN with reasonable size, static
channel assignment may also lead to low spectrum utility efficiency because it becomes hard to adapt the change in the traffic while the network size is large. Static channel assignment has to be conducted carefully, as this kind of strategy may get the network partitioned because the number of neighboring MRs allowed for connection is constrained by the number of available channels. In a worst case, some essential links may be missing, which leads to network partition. Dynamically switching of radios on different channels can help for better performance with adaption to the traffic changes. Pure dynamic channel assignment also suffers from some critical problems due to frequent switching. Section 4.2 covers the related works on the static and dynamic channel assignment.

To meet the requirements in a self-organizing WMN, we propose an approach with distributed channel assignment. It is featured with division on multiple radios into two sets: the receiving radios and the forwarding radios. The receiving radios statically operate on certain channels for a long duration and is mainly used to receive packets. The forwarding radios are designed to switch dynamically over different channels and are mainly used to forward the packets. The radios can be adaptively switched from one set to the other as per the traffic demands and changes. Using radio partition can effectively avoid synchronization problem in channel switching which is typical issue in a dynamic channel assignment. Our approach ensures the network connectivity and adjusts the channels used by the radios according to the traffic. The radio switching between the sets enables the control of MR throughput and facilitates traffic balancing in the given network. The detailed discussion of radio management is given in Section 4.4.

To further improve network performance, we implement a queue management in multi-radio multi-channel WMNs. The radio switching from one channel to an-
other will introduce the time delay due to hardware limitation. This delay is called switching delay, which varies from several milliseconds to hundred milliseconds [39] [40] for current commercial radios. It is relatively small portion and can be ignored in a static channel assignment as the radio does not switch often in a static channel assignment. However, this a relative large proportion in dynamic channel switching because the radio is switched frequently. During the channel switching, there is no transmission is available. Therefore, our approach manages the queue and adjusts the packet forwarding sequence for forwarding radios so as to eliminate the overhead. The detailed investigation is given in Section 4.5. Extensive simulation in Section 5.5 illustrates high efficiency and very low overhead of the proposed schemes. This chapter is concluded in Section 5.6.

4.2 Related Work

To improve the throughput and reduce the delay, multiple radios equipped by MRs need to appropriately utilize the channels. Several approaches have been proposed in literature. The radios in [12] [28] are distributed to different channels and are operated on the assigned channels for a long duration. It intends to minimize the interference in every channel and maximize the throughput from a global view. Implementing such a static channel assignment has some drawbacks and limitations. Firstly, the number of channels operated by a MR is constrained by the number of the MR radios. If two neighboring MRs do not have a common channel, the link between them will be lost. Removal of some essential links in the network will lead to the network partition problem. So, the channel assignment has to be done very carefully to maintain the connectivity. Furthermore, the efficiency of the static
channel assignment may be low since it is not able to adjust the assignment to the network changes in time. In addition, the re-assignment of channels is not an easy task in a large scale WMN. High management message overhead is introduced with an increase in the network size.

On the other hand, some approaches such as [31] [41] enable radios to be switched onto different channels. The negotiation of hopping schedule [31] and the agreement of the channels [41] require radios to meet at the default channel at predefined time intervals. It has higher efficiency as compared to a static approach and is able to utilize multiple channels by a single radio. The challenging issue is the synchronization of switching operation. Two MRs intending to communication need to switch radios on the same channel at the same time. Less accurate synchronization negatively impacts the throughput. High cost of precision equipment such as GPS, is not yet feasible for all MRs. In addition, the network overhead will be high when the channel switching is frequent.

4.3 Problem Formulation and Modeling

A WMN backbone consists of MRs which are interconnected with wireless links and provides the network access service for MCs (shown in Figure 4.1). Due to different transmission distance and data rate requirements for the connections in the backbone network and the accesses of MCs, we assume an MR to use separate radios and channels for each kind of communications and they do not mutually affect each other. Our focus is the transmission between interconnected MRs in the backbone network.

To transmit packets between a pair of MRs, e.g., $u$ and $v$, in a multi-channel
environment, several conditions need to be satisfied: (i) The distance between any two direct communicating MRs is less than the critical transmission distance $r_T$. If two MRs are within transmission distance of each other, the link between them is denoted by $uv$; (ii) To transmit packets on link $uv$, each of the MRs needs to operate a radio on a common channel, e.g., channel $k$; (iii) To successfully receive a packet, the signal to interference ratio (SIR) at the receiving radio must be above a threshold value. Due to exponential attenuation of signals, other radios operating on the same channel must have enough distance away from the receiving radio and such critical interference distance is denoted by $r_I$. Inappropriately using multiple radios may cause the network partition and adversely hurt the network performance. To effectively operate multiple radios on the right channels, there are three major issues to be considered for multi-radio MRs: channel assignment, channel switching synchronization, and the switching delay.
Channel Assignment for Multi-radio MRs

There are two kinds of arrival traffics: one is from the MCs and the other is from other MRs in the backbone network. Part of the traffic from other MRs is transmitted to the MCs. The rest is buffered in the queues with the traffic from the MCs, which will be forwarded to the next hop MRs. Because the number of simultaneous transmission within an interference region over a channel is limited, channel assignment is to operate radios to increase the number of simultaneous transmissions. Meanwhile, the assignment must avoid the network partitioning problem and should adjust the assignment to the changes in the network traffic.

Channel Switching Synchronization

Switching synchronization not only needs two communicating MRs to switch the radios to a channel simultaneously, but also requires low overhead and high accuracy. There are two kinds of switching synchronization methods: out-of-band and in-band. Out-of-band method utilizes external information for each MR to acknowledge the time, such as Global Positioning System (GPS). GPS technology provides accurate time information; however it is still expensive to equip every MR in the network with a GPS receiver and it is hard to guarantee that all MRs intending to join in the network is also equipped with a GPS receiver. In-band method seeks to synchronize the switching by beacon signals exchanged among the MRs and is cost effective. As compared to the method based on the GPS technology, beacon signal suffers from multiple factors (such as propagation delay of beacon signals) that could decrease the accuracy.
Switching Delay

It is negligible in a static channel assignment because it is a very small fraction as compared to the long duration that a radio remains on using a switched channel. However, the switching delay becomes comparable in a dynamic radio scheme due to frequent channel switching.

4.4 Radio Management

Dynamic channel assignment requires the radios used by the communicating MRs to be switched to the same channel at the same time so as to satisfy the necessary conditions of data transfer in a multi-channel environment. To address the synchronization problem, we divide the radios into two sets: the receiving radios (RRs) and the forwarding radios (FRs). RR is used to receive packets from other MRs and FR is used to transmit packets. While MR has packets to send to another MR, it switches one of the FRs to a channel used by the destination MR. For example, MRs $A$, $B$, and $C$ each has two radios, shown in Figure 4.2. One of the radios of each MR is set to RR and the other one is set to FR. Suppose the RR of $A$, $B$, and $C$ is assigned on channels 1, 2, and 3, respectively. If $A$ has packets to send to $B$, it switches its FR to channel 2, because $B$ already has a RR on channel 2, they can communicate immediately without any synchronization process. If $A$ has another packet to send to $C$, it can just switch the FR to channel 3 and transits the packets to the RR of $C$, which is already on channel 3. $B$ and $C$ communicate to other MRs in a similar way. Each RR can broadcast beacon signals to notify the existence on a channel. Hence, FRs of MRs can acknowledge the channel usage information of the
neighboring MRs.

One of the benefits in dividing the radios into RRs and FRs is that it can adaptively adjust the radios for the incoming and outgoing traffics. The main concern of other existing approaches is to minimize the interference, which may force the radio not to fully utilize. For example, there are six MRs from A to F. The traffics from A, B, and C are relayed by MR D to E or F. Suppose MRs A, B, and C have the same traffic demands, i.e., $\lambda_{AD} = \lambda_{BD} = \lambda_{CD} = \lambda_p$, and the FR service rate of MR D is $\mu_{DE} = \lambda_p$. Then, MR D will be highly congested because the packet forwarding rate is much lower than the incoming traffic rate and the maximum throughput is only $\lambda_p$. We can balance the radio assignment for incoming and outgoing traffic by switching a radio from the RR set to the FR set. Hence, the throughput can reach close to $2\lambda_p$.

WMN is different from an ad hoc network that each MR gets the traffic generated by the MCs. The incoming traffic of a MR $u$ is denoted by $\lambda_{iu}$. The incoming traffic contains the traffic from the MCs served by MR $u$, denoted by $\lambda_{iu0}$. The outgoing traffic of MR $u$ is denoted by $\lambda_{gu}$, which also includes the traffic to the MCs, denoted by $\lambda_{gu0}$. The assignment of radios should balance the traffic so as to maximize MR throughput. Therefore, when the traffic is not heavy, the ratio of the number of RR and the number of FR is:

$$\frac{|RR|}{|FR|} \approx \frac{\lambda_{iu} - \lambda_{iu0}}{\lambda_{gu} - \lambda_{gu0}}$$

(4.1)

The adjustment mechanism also supports the traffic control by avoiding the congestion at the MR. When the number of buffered packets is above a pre-determined congestion threshold $\zeta$, the MR switches one of the radios from the RR set to the FR set. With such a switching, the maximum allowed incoming traffic is reduced and
the packet forwarding rate is increased. Because of the reduced number of RRs, the neighboring MRs forwarding packets to the MR has to decrease their sending rate or find an alternative path. It is possible to balance the traffic in the whole network. When the number of buffered packets is below another pre-determined threshold $\eta$, which implies that the MR can accept more packets, MR switches one of the FRs to the RRs.

