A Dissertation submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

Scalable MAC Protocol for D2D Communication for future 5G Networks

Bushra Ismaiel

2018

University Of Technology Sydney,
Faculty Of Engineering and Information Technology, School Of Electrical and Data Engineering

Supervisor

Dr.Mehran Abolhasan

Co-Supervisor

Dr.Daniel Franklin

Declaration Of Authorship

I, Bushra Ismaiel, declare that this thesis titled, Scalable MAC Protocol for D2D Communication for future 5G Networks, and the work presented in it are my own. I confirm that:

- The work is done solely while in candidature for a research degree at this University.
- The work done in this thesis has not been previously submitted/published for a degree.
- The work of others have been quoted, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- Any help that I received in my research work and the preparation of thesis itself has been acknowledged.

'This research is supported by the Australian Government Research Training Program'

Signature of Student: Production Note:
Signature removed prior to publication.

Date: 15^{th} October 2018

Table of Contents

List of Figures

List of Tables

List Of Abbreviations

D2D	Device-to-Device
QoS	Quality-of-Service
LTE-A	Long term evolution advance
5G	Fifth generation
MAC	Medium access control
eNB	Evolved node B
PCF	Point coordinated function
DCF	Distributed coordinated function
CP	Contention period
CFP	Contention free period
TDMA	Time division multiple access
FDMA	Frequency division multiple access
CSMA/CA	Carrier sense multiple access collision avoidance
CSMA/CD	Carrier sense multiple access collision detection
RTS	Request-to-send
CTS	Clear-to-send
NAV	Network allocation vector
ACK	Acknowledgement
SIFS	Short Interframe Space
DIFS	DCF short interframe space
PIFS	PCF short interframe space
TXOP	Transmission opportunity
TBTT	Target beacon transmission time

List Of Abbreviations

WLAN	Wireless local area network
SC-MP	Scalable MAC Protocol
LTE-U	Long term evolution unlicensed
LTE-LAA	Long term evolution licensed assisted access
LWA	Long term evolution and Wi-Fi aggregation
SDN	Software defined network
BS	Base station
WBS	Wireless local area base station
HCF	Hybrid coordination function
LBT	Listen before talk

List Of Parameters

r_k	Polling state of user$_k$
v_k	Allocation of time slot to voice traffic of user$_k$
m_k	Time slot is allocated for video/multimedia traffic to a Wi-Fi user after satisfying the channel condition B or allow D2D communication to a Wi-Fi user if the neighbour has already downloaded the video/multimedia file and satisfy the channel condition A
d_k	Allows user to do D2D communication in Markov process.
P_V	Probability for voice traffic
P_M	Probability for video/multimedia traffic
P_N	Probability that video is not downloaded by any neighbour
P_S	Channel condition B is not satisfied to allocate time slot
P_D	Channel condition A is not satisfied for D2D communication
N_n	Total number of neighbours that downloaded the video
t_v	time spent to transmit the voice data
t_m	time spent to transmit the video/multimedia data
t_d	time spent to transmit the D2D data
d_v	data rate for voice
d_m	data rate for video/multimedia
d_d	data rate for D2D communication
$\mathbf{Q}$	Semi-Markov process transition matrix
Γ	Semi-Markov process diagonal matrix
ϕ	eigen value of matrix H
T_r	Duration of r in semi-Markovian model

List Of Parameters

T_v	Duration of voice traffic in semi-Markovian model
T_m	Duration of video/multimedia traffic in semi-Markovian model
T_v	Duration of voice traffic in semi-Markovian model
T_m	Duration of video/multimedia traffic in semi-Markovian model
$M_k(t)$	Moment generating function
$y_{n,k}$	binary variable to select the band
$\mathbf{B}_n$	Total bandwidth
$C_{n,k}$	Effective capacity of licensed/unlicensed band of user$_k$
$\theta_{n,k}$	Quality-of-service of licensed/unlicensed band of user$_k$
$\alpha_{n,k}$	Bandwidth allocated to user$_k$ of licensed/unlicensed band
D_{th}^k	Delay threshold of user$_k$
$\gamma_{n,k}$	Signal-to-interference noise ratio of user$_k$ in licensed/unlicensed band
$\mathbf{R}_k$	Minimum data rate of user$_k$
P_{th}^k	Probability threshold of delay bound
$\mathbf{M}$	Markov process transition matrix

ABSTRACT

Due to the steep growth in mobile data traffic, it will be a challenge for 5G networks to ful-fill the requirement using limited resources in licensed spectrum. However, the joint deployment of smaller cells in the Macro-cell has attempted to overcome this issue. It is observed that users are adversely affected by limited resources in the licensed band. Due to the scarcity of resources in the licensed band, it is better to deploy a small cell operating at an unlicensed spectrum like WLAN. Establishing Device to Device communication (D2D) in the cooperative deployment of cellular networks and WLAN can accommodate the on growing user data demand by intelligently allocating the resources, hence, forming a centralized control in a distributive manner.

This Thesis gives a detailed overview of all the LTE technologies operating in an unlicensed band which includes; LTE-U, LAA, LWA, and MuLTEfire. The technologies are compared with extensive simulation and further D2D communication is applied in these technologies to observe their behaviour.

This Thesis also introduces a three-tier architecture for next generation 5G networks which can offload traffic from cellular networks to WLAN in a dense environment. It proposes a Scalable MAC Protocol (SC-MP) to efficiently allocate resources for Wi-Fi users with D2D communication. SC-MP will allocate WLAN resources to the normal users in a centralized and efficient manner based on a novel PCF strategy, which will develop a centralized control in a distributive manner. The