The successful reception of a bit is determined by the SIR at the receiver. The success of a link transmission is mainly determined by the reception of data packets. In our scheme, it is important to measure the SIR at the RR. Therefore, for a duration, MR can directly measure the interference level from the physical layer and decide which channel to select. If MR has more than one RRs, say $m_r$, it selects sequentially one among $m_r$ RRs that has the minimum interfering channel. MR does not operate two RRs on the same channel, because two RRs cannot work simultaneously at the same channel. $m_r$ RRs can operate on $m_r$ different channels. Each RR use beacon signals to inform its existence on the channel. So, other MRs can acknowledge it and switches FR to the channel for transmission. The traffic volume at the RR is indicated by the frequency of the beacon signals. MR may have several RRs and the incoming traffic at different RRs can be hopefully balanced. MR
uses different frequency for the beacon signals: higher frequency means lower traffic and lower frequency means higher traffic. Therefore, other MRs want to connect to a given MR can observe the frequency of its beacon signals and decide which RR of the MR it should connect to and switch a FR to the channel. Therefore, the proposed radio management mechanism avoids switching synchronization problem and supports traffic balancing and congestion avoidance.

4.5 Multi-channel Queue Management

Since FRs dynamically operate on different channels, the sequence of packets in the queue affects the performance of the FR. The worst case is when any two consecutive buffered packets are transmitted using different channel. The radio has to be switched for every packet transmission, which causes high switching overhead.

Let us first observe the queue management for a single FR, as shown in Figure 4.3(a). MR $u$ has received packets from four different channels to forward, which are denoted by the number 1, 2, 3, and 4 respectively. Suppose MR $u$ has a single FR that is dynamically switched to different channel for sending the packets. MR $u$ has to switch for every two consecutive packets, forwarding using different channels. However, if the buffer can adjust the sequence of packets and group those packets forwarded together over the same channel, the number of switching times can be reduced.

Given the total number of channels $c$, the packet arrival process is independent identical distributed (i.i.d.), and the packet arrival is a Poisson arrival process. The arrival rate of the packets over channel $i$ is denoted by $\lambda_i$ and the summation of traffic rate over all channels is $\lambda = \sum_{i=1}^{c} \lambda_i$. The channel switching occurs when
two consecutive packets are on different channels. The packet on the channel $k$ is denoted by $Pkt_k$. Conditional probability that the radio switched from channel $i$ to channel $j$ is denoted by $p_{ij}$. So, the probability that no packet on other channels exists between two consecutive $Pkt_i$ packets whose interval is $t$:

$$p_{ii|t} = e^{-(\lambda_i-\lambda)t}.$$  \hfill (4.2)

The probability distribution function (pdf) of the interval between two consecutive $Pkt_i$ packets is:

$$f_i(t) = \lambda_i e^{-\lambda_i t}.$$  \hfill (4.3)

So, the probability that no packet exists on other channels between two consecutive $Pkt_i$s is:

$$p_{ii} = \int_0^t p_{ii|t} f_i(t) dt = \frac{\lambda_i}{\lambda}.$$  \hfill (4.4)

The probability that one FR is switched from channel $i$ to other channels is:

$$\sum_{j \neq i} p_{ij} = 1 - p_{ii} = \frac{\lambda - \lambda_i}{\lambda}.$$  \hfill (4.5)

The channel switching is a stochastic process. Every channel is assumed to be in a state of the stochastic process. The steady state probability that a radio remains on a channel can be solved by the balance functions and the steady probability that the radio remaining on channel $i$, denoted by $\pi_i$, is given by:

$$\pi_i = \frac{\lambda_i}{\lambda}.$$  \hfill (4.6)

So, the channel switching rate is expressed by:

$$F_{sw} = \lambda \sum_{i=1}^c \pi_i (1 - p_{ii}) = \lambda - \sum_{i=1}^c \frac{\lambda_i^2}{\lambda}.$$  \hfill (4.7)
Figure 4.3: Queue Management

Our proposed queue management is to divide a queue into $c$ queues. Each queue buffers the packets forwarding on the same channel, as shown in Figure 4.3(b).

The time that a radio remains on a channel is proportion to the steady state probability obtained above. Once FR is switched to the channel $k$, it remains on the channel with a randomly selected time which follows the exponential distribution with parameter $\mu_k = \frac{1}{T_{\text{mean}}}$, where $T$ is the mean period of a radio switched through overall channels. During the time interval that FR remains on channel $k$, if there is no other packet arriving at the queue for channel $k$, the FR is switched to the next queue having packets.

Now, let us investigate the queue management for multiple FRs. Conventional method uses multiple queues separately. Each queue is associated to a radio and the packets are distributed to one of the queues as shown in Figure 4.3(c). With dynamic channel assignment, the radio in such a scheme still needs to be switched among all the channels. Our management still divide a queue into $c$ queues, and each one of FRs charges for one or several queues. As the example in Figure 4.3(d) shows, each of the two FRs takes care of two queues. To fully utilize the FRs, it
needs to balance the outgoing traffic among the FRs. We use a greedy algorithm to balance the outgoing traffic. First, taking all queues as a candidate set and sorting it by decreasing order of the traffic arrival rate of the queue. Then, assigning the first one, i.e., the queue having the heaviest traffic, to one of the FRs and remove the queue from the candidate set. The FR assigned to the queue takes the traffic arrival rate as the weight. The assignment of the second queue is similar. Taking the queue having the heaviest traffic in the candidate set and assign it to the FR with the lightest weight. The FR assigned queue accumulates the arrival rate for the weight. The process repeats until the candidate set is empty.

When FR needs to send packets, it is switched to the channel which the receiver operates on. If the destination MR has more than one RRs, it selects the one with the highest frequency beacon signals. So, the traffic can be approximately balanced at the receiver.
Figure 4.5: Switching Rate

Figure 4.6: Delay at MR

80
4.6 Performance Evaluation

We determine the probability of connectivity and the channel switching overhead using simulation. Figure 4.4 compares the probability of connectivity. The total number of MRs is varied from 50 to 1000. MRs are deployed in a unit disk with the Poisson point process [42]. The critical transmission distance $r_T$ is set to $\sqrt{\frac{a \log n}{\pi n}}$, detailed in [42] [16], where $a$ is set to 1.2, 1.5, and 2.0 in the simulation. The probability of connectivity means the probability that any MR has at least one path to any other MR in the network. Suppose the static channel assignment used in the simulation selects different channels for the connected MRs, so the number of neighbors that MR can connect to is no more than the radio number. The probability of connectivity for the static channel assignment with different number of radios: $m = 4$ and $m = 6$ are compared with the dynamic channel switching with $m = 2$. From Figure 4.4, dynamic channel switching has better connectivity and is able to fully utilize the multi-channel resource. The connectivity is improved by increasing the number of radios for the static channel assignment. The transmission distance plays an important role for the connectivity. When the parameter $a$ increases, the connectivity of static channel assignment is close to that of the dynamic channel switching, which means that they have a similar channel utility.

Figure 4.5 shows the switching rate with different arrival traffic pattern. The arrival traffics are over three channels, which are denoted by $\lambda_1$, $\lambda_2$, and $\lambda_3$. The arrival rates are the same for traffic over three channels, which is denoted by traffic pattern $Tf_1$. In this pattern, each arrival traffic is increased by 0.5 in each step. From Figure 4.5, the switching rate for one queue state confirms that the analysis provided by Eq. (4.7), and the switching rate of proposed queue management is
below that of a single queue situation. When the traffic is light, the switching rates
of a single queue and multi-queue are similar. When the traffic is heavy, the benefit
of our queue management scheme is very obvious. This is because that the proposed
queue management saves switching time by organizing packets together over the
same channel. After a radio processes a packet in the queue, if there is no new
arriving packet over the same channel, it is switched to the next channel. So, while
the packet arrival rate is low, the switching times of the radio is close to that in the
single queue situation. When the traffic is heavy, the radio can continue to serve for
several packets over the same channel for a time duration rather than to be switched
across the channels. So, the switching overhead is obviously reduced. In traffic
pattern $T_f_2$, two of the traffics are started from 1.0 and increased by 0.5 for every
step, the rest one keeps at 3.0. Traffic pattern $T_f_3$ only increases one of the traffic
arrival rate and keeps the other two traffics constant. Both $T_f_2$ and $T_f_3$ show the
same trend as $T_f_1$.

The delay at MR is shown in Figure 4.6. The proposed queue management has
less packet delay than that of the single queue situation. The delay can be reduced
further by minimizing the switching rate. In addition, the curves in Figure 4.6 shows
that the delay is reduced drastically when the traffic load is heavy.

4.7 Conclusion

This chapter proposes an effective radio partitioning and an efficient queue man-
agement scheme for a WMN. Our schemes operate parts of the radios dynamically
so as to adapt the traffic changes and ensure the network connectivity. It does not
need the switching synchronization usually used in a dynamic channel assignment
and reduces the switching overhead by adjusting the forwarding sequence of the packets. The novel radio management divides radios into RR set and FR set, which enhances the throughput and supports the traffic control. The total traffic arrival rate and departure rate can be controlled by shifting the radios between the two sets. It also facilities automatic traffic balancing in a network. The simulation results show that our proposed management can effectively decrease the switching rate and reduce the packet delay. The scheme is distributed and is able to change the channel assignment according to the SIR at MR location. It does not need to exchange much information and can be quickly self-configured and organized by scanning the usage of channels in the transmission region.
Chapter 5
Fractional Frequency Reuse in WMN Multi-Cell

5.1 Introduction

WMN system is able to deliver convenient network access service for MCs, by multi-cell coverage. To secure seamless connections, the coverage of cells is required to be continued, which inevitably leads to coverage overlap at cell edges. MCs located at these overlapped areas may have lower signal to signal to interference and noise ratio (SINR) because of weak signals from long distance serving MRs and serious interference from neighboring MRs. Frequency reuse [43] has been introduced in multi-cell networks to avoid ICI, where the frequency band is partitioned into multiple sub-bands. A cell uses one of the sub-bands and identical sub-band is only reused by the cells at certain distance away, defined as reuse distance. Thus, the interference from the same frequency sub-band can be mitigated by longer reuse distance. Frequency reuse factor represents the rate (cells/sub-band) of the identical frequency sub-band reused in networks. Higher reuse factor means
longer reuse distance, which has less ICI. However, it also leads to more division of frequency band and each sub-band has less bandwidth. Under such circumstance, spectrum may not be sufficiently utilized, because it does not consider non-uniform distribution attribute of ICI, which is aggravated with the distance to MR and most serious at cell edge.

In order to mitigate cell edge interference and enhance spectrum efficiency, fractional frequency reuse (FFR) starts to be fully utilized by the next generation systems, e.g. IEEE 802.16 [4] and the third generation partnership project long term evolution (3GPP LTE) advanced [44]. Different from constant frequency reuse in conventional multi-cell cellular systems, orthogonal frequency-division multiplexing (OFDM) technology [45] employed by the next generation systems is capable of delivering different frequency reuse factors on frequency partitions. Based on interference levels, a geometric cell can be divided into cell center and cell edge zones. Cell edge zone is the coverage area overlapped by adjacent cells, where serious interference may be present. The rest coverage area around MR is defined as cell center zone. According to the level of ICI in zones, it is desired to implement a higher frequency reuse factor for cell center zone and a lower frequency reuse factor for cell edge zone.

Due to presence of a number of MCs and variability of traffic loads in multiple cells and their zones, fixed or pre-determined FFR is not able to fully utilize radio resources. Conventional FFR, whose radio resource allocated on zones is fixed, cannot switch partitioned radio resource between zones for traffic load changes. Radio resource or transmit power in conventional FFR is not able to be adjusted for traffic changes as well. Spectrum efficiency will be lost because of the absence of any intra-cell and/or inter-cell frequency reuse adjustment. To overcome these deficiencies, a
Dynamic FFR is highly desirable for commercial multi-cell systems.

Dynamic FFR should adaptively allocate and/or adjust frequency reuse factor and bandwidth of frequency partitions in different cells and zones with low computation complexity. There are many challenging problems to perform these functions. Firstly, evaluation and comparison of FFR schemes need accurate models to characterize their throughput bounds in multi-cell networks, considering MS geometry location, ICI fluctuations in multi-cell networks, etc. Secondly, spectrum partitions used for different zones have different reuse distances. If the spectrum partitions used for the same type of zones in two neighboring cells exist overlap, serious ICI will be generated. In addition, jointly with the adjustment of transmit power, the problem becomes more complex. Optimal allocation in a multi-cell network is a NP-hard problem [46].

In this chapter, we establish an analytical framework for FFR schemes, which also provides a foundation in deriving our scheme. Throughput of randomly deployed MCs in multi-cell networks is estimated with taking effects of frequency band partition and resource allocation of FFR into account. In result, FFR scheme can be evaluated by its cell throughput which is obtained from subchannel capacity. Then, we propose an innovative differentiable FFR (D-FFR) scheme that enables each MR to have differentiable spectrum partition and resource allocations for cells and their zones. Our D-FFR provides a flexibility in ICI mitigation and elasticity in MAC layer QoS support.
5.2 FFR Model and Analysis

FFR can be implemented either on cells (with omni-antenna) or on cell sectors (with directional antenna). When FFR is implemented on cells, every cell is divided into cell center and edge zones. A fraction of frequency resource, i.e. spectrum partition, is allocated to cell zones with designed reuse distance. In contrast, when FFR is used for sectors, a fraction of frequency resource is used at edge zone of every sector with designated reuse distance, where different sectors of a same cell may use different fractions of the frequency band. To simplify the illustration of our proposed mechanism, we analyze the implementation of FFR on cells. However, our results are also equally applicable to FFR utilization on sectors.

5.2.1 System Model

We consider a WMN network to consist of $K$ MRs and $N$ MCs. The set of MRs in a multi-cell network is denoted by $\mathcal{K} = \{1, ..., K\}$. On a two dimensional area $\mathbb{R}^2$, the coverage of a single MR is a circle centered at the MR with radius $R_T$, where MC can decode frame preamble sent by MR, as shown in Figure 5.1. The inter-MR distance, also called inter-site distance, is denoted by $D_s$. Assume $K$ MRs are connected to each other by a wired backhaul. When multiple MRs constructs multi-cell coverage, the cell of a MR is a hexagonal area centered at the MR, where inradius $R_{HI}$ of the hexagonal is $\frac{D_s}{2}$, and circumradius, denoted by $R_{HC}$, is $\frac{D_s}{\sqrt{3}}$. MC located in the area overlapped by two MR coverages, is able to receive the broadcast messages and obtain identifications (MR-IDs) of both MRs. For example, $MC_1$ in Figure 5.1 can receive the messages from $MR_1$ and $MR_2$. However, MC is only allowed to register and exchange data with one of the MRs, which is called serving
(or anchor) MR and the other MR is called target MR. In general, the overlapped area is also used for MC to perform handover when it migrates access connection from a serving MR to a target MR. We assume MC in the network selects the MR with the minimum Euclidean distance as the serving MR, and executes handover procedure to connect to the target MR when the target MR becomes closer than the serving MR. Similarly, MC located in the area covered by three MRs is able to acknowledge three MR-IDs, like $MC_3$ in the figure.

For FFR implementation, we define cell edge and center zones as follows:

- **Cell edge zone** is such an area in a cell that MC in the cell can receive broadcast messages from more than two MRs.

- **Cell center zone** is such an area in a cell that MC in the cell only can receive the broadcast messages from the serving MR.

In our analysis, cell center zone can be approximated by the area within a circle centered at an MR with radius $R_c$, as the dotted circle in Figure 5.1, which approximates the shaded area.

Frequency resource $\mathcal{F}$ used for data transmission is divided into orthogonal modulated subcarriers. A group of subcarriers can be assigned to each MC for communication between MR and MC. Each MR allocates downlink subchannels in such a set $\mathcal{F} = \{1, ..., F\}$ for MC in the cell, where every subchannel $i \in \mathcal{F}$ is a group of subcarriers. There are two ways to group subcarriers into subchannels: contiguous and distributed permutation. A subchannel with contiguous permutation has continuous frequency bandwidth provided by contiguous subcarriers. Distributed subcarrier permutation uses a permutation formula to determine the subcarriers in a subchannel, which may be distributed on the frequency band. The formation of subchannels can utilize either contiguous or distributed permutations. While using
Figure 5.1: Multi-cell Coverage
distributed permutation, we assume each subchannel consists of a fixed number of subcarriers and the permutation formula is same for each MR.

To adaptively allocate the subchannels on different zones, it needs to know ahead of the belonging of the zones of MCs. MC in the network is able to record the MR-IDs obtained from its received messages, which forms a diversity set. Diversity sets of MCs are delivered to the serving MR for acknowledging the approximate location of the MCs. Take the MCs in Figure 5.1 as example. \(MC_1\) can receive the broadcast messages from both \(MR_1\) and \(MR_2\), so its diversity set is \(H_{MC_1} = \{MR_1, MR_2\}\). Similarly, the diversity sets of \(MC_2\) and \(MC_3\) are \(H_{MC_2} = \{MR_1\}\) and \(H_{MC_3} = \{MR_1, MR_2, MR_3\}\), respectively. MCs transmit the diversity sets to their serving MR, so the serving MR can deduce the zone belongings of MCs.

### 5.2.2 Cell Throughput

The correct detection of signals at MC depends on its SINR. While two MRs use the same subchannel for communication with different nearby MCs, the interference could cause both MCs to lose the message. To understand the interference problem, we need to measure the SINR at MC. The SINR for MC located at point \(s\) in the cell of MR \(u\) over subchannel \(i\) can be expressed by:

\[
\xi^{(u)}_i(s) = \frac{P^{(u)}_i L^{(u)}_i(s)}{\sum_{v \in U_i - u} P^{(v)}_i L^{(v)}_i(s) + N_o},
\]

(5.1)

where \(U_i\) denotes the set of MRs using subchannel \(i\), \(P^{(v)}_i\) denotes the transmit power of MR \(v\) over subchannel \(i\), \(L^{(a)}_i(s)\) denotes the path loss from MR \(u\) to the MC located at point \(s\), and \(N_o\) denotes the thermal noise power. With higher transmit power used by the serving MR, MC can have higher SINR. As Eq. (5.1) shows, \(\xi^{(u)}_i(s)\) can
be increased by increasing $P_i^{(u)}$. However, it also leads to higher interference on
the MCs located in neighboring cell $v$, decreasing $\xi_i^{(u)}$ due to increase at the denom-
inator. So, the optimization problem needs to consider network’s comprehensive
conditions.

MC may stay in a cell for certain time period or cross the cell. We assume MC
stays at a point with certain probability and denote the probability density function
(PDF) of MC stay at point $s$ of cell $S$ ($s \in S$) by $\varphi(s)$, which satisfies $\int_S \varphi(s)ds = 1$. Then, we denote the probability that MC located at point $s$ is allocated subchannel $i$ by $\varphi_i(s)$, which satisfies:

$$\sum_{i \in \mathcal{F}} \int_S \varphi_i(s)ds = 1. \quad (5.2)$$

MCs served by MR $u$ is i.i.d. in a cell $u$. The probability that subchannel $i$ will
be allocated in cell $u$ can be computed by:

$$\phi_i^{(u)} = \int_S \varphi_i^{(u)}(s)ds, \quad (5.3)$$

where $\varphi_i^{(u)}$ denotes the probability that MR $u$ allocates subchannel $i$ in its cell.

Then, we take geometric attribute of MC into account to develop the equation
to compute the average achievable cell throughput of cell $u$, denoted by $C^{(u)}$, which
can be estimated by:

$$C^{(u)} = \sum_{i \in \mathcal{F}^{(u)}} \int_S \log(1 + \xi_i^{(u)}(s))\varphi_i^{(u)}(s)ds, \quad (5.4)$$

where $\mathcal{F}^{(u)}$ denotes the set of subchannels used by cell $u$.

Eq. (5.4) estimates cell throughput of reuse-1 scheme, where all subchannels are
used by every cell with same transmit power. To model and compute cell throughput
in FFR, Eq. (5.4) needs to be further extended with considering following factors.
The set of subchannels used by cells will vary with cells. For MR in set $\mathcal{K}$, the usage of a subchannel is denoted by the probability computed by Eq. (5.3). With different schemes, the usage of subchannels is different, which leads to different cell throughput. The power level of an identical channel at different cells may be different due to power allocation of cells. The calculation of achievable cell throughput should account this effect as well, together with the PDF of subchannel usage. Thus, Eq. (5.4) is rewritten as:

$$C^{(u)} = \sum_{i \in \mathcal{F}(u)} \int_{S} \zeta^{(u)}_{i,\mathcal{K},\mathcal{P}}(s) \varphi^{(u)}_{i}(s) ds,$$

where

$$\zeta^{(u)}_{i,\mathcal{K},\mathcal{P}}(s) = \log(1 + \frac{P_{L}^{(u)}(s)}{\sum_{v \in \mathcal{K} - u} P_{L}^{(v)}(s) + N_{o}}).$$

In Eq. (5.6), $\mathcal{P}_{i}$ denotes the power allocation vector for MR set $\mathcal{K} = \{1, ..., K\}$ on subchannel $i$ and $P_{L}^{(u)}(s) = P_{L}^{(u)} s_i L_i^{(u)}(s)$.

### 5.2.3 FFR Scheme

As FFR divides cell coverage area into two zones: edge zone $S_e$ and center zone $S_c$, the group of subchannels allocated to edge zone is called edge sub-band and the group of subchannels allocated to center zone is called center sub-band. Edge sub-band has higher frequency reuse factor, e.g., three, and center sub-band has frequency reuse factor one. In other words, the sub-bands for the adjacent edge zones are non-overlapping and the sub-bands for the center zones of adjacent cells utilize overlapped resource. To reduce the interference among center zones which have frequency reuse factor of one, the transmit power used for center sub-band is usually
lower than that used for edge sub-band. According to the method of frequency reuse in center zone, FFR can be divided into two major categories: hard FFR and soft FFR. Hard FFR has dedicated subchannels for center zone, which are not allocated to any cell edge zones. In soft FFR, the center zone of a cell reuses the frequency resource used by its adjacent edge zones. To mitigate any possible interference, its transmit power is lower than that of dedicated center frequency resource. Soft FFR can be further divided into two types: soft FFR A and soft FFR B, depending on if it has dedicated center frequency resource.

Figure 5.2(a) shows a **hard FFR scheme** with frequency reuse factor three at edge
zone and frequency reuse factor one at center zone. In hard FFR, center zones of cells use the same center sub-band. The MCs in the center zones of adjacent cells share the same subchannels, while the MCs in the edge zones of adjacent cells are allocated different subchannels. The MCs in the edge zone do not have the ICI from adjacent cells and can have higher SINR than that in reuse-1 scheme. Not all the frequency band is used by a cell in hard FFR. The empty bandwidth in Figure 5.2(a) is the frequency sub-band not used by a cell in order to satisfy edge subchannel allocation constraints. The vertical dotted lines in Figure 5.2 indicate the separation of the subchannels. Obviously, the sub-bands of adjacent cell edges are non-overlapping. The set of subchannels allocated for cell center zones is denoted by $\mathcal{F}_c$. The subchannels allocated for three different edge zones are denoted by $\mathcal{F}_e$, $\mathcal{F}_e$, and $\mathcal{F}_e$, which are non-overlapping and the total amount of bandwidth of three sub-bands complements the center bandwidth. The vertical axis indicates the transmit power used for the subchannels. The subchannels allocated for the edge zones use power $P_e$, which is higher than power $P_c$ used by the subchannels allocated for the center zones. From Eq. (5.5), the center zone throughput at cell $u$ can be estimated by:

$$C^{(u)}_{He} = \sum_{i \in \mathcal{F}_c(u)} \int_{S_c} \zeta_{i,\mathcal{K}(\mathcal{F}_e(u)),P_i(c)}^{(u)}(s) \varphi_i^{(u)}(s) ds,$$

(5.7)

where $\mathcal{K}(\mathcal{F}_e)$ means all MRs use the same transmit power $P_c$ on subchannel $i$. The throughput of the edge zone is computed by:

$$C^{(u)}_{He} = \sum_{i \in \mathcal{F}_e(u)} \int_{S_e} \zeta_{i,\mathcal{K}(\mathcal{F}_e(u)),P_i(e)}^{(u)}(s) \varphi_i^{(u)}(s) ds,$$

(5.8)

where $\mathcal{K}(\mathcal{F}_e(u)) := \bigcup_{i \in \mathcal{F}_e(u)} \mathcal{K}(i)$, and $\mathcal{P}_i(e)$ means all MRs in set $\mathcal{K}(\mathcal{F}_e(u))$ use transmit power $P_e$ on subchannel $i$. 94
The cell throughput of hard FFR scheme is the summation of center and edge zone throughput. The cell throughput of cell $u$ can be expressed by:

$$C_{H}^{(u)} = C_{Hc}^{(u)} + C_{He}^{(u)}. \quad (5.9)$$

Figure 5.2(b) shows a soft FFR A scheme. Different from hard FFR, the empty bandwidth is reused in soft FFR A with lower power $P_{e2}$, which can provide higher center bandwidth resource usage. However, it may also cause more interference to the MCs in the center zone and the MCs in the edge zone of the adjacent cells. The cell throughput of center zone have two portions, where the first one is same to that in hard FFR and computed by:

$$C_{SAc1}^{(u)} = C_{Hc}^{(u)} = \sum_{i \in \mathcal{F}_c} \int_{S_c} \zeta_{i,K,P_i(c)}^{(u)}(s)\varphi_i^{(u)}(s)ds. \quad (5.10)$$

The second portion, gaining from the reuse of subchannels used by neighboring edge zones, is calculated by:

$$C_{SAc2}^{(u)} = \sum_{i \in \mathcal{F}_e - \mathcal{F}_c - \mathcal{F}_a} \int_{S_e} \zeta_{i,K,P_i}^{(u)}(s)\varphi_i^{(u)}(s)ds, \quad (5.11)$$

where the element in $\mathbb{P}_i$ is equal to either $P_e$ or $P_{e2}$.

The total center zone throughput is the summation of two portions obtained above, which is expressed by:

$$C_{SAc}^{(u)} = C_{SAc1}^{(u)} + C_{SAc2}^{(u)} = C_{Hc}^{(u)} + C_{SAc2}^{(u)}, \quad (5.12)$$

which is higher than that of hard FFR.

In soft FFR A, the throughput of edge zone is obtained by:

$$C_{SAe}^{(u)} = \sum_{i \in \mathcal{F}_e} \int_{S_e} \zeta_{i,K,P_i}^{(u)}(s)\varphi_i^{(u)}(s)ds. \quad (5.13)$$
Figure 5.2(c) shows a soft FFR B scheme. In soft FFR B, there is no dedicated center sub-band and the sub-channels for center zone are overlapped with the adjacent cell edge subchannels. The throughput of center zone is computed by:

\[
C^{(u)}_{SBc} = \sum_{i \in \mathcal{F} - \mathcal{F}_e^{(u)}} \int_{S_c} \check{c}_{i,K,i}^{(u)}(s) \psi_i^{(u)}(s) ds. \tag{5.14}
\]

The throughput of edge zone is calculated by:

\[
C^{(u)}_{SBe} = \sum_{i \in \mathcal{F}_e^{(u)}} \int_{S_e} \check{c}_{i,K,i}^{(u)}(s) \psi_i^{(u)}(s) ds. \tag{5.15}
\]

Comparing Eq. (5.13) with Eq. (5.15), soft FFR A and soft FFR B would have the same edge zone throughput if they have the same number of subchannels. However, from Figure 5.2, it is clear that the sub-band of soft FFR B is wider than that of soft FFR A, due to the usage of sub-band \( K_c \) in soft FFR A. Therefore, \( C^{(u)}_{SBc} > C^{(u)}_{SAe} \). Similarly, the computation for center zone throughput of soft FFR B is same for soft FFR A. Soft FFR B is a special case of soft FFR A without dedicated center zone subchannels.

From the cell throughput obtained above, we can find the relation of different schemes. For center zone throughput, we have:

\[
C^{(u)}_{SAc} = C^{(u)}_{SBc}(\mathcal{F}_e^{(u)}) + C^{(u)}_{He}, \tag{5.16}
\]

where \( C^{(u)}_{SBc}(\mathcal{F}_e^{(u)}) \) is the soft FFR B center zone throughput by limiting the overall subchannels to \( \mathcal{F}_e^{(u)} \).
In the computation of edge zone throughput, we have:

\[
\zeta_{i,K,P_i}^{(u)}(s) = \log(1 + \frac{P L_i^{(u)}(s)}{\sum_{v \in K - u} P L_i^{(v)}(s) \phi_i^{(v)} + N_o})
\]

\[
= \log(1 + \frac{P_e L_i^{(u)}(s)}{\sum_{v \in K \setminus \mathcal{F}_i^{(u)} - u} P_e L_i^{(v)}(s) \phi_i^{(v)} + \sum_{w \in K \setminus K(\mathcal{F}_i^{(u)}) - u} P_c L_i^{(w)}(s) \phi_i^{(w)} + N_o})
\]

\[
\leq \log(1 + \frac{P_e L_i^{(u)}(s)}{\sum_{v \in K \setminus \mathcal{F}_i^{(u)} - u} P_e L_i^{(v)}(s) \phi_i^{(v)} + N_o})
\]

\[
\leq \zeta_{i,K(\mathcal{F}_i^{(u)}) \setminus P_i}^{(u)}(s).
\]

Since the edge zone throughput of different types of FFR is expressed by:

\[
C_e^{(u)} = \begin{cases} 
\sum_{i \in \mathcal{F}_i^{(u)}} \int_{S_e} \zeta_{i,K(\mathcal{F}_i^{(u)}) \setminus P_i}^{(u)}(s) \varphi_i^{(u)}(s) ds, & \text{Hard FFR} \\
\sum_{i \in \mathcal{F}_i^{(u)}} \int_{S_e} \zeta_{i,K,P_i}^{(u)}(s) \varphi_i^{(u)}(s) ds, & \text{Soft FFR A.}
\end{cases}
\]

(5.18)

We can deduce from Eqs. (5.17) and (5.18) that hard FFR has equal or higher edge throughput than the soft FFR A, \(C_{e,H_e}^{(u)} \geq C_{e,S_Ae}^{(u)}\).

From the analysis above, we know that the amount of radio resource allocated to cells and zones is fixed in hard FFR or soft FFR scheme, which results in fixed designed cell throughput. A practical system requires differentiable cell throughput satisfying design purpose. On the other hand, cell throughput has fixed division in center zone throughput and edge zone throughput in above schemes. Hard FFR and soft FFR have shown different advantages in edge and center zone throughput. It is also desirable to have a mechanism to flexibly adjust cell throughout division in edge and center zones.
5.3 D-FFR Modeling

Resource allocation in FFR needs to satisfy certain constraints, otherwise ICI may be aggravated. Figure 5.3 shows an example of resource allocation in FFR. Suppose total 12 subchannels, which are indexed by hexadecimal numbers (from ‘1’ to ‘c’), are allocated for an OFDMA network constructed by 9 cells. The frequency reuse factor for the edge zones is 3, which means every 3 cells reuses a complete edge frequency sub-band. The first letter of cell notation means the same group of cells use the edge frequency resource. For example, for three cells in Figure 5.3 beginning with letter ‘A’, their edge radio resources form a complete edge frequency sub-band. Such a group of cells forms a cluster.

With fixed FFR shown in Figure 5.3, there is no adjacent cell edge zones using a same channel, which ensures low ICI. The problem may happen when it adjusts the subchannel allocation among different zones. Obviously, individual cell resource allocation will cause serious ICI due to using same subchannels in adjacent cell edge zones. Some work argues to adjust allocation in a cluster, which secures none-overlap edge resource allocation for cluster. Figure 5.3(b) shows the problems that cluster A adjusts its cluster resource allocation without coordinating with other clusters. While cluster A extends subchannel 4 for the edge sub-band of cell A₁, it reduces the edge zone sub-band of cell A₂ by removing subchannel 4, so as to avoid serious ICI in cluster A. However, subchannel 4 will seriously interfere with the usage of this subchannel at the edge zones of adjacent cells C₂ and D₂. The resource adjustment in cell center zone will also cause similar problem. If cluster A extends subchannel 9 for the center sub-band of one of cells in the cluster, other cells in cluster A has to switch subchannel 9 to their center zone to avoid ICI in cluster.
However, it still poses seriously interference to other clusters due to the same sub-channel used by adjacent cells $B_3$ and $C_3$. The interfered subchannels are denoted by the shaded boxes in the figure.

Figure 5.3(c) shows an allocation result solved by our proposed D-FFR scheme. Subchannels 1, 2, 3, and 4 are allocated for cell $A_1$ edge zone. The subchannels of other cells are adjusted to mitigate interference as ensure the number of needed subchannels at each cell as well. Cells $A_2$ and $A_3$ are able to utilize subchannels 9 and $a$, respectively, for the center zone throughput and adjust the downlink transmit power according to the subchannel usage of the adjacent cell edge zones. Such frequency resource rearrangement and downlink transmit power adjustment enable each cell to have different amount of resources for each partition according to its requirement and employ the residual capacity of the subchannels. Comparing with allocation results in Figure 5.3(b), the amount of resource allocated in each cell and its zone is similar, but our method has much lower ICI and higher spectrum efficiency.

In the rest of this section, we model the allocation problem and show the constraints in mitigating ICI. Sequentially, our heuristic algorithm satisfying the constraints in the model with proposed objective is described in next section, which is able to perform the allocation as the example in Figure 5.3(c) shows. To simplify the analysis, we illustrate our method with the allocation of downlink resource and transmit power allocation. However, the principles in our method is also applicable to uplink.

**Objective:** The allocation objective in D-FFR scheme is to maximize the throughput with different edge/center throughput divisions in cells. The problem can be modeled as follows. The network is modeled by $G(V, E)$, where $u \in V$ denotes a
cell in the network and $uv \in E$ denotes the inter-cell interference relationship. If downlink signals from MR $u$ interfere with downlink signals from MR $v$, an edge $uv$ is added in set $E$ to show such a conflict. The downlink traffic load at a cell $u$ is denoted by $D^{(u)}(t)$. Since we consider the allocation at a time slot $t$, we ignore $t$ in the notation and use notation $D^{(u)}$ for simplification. $D^{(u)}$ can be further divided into edge zone traffic load, $D^{(u)}_e$, and center zone traffic load, $D^{(u)}_c$. The subchannel allocation at MR $u$ on subchannel $i$ is denoted by a binary variable $A^{(u)}_i$. The selection of transmit power is denoted by $P^{(u)}_i$, which takes the value from one of three values: $P_e$, $P_{c1}$, or $P_{c2}$, where $P_e > P_{c1} > P_{c2}$. Resource allocation and transmit power determination should satisfy the following constraints:

**Edge zone allocation constraint**: It models the ICI between adjacent cell edge zones and is used for ICI mitigation. By this constraint, adjacent cell edge zones are
prevented from using the same subchannel. The constraint can be expressed by:
\[
A_i^{(u)} + A_i^{(v)} \leq 1, \quad \forall i \in \mathcal{F}, \forall uv \in E. \quad (5.19)
\]

**Transmit power constraint:** It models and constrains the transmit power used by adjacent cell on the same subchannel. To prevent the subchannel from being excessively reused in adjacent cells, the power levels of transmitter at certain distance need to be below a defined threshold. Considering both hard FFR and soft FFR cases, the maximum allocable transmit power for adjacent cells on a same channel should be no more than \(P_e + P_{c2}\). This constraint serves as the upper bound for the resource reuse in adjacent cells:
\[
P_i^{(u)} A_i^{(u)} + P_i^{(v)} (1 - A_i^{(v)}) \leq P_m, \forall i \in \mathcal{F}, \forall uv \in E, \quad (5.20)
\]
where \(P_m = P_e + P_{c2}\). If no soft reuse is allowed in a system, it makes \(P_m = P_e\).

**Edge zone resource reservation factor:** To control the allocation between the zones in a cell, we set two bounds for the division of the frequency. The first one is the edge zone throughput reservation factor \(B_L^{(u)}\), which is used to reserve radio resources for cell edge zone and serves as edge zone throughput lower bound. Without this factor, some edge zones may have very limited resources and have negative effect to the whole network performance. This factor is also necessary when the traffic load is light. It can allocate more resource for edge zone to have the designed throughput proportion. The allocation to support lower bound is given by:
\[
\sum_{i \in \mathcal{F}} A_i^{(u)} \geq B_L^{(u)}, \quad \forall u \in V. \quad (5.21)
\]

**Edge zone resource constraint factor:** \(B_U^{(u)}\) is used to control the maximum amount of subchannels allocated for edge zone, which serves as edge zone throughput upper bound. Different from \(B_L^{(u)}\) that attempts allocate more resources to the
edge zone when the traffic load is light, $B_U^{(u)}$ provides more resource to the center zone while the traffic load is heavy. On the other hand, this factor is able to reserve resources for center zone to avoid negative effect to adjacent cell edge zones. As edge zone allocation and transmit power constraints show, a big amount usage of subchannels by a cell edge zone may result scarcity of radio resource in adjacent cell edge zones. The constraint can be expressed by:

$$\sum_{i \in \mathcal{F}} A_i^{(u)} \leq B_U^{(u)}, \quad \forall u \in V. \quad (5.22)$$

While executing the algorithm, the actual division of resources for two zones will exist between two bounds. Figure 5.4 shows the adaptive subchannel allocation between the two bounds, where vertical axis ($P$) denotes power level and horizontal axis ($f$) represents frequency resource. By setting $B_L^{(u)}$, MR $u$ is able to reserve edge zone resources. Also, by setting $B_U^{(u)}$, the maximum allocation for the edge zone is constrained. On the other hand, by controlling the maximum amount of resource used for edge zone, the resource is also implicitly reserved for center zone. In addition, the configuration of $B_L^{(u)}$ and $B_U^{(u)}$ can be potentially considered to support QoS traffic classes defined in MAC layer. While reserving relative larger edge resource for a small number of MCs in edge zone, each MC would have wider bandwidth, shorter time delay, etc.

### 5.4 D-FFR Scheme and Analysis

Our scheme involves three major steps that satisfy both requirements and constraints discussed in last section. We firstly construct a node weighted constraint graph, which reflects the constraint relationship among the cells in the network.
Then, we iteratively search the maximal independent node sets for subchannel allocation. After that, the power levels for subchannels at different zones are determined according to subchannel usage conditions and ICI of adjacent cells.

5.4.1 Node Weighted Constraint Graph

As discussed above, multi-cell network is denoted by a constraint graph $G(V, E)$, where node $u \in V$ represents the MR and edge $uv \in E$ connecting pairs of nodes represents the subchannel allocation constraints between the MRs. The construction of graph depends on the network topology, i.e., the locations of the MRs for interference estimate. The nodes in the graph are weighted according to the traffic loads and the traffic loads in the edge zones are of special interest. For a node $u \in V$, the node weight, denoted by $w(u)$, is equivalent to the edge zone traffic load.

5.4.2 Maximal Independent Node Set

Subchannels are firstly allocated to the cell edge zones in the network. It iteratively searches the maximal independent node set and allocates subchannels for the nodes until all subchannels have been examined or the edge resource requirements are satisfied. An independent node set is a subset such that no two nodes in the
subset are adjacent in the graph. A maximal independent node set is a set that is not a subset of any other independent node set. We iteratively search such a maximal independent node set in the weighted constraint graph to allocate resources for the nodes so that ICI can be mitigated.

The purpose of maximal independent node set search is not just to find the maximal number of MRs that can perform transmission simultaneously, but also to find the group of MRs which mostly needs to schedule downlink transmissions. The weight of node in weighted constraint graph indicates such a requirement. This is important while some MRs are heavily loaded. They can have more opportunity to obtain subchannel for their transmissions.

The algorithm searches on the node weighed constraint graph constructed by the MRs indicating the resource requirements. For MR without requirements or having been assigned required subchannels, the corresponding node and corresponding edge(s) could be removed from the graph. Graph \( G(V, E) \) is constructed for the MRs requiring edge zone resource allocation. The algorithm searches the node that can be included in independent node set \( I \) where no node in the set is adjacent to each other. To illustrate our proposed algorithm, we give the following definitions.

Given graph \( G(V, E) \) and set \( I \) formed by a subset of node set \( V \), if a node \( w \) is not in set \( I \) but adjacent to a node \( v \) in set \( I \), we say \( w \) is adjacent to set \( I \).

Given graph \( G(V, E) \) and set \( I \) formed by a subset of node set \( V \), we denote the set of all nodes in \( V \) that are adjacent to set \( I \), by \( J(I) \), which is expressed by:

\[
J(I) = \{u | uv \in E, u \notin I, v \in I\}. \tag{5.23}
\]

It is obvious that \( I \cap J(I) \) is empty. When we have a set \( I \) and its adjacent node set \( J(I) \), we can easily determine if a node \( u \) is independent of set \( I \), which implies
that node $u$ is not adjacent to any nodes in set $I$. In the graph, we can find node set in which the node is adjacent to any node in set $I \cup J(I)$. This node set is denoted by $J(I \cup J(I))$, which means the set of nodes adjacent to set $I \cup J(I)$.

For a node $u$ not in $I \cup J(I)$, if it is adjacent to $I \cup J(I)$, the node which node $u$ is adjacent to must be in set $J(I)$. This statement can be easily proved by contradiction. If $u$ is adjacent to a node in $I$, it will be in set $J(I)$, which contradicts to the assumption.

**Algorithm:** At the beginning, independent node set $I$ is initiated as an empty set and takes the node with the maximum weight as the first element. Then, the corresponding set $J(I)$ can be formed. Set $J(I \cup J(I))$ can be sequentially formed. The algorithm takes node $u$ which has the maximum weight in $J(I \cup J(I))$ to add in set $I$. Then, it updates sets $J(I)$ and $J(I \cup J(I))$ since a new node is added in $I$. The algorithm repeats such procedures until $J(I \cup J(I)) = \emptyset$, which means there is no independent node in the graph that is out of set $I$. We denote the obtained feasible solution by $I^*$. 

Based on $I^*$, we examine if the obtained independent node set is the highest weighted one we can find, where the weight of set is defined as the accumulation of the weights of the elements in the set. If not, the examination procedure will give a new higher weighted independent node set at the end. The algorithm searches a node $v, v \in V - I^*$ such that node $v$ has and only has an adjacent node $w, w \in I^*$. All such nodes that have not been selected in $I^*$ form a node set $L$. If such a node cannot be found, the independent node set $I$ is the maximal independent node set in our search process. Otherwise, we form a new independent node set. It picks up the the maximum weight node from set $L$, for example, node $v$. Then, add $v$ in set $I^*$ and remove the adjacent node $w$ in $I^*$. After this, it searches the node that
does not adjacent any nodes in \( I^* \). For a set of such nodes, it picks up the node with the maximum weight and adds it into set \( I^* \). It continues until no such a node can be found. The new formed independent node set is used to compare with the previous one to see which one has higher weight. Repeat this on the independent node set, until all nodes in set \( L \) is visited. If a higher weighted set is found, it will be examined in the same way. Otherwise, it is terminated and the maximal independent node set is obtained.

After obtaining the maximal independent node set, the algorithm allocates available subchannel to the nodes in the maximal independent node set \( I^* \). The node weighted constraint graph is then updated for the next iteration. The weights of the nodes in the set are reduced since they are planned to obtain the resources for the edge traffic loads. For each node in the set, if the node weight is non-positive and the allocated resource is above the edge sub-band lower bound, the node and corresponding edge(s) will be removed from the node weighted constraint graph. In the case that the node reaches the maximum allowed resource, the node and corresponding edge(s) are also removed. The iteration terminates until the node weighted constraint graph is empty or all the resources have been used up. The center sub-band of each MR is determined by the complementary set of all subchannels allocated to the edge sub-band.

**Complexity:** The computational complexity in determining all the nodes adjacent to set \( I \) is no more than \( \min\{6|I|, |V| - |I|\} \) since the maximum degree of nodes is no more than six. Similarly, to find set \( J(I \cup J(I)) \) is no more than \( \min\{6|J(I)|, |V| - |J(I)|\} \). To find the maximum weighted node in \( J(I \cup J(I)) \), the computational complexity is no more than \( |J(I \cup J(I))| \). The cardinality of set \( I \) is increased from 1 but no more than \( |V| \). The number of steps of all these
operations is \( O(n^3) \), where \( n = |V| \).

Searching if there exists a higher weighted independent node set from set \( J(I^*) \) and determining \( J(I^*) \) from \( G(V, E) \) needs no more than \( 6|I^*| \) steps. To determine if the node in \( J(I^*) \) has and only has one adjacent node in \( I^* \) takes no more than six steps. After finding such a node and exchange it with the corresponding node in \( I^* \), the computational complexity in the rest procedures to find a new independent node set is \( O(n^3) \) as discussed above. Then, it performs at most \( n \) times such exchange and search. So, the computational complexity is \( O(n^4) \).

**Example:** We use the example in Figure 5.5 to show the algorithm procedures. Figure 5.5(a) is the weighted graph transformed from Figure 5.3(a). For the node sequence \( V = [ A1, A2, A3, B1, B2, B3, C1, C2, C3, D1, D2, D3 ] \), the initial node weight vector is \( w(V) = [ 2, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 1 ] \), as shown in Figure 5.5(a). For the
edge zone throughput reservation factors, suppose the factor of node $B_2$ is set to $B_L^{(B_2)} = 2$ and others are set to zeros. For the edge zone throughput constraint factors, suppose the factor of node $D_1$ is set to $B_U^{(D_1)} = 1$ and others are set to 6.

During the independent node set selection, $A_1$ is firstly selected to add in set $I$, then we have $I = \{A_1\}$, $J(I) = \{A_2, A_3, B_3, C_2, C_3, D_2\}$, and $J(I \cup J(I)) = \{B_1, B_2, C_1, D_1, D_3\}$. We select a node from set $J(I \cup J(I))$, for example, $B_2$, to add it in set $I$. Then, the sets are updated as $I = \{A_1, B_2\}$, $J(I) = \{A_2, A_3, B_1, B_3, C_2, C_3, D_2\}$, and $J(I \cup J(I)) = \{C_1, D_1, D_3\}$. It repeats until $J(I \cup J(I)) = \emptyset$. We can have an independent node set $I^* = \{A_1, B_2, C_1, D_1\}$ as shown in Figure 5.5(b).

After the independent node set examination, we know that $I^*$ is already the maximal weighted independent node set. An updated weighted graph can be generated afterward, whose updated weight vector is $w(V') = [1,1,1,1,0,1,0,1,1,1,1,1]$. From set $I^*$, we have the corresponding allocation vector $A_1^{(V)} = [1,0,0,0,1,0,1,0,0,1,0,0]$.

When compare the resource allocation with the edge zone throughput lower bounds of those nodes having non-positive weight, we can find $A^{(B_2)} < B_L^{(B_2)}$ and $A^{(C_1)} > B_L^{(C_1)}$, which means $B_2$ has not reached its edge zone throughput lower bound. So, $C_1$ is removed from the graph but $B_2$ is kept for the next assignment. Also, we check the nodes that have positive weight and compare them with their edge zone throughput upper bounds. We can find that $A^{(D_1)} = B_U^{(D_1)}$ while others are still less than their edge zone throughput upper bounds. Thus, we also remove $D_1$ from the graph. The remaining graph is shown in Figure 5.5(c).

Then, we repeat the above procedures to find the independent node set for the next channel assignment. During the search of independent node set, we may ob-
tain a set \( I^* = \{A1, B1, D3\} \), as shown in Figure 5.5(d). With the examination algorithm, we can find an independent node set with higher weight, which is \( I^* = \{A2, B1, C3, D3\} \), shown in Figure 5.5(e). The later set has higher weight and is selected as the maximal independent node set for the second subchannel. Figure 5.5(f) is the remaining graph after the second channel assignment. These procedures are repeated until all the channels are assigned.

### 5.4.3 Transmit Power Allocation

After the subchannel allocation for zones, some residual capacity of subchannel used by adjacent cells still can be utilized. For cell \( u \) using a subchannel \( i \) in the center zone, if \( i \) is not used by adjacent cell edge zones, \( u \) can use the subchannel with transmission power \( P_{c1} \) instead of \( P_{c2} \) (\( P_{c1} > P_{c2} \)). Every subchannel allocated to a cell’s center zone needs to check the usage of the subchannel in the adjacent cell edge zones and then adjusts the transmit power so as to maximize the throughput in the center zone.

### 5.4.4 Analysis

Different from the analysis of the FFR schemes considered in Section 5.2, frequency sub-bands in the zones is not a fixed value. Center zone subchannel set \( \mathcal{F}_c \) in Eq. (5.10) is changed to \( \mathcal{F}_{c1}^{(u)} \), since each cell has different center zone subchannel configuration. The center zone throughput for \( \mathcal{F}_{c1}^{(u)} \) and \( \mathcal{F}_{c2}^{(u)} = \mathcal{F} - \mathcal{F}_{c1}^{(u)} - \mathcal{F}_{e}^{(u)} \) are computed by:

\[
C_{Dc1}^{(u)} = \sum_{i \in \mathcal{F}_{c1}^{(u)}} \int_{S_e} \tilde{c}_{i,k}(s) \tilde{x}_i(s) \varphi_{i}^{(u)}(s) ds,
\]

(5.24)
and

\[
C_{Dc2}^{(u)} = \sum_{i \in \mathcal{F}} \int_{S_e} \zeta_{i,\mathcal{K},\mathcal{P}_e}^{(u)}(s) \varphi_i^{(u)}(s) ds. \tag{5.25}
\]

Due to the impacts of edge zone resource reservation and constraint factors, center zone throughput, \(C_{Dc}^{(u)} = C_{Dc1}^{(u)} + C_{Dc2}^{(u)}\), is constrained as given by Eq. (5.26):

\[
C_{Dc}^{(u)} \mid \mathcal{F}_{c1} = \emptyset, \mathcal{F}_e = B_U^{(u)} \leq C_{Dc}^{(u)} \leq C_{Dc}^{(u)} \mid \mathcal{F}_{c1} = |\mathcal{F}| - B_L^{(u)}.
\tag{5.26}
\]

From Eq. (5.24), Eq. (5.25), and Eq. (5.26), we can have some interesting results. When \(B_L^{(u)}\) and \(B_U^{(u)}\) are equal to \(|\mathcal{F}_{SAe}^{(u)}|\) of soft FFR A and if we assume every subchannel has the same usage condition, we can have:

\[
C_{Dc}^{(u)} \mid |\mathcal{F}_{c1} = \emptyset, |\mathcal{F}_e = |\mathcal{F}_{SAe}^{(u)}| = C_{Dc}^{(u)} = C_{Dc1}^{(u)} + C_{Dc2}^{(u)} = C_{SAc}^{(u)} + C_{SAc}^{(u)} = C_{SAc}^{(u)}.
\tag{5.27}
\]

If we set \(B_L^{(u)} = B_U^{(u)} = |\mathcal{F}_{SBc}^{(u)}|\) under same assumptions, we can let D-FFR center zone throughput be:

\[
C_{Dc}^{(u)} \mid |\mathcal{F}_{c1} = \emptyset, |\mathcal{F}_e = |\mathcal{F}_{SBc}^{(u)}| = C_{SBc}^{(u)}.
\tag{5.28}
\]

For edge zone, \(\mathcal{F}_e^{(u)}\) in D-FFR is possible to take any subchannels from set \(\mathcal{F}\) which is different from previous FFR schemes that only take the set from the fixed set \(\mathcal{F}_{c1}, \mathcal{F}_{c2},\) or \(\mathcal{F}_{c3}\). So, edge zone throughput for D-FFR is calculated by:

\[
C_{De}^{(u)} = \sum_{i \in \mathcal{F}_e^{(u)}} \int_{S_e} \zeta_{i,\mathcal{K},\mathcal{P}_e}^{(u)}(s) \varphi_i^{(u)}(s) ds. \tag{5.29}
\]

In addition, edge zone throughput is constrained by:

\[
C_{De}^{(u)} \mid |\mathcal{F}_e^{(u)}| = B_L^{(u)} \leq C_{De}^{(u)} \leq C_{De}^{(u)} \mid |\mathcal{F}_e^{(u)}| = B_U^{(u)}.
\tag{5.30}
\]
We also can have special setting to let D-FFR edge zone throughput to reach soft FFR A or soft FFR B. If \( B_L^{(u)} = B_U^{(u)} = |F_{SAe}^{(u)}| \) and every subchannel has same conditions, we have:

\[
C_{De}^{(u)} |_{B_L^{(u)} = B_U^{(u)} = |F_{SAe}^{(u)}|} = C_{SAe}^{(u)}. \tag{5.31}
\]

If change \( B_L^{(u)} \) and \( B_U^{(u)} \) to \( |F_{SBe}^{(u)}| \), we have:

\[
C_{De}^{(u)} |_{B_L^{(u)} = B_U^{(u)} = |F_{SBe}^{(u)}|} = C_{SBe}^{(u)}. \tag{5.32}
\]

In summary, with the setting of edge zone throughput factors, we can achieve different throughput division on zones, which includes the performance of soft FFR A and soft FFR B. If we do not use soft reuse, it also can achieve hard FFR. With the adaptive adjustment algorithm in our proposed D-FFR, it ensures the system can still achieve high spectrum efficiency with differentiable setting on each cell. By Eq. (5.26) and Eq. (5.30), we can know the lower and upper bounds of center and edge zone throughput by our D-FFR.

### 5.5 Performance Evaluation

In this section, we investigate the performance of our proposed schemes using extensive simulation results with following observations: adaptive adjustment regarding the changes in the traffic loads, ability to control the throughput between the edge zone and the center zone, the interference mitigation in cells with different edge zone and center zone throughput divisions, and the intelligent utilization of adjacent cell capacity.

Figures 5.6, 5.7 and 5.8 show the effect of these adjustments on the edge and the center throughputs. These include throughput per cell, throughput in edge zone, and
Figure 5.6: Throughput per Cell

Figure 5.7: Throughput in Edge Zone
Figure 5.8: Throughput in Center Zone

Figure 5.9: Allocation Adjustment
Figure 5.10: Throughput in Edge Zone

Figure 5.11: Throughput in Center Zone

114
Figure 5.12: Overall Throughput

Figure 5.13: Throughput in Edge Zones
throughput in center zone, with changes in the traffic loads. The MCs are assumed to be uniformly distributed in the cells and each MC requires a unit resource. As discussed before, soft FFR B is the extreme case of soft FFR A, which maximizes the edge zone throughput. Hard FFR has similar performance as soft FFR A, but soft FFR A has higher throughput when the traffic load is heavy. The actual throughput division of edge and center zones is adaptive to the traffic load, but the constraint factor can limit the maximum amount of the resource allowed for the allocation to edge zone. As shown in Figure 5.6, with different configurations of constraint factor, the overall cell throughput with our scheme are able to span between the cell throughput curves of soft FFR B and soft FFR A.

Figure 5.7 shows the increase of the edge zone throughput. When the edge zone throughput constraint factor is relaxed, the maximum edge zone throughput approaches the soft FFR B scheme. This is because no constraint on the resource allocation to the edge zones. When the traffic load is heavy, more resource is allo-
cated to the edge zones. For a cell, the increase of the edge zone throughput is at the cost of decrease in the center zone throughput. By controlling the edge zone constraint factor, the division of cell throughput between center and edge zones can be adjusted by our scheme. When the constraint factor $B_U$ is decreased with sequence 10, 8, 7, and 5 in the figure, the edge zone throughput is also decreased. In contrast, the corresponding center zone throughputs are increased. Figure 5.8 shows the corresponding center zone throughput with the same configuration. The proposed scheme has much higher center zone throughput as compared to that of soft FFR B, even with the tight constraint factor on the edge zone throughput. This is because the proposed scheme has dedicated center zone sub-bands. While the traffic load is light, this feature can support higher center zone throughput.

From Figures. 5.6, 5.7, and 5.8, when D-FFR sets $B_U(5)$, its cell throughput and throughput division on zones of D-FFR are like that of soft FFR A, which means soft FFR A is a special case of our D-FFR. When D-FFR does not set the constraint on edge zone throughput $B_U(\infty)$, its edge zone throughput is like that of soft FFR B, but D-FFR has higher center zone throughput. Same to the analytical results, these simulation results also shows that our scheme can effectively perform the adjustment of the throughput division between edge and center zones.

Figure 5.9 compares the edge zone throughput with three different adjustment methods. Given the cells located as the configuration, we increase the traffic load in the edge zone of cell 1 (the center cell). The reason to adjust the center cell is to ignore the side effect in the simulation. The edge zone of the center cell has the same interference from all directions. To deal with the traffic increase at cell 1 edge, “Soft FFR A -AllCellAdj” is to adjust all cell edge sub-bands at the same reuse distance with same configuration, e.g., the cells using the same sub-band as cell 1 extend the
subchannels same to cell 1. “Soft FFR A -IdvCellAdj” only extends the edge sub-band of cell 1. “Soft FFR A -IdvClstAdj” reduces the edge sub-band of cell 2 (in the same cluster) to surrender subchannel(s) for the edge sub-band of cell 1. From Figure 5.9, we see that the proposed scheme has the best improvement, which increases with the traffic load. Other adjustment methods violate the constraints discussed in Section 5.3. So, the edge zone throughput is not able to be increased. It is even seriously decreased in “Soft FFR A -AllCellAdj”. Our proposed D-FFR scheme considers the constraints and is able to achieve appropriate resource allocation with lower computational complexity.

Figures. 5.10 and 5.11 show the adjustment of throughput division between the edge zone and the center zone in a single cell. We take the six cells around the center cell for observation, because they have similar interference conditions under the same configuration. We give an index number for each cell as 2, 3, 4, 5, 6, and 7 (the center cell is numbered by 1). Then, we adjust one of the cells, i.e., cell 2. By decreasing the edge zone throughput constraint factor, the edge zone throughput of cell 2 is decreased as expected. In contrast, the center zone throughput is able to adaptively increase as shown in Figure 5.11, because the scheme exploits the resource released by the edge zone. Due to the adaptive allocation of our scheme, the decrease of edge zone throughput does not have much effect on other cell’s throughput. The scheme algorithm adaptively adjusts the subchannel allocation to avoid the potential interference caused by the behavior that the decrease of the edge zone throughput at cell 2. As the figures show, other cells can still maintain a similar performance while cell 2 makes the adjustment. Thus, the throughout of different cells and zones can be differentiable.

Figures. 5.12, 5.13, and 5.14 illustrate the network performance when the traffic
load is not uniformly distributed in the cells. Through these figures, we show the ability of the proposed scheme utilizes the subchannels that are not used by the edge zones of adjacent cells. Figure 5.12 shows the overall throughput of the network while increasing the traffic load at the edge zone of the center cell. While the traffic load increases, soft FFR A stops increasing due to reach its maximum resource allowed to allocate in the edge zone. Soft FFR B also reaches its maximum edge zone throughput at a latter point. The proposed scheme still can increase the edge zone throughput due to the ability to use the subchannels not allocated by adjacent cells at the edge zones. Because bounded by the edge zone throughput constraint factor, the throughput of D-FFR will stop at an upper bound as shown in Figs. 5.12 and 5.13. Given a higher edge zone throughput constraint factor, the edge zone throughput is also larger. The edge zone throughput curve for the algorithm without edge zone throughput constraint can keep increasing as shown in Figure 5.13. The center zone throughput for the one without edge zone throughput constraint will decrease, because the throughput increase of the edge zone begins to hurt the throughput at the center zones of cells. Therefore, appropriate setting of the factor enables the utilization of unallocated adjacent cell subchannels to be within a reasonable region.

5.6 Conclusion

Three major types of FFR schemes: hard FFR, soft FFR A, and soft FFR B, have been modeled and compared. To support differentiable fractional resource allocation at cells in a multi-cell WMN, D-FFR scheme is proposed and analyzed with featured discrete partition on frequency band. By considering ICI mitigation and resource allocation constraints, the proposed D-FFR scheme is able to adaptively allocate
different amount of resource for cell zones to support differentiated traffic loads and enhance network performance by optimal operations on resource arrangement and power control. The analytical and simulation results show that our proposed D-FFR can effectively perform differentiable cell and zone throughput with adaptively fractional resource allocation satisfying various constraints.
Chapter 6

Conclusion and Future Work

6.1 Conclusion

As technology evolved from both wireless multi-cell and multi-hop systems, WMN takes advantages of both of them, which offers a promising perspective in providing ubiquitous and pervasive broadband wireless access (BWA). Due to the issues inhered from multi-hop transmission, network capacity is decreased with the network size, and network delay is increased with the network size. Different from a pure ad hoc network, WMN is able to adjust the capacity by controlling MR deployment, channel assignment, transmission power, etc. Our dissertation obtains asymptotic bounds of capacity and delay in a WMN and discussed their relationship. The results provide the range of network delay and capacity with different types of routing strategies. A comparison of results indicate that different routing strategies can be selected according to the scale of a WMN. Furthermore, some of our results are applicable to the analysis of ad hoc networks.

The procedures of deriving asymptotic bounds indicates that the interference is the main problem affecting the capacity, so channel assignment is shown to be crit-
ical the performance of WMN. Extending from the analytical work, we proposes two types of channel assignment schemes: centralized channel assignment and distributed channel assignment. Centralized channel assignment is suitable for ISP deployed WMN, which can manage each MR to achieve globe optimization for higher network capacity and lower network delay. On the other hand, distributed channel assignment meets the requirements of self-organizing WMN, where MRs connect to each other by ad hoc fashion. To address the channel scanning and the frequent channel switching problems special to self-organizing procedures, we also propose additional queue management method for distributed channel assignment.

Besides the channel assignment in a WMN backbone, the resource management in MR cells is also essential. Multi-cell network architecture is formed by network access coverage of MRs. To improve spectrum efficiency, spatial reuse of frequency is utilized by multi-cell architecture, which inevitably introduces ICI and compromises quality of service. As the frequency resource is spatially reused, mobile stations (MSs) may be interfered by additive signals from multiple neighboring cells operating on the same radio at the same time. To mitigate ICI, frequency reuse in previous multi-cell systems divides frequency band into multiple orthogonal partitions and identical partition is only reused at cells with certain distance away. The number of partitions, defined as frequency reuse factor, can be used to identify the space distribution of frequency resource. Higher frequency reuse factor can reduce ICI significantly; however, it also greatly decreases the spectrum efficiency due to less frequency resource available at each cell.


6.2 Future work

There are several further research work we can undertake and some of them to follow are:

Capacity and delay analysis in mobile MRs. While MRs are deployed on mobile objects such as vehicle, airplane, and so on, the capacity and message delay will be different from the results we obtained in a static WMN. For mobile MRs, it also can consider to use carry-forward method to delivery message, where the message received by a moving MR will be buffered for a while before it sends to a next hop MR. In such a case, the movement of MRs could be used to help the delivery of messages. This method may be not preferred by delay-sensitive application. However, it could provide higher capacity for a network, which is good for delay-tolerant applications. The investigation on the balance of capacity and delay can be further conducted and the results can apply to many areas.

Channel assignment for mobile MRs. The channel assignment for mobile WMN is also a very challenging issue in future work. It requires a distributed channel assignment scheme and able to perform fast channel assignment with high frequent channel switch rate. The distributed channel assignment proposed in Chapter 5 can be further developed and used in mobile WMN.

Fractional frequency reuse in self-organizing WMN. Current FFR is based on the regular topology, which has predictable number of neighboring MRs. While the MRs are randomly deployed and self-organizing, the implementation of FFR is much more complicated. To investigate the ICI mitigation under such situation, an efficient protocol is needed for MR to discover the frequency usage situation of neighboring MRs and utilize the information to form a self-organizing FFR.
Acronyms

AP  access point  
BS  base station  
BWA  broadband wireless access  
CSMA  carrier sense multiple access  
CDMC  center for distributed and mobile computing  
FFR  fractional frequency reuse  
FR  forwarding radio  
GSM  global system for mobile communications  
ICI  inter-cell interferre  
IGW  Internet gateway  
ISP  Internet service provider  
IP  Internet protocol
MAC  media access
MANET  mobile ad hoc network
MC  mesh client
MR  mesh router
OFDMA  orthogonal frequency-division multiplexing access
QoS  quality of service
PDF  probability density function
PHY  physical
RR  receiving radio
SNR  signal to noise ratio
SINR  signal to interference and noise ratio
SIR  signal to interference ratio
VoIP  voice over IP
WLAN  wireless local area network
WMN  wireless mesh network
Bibliography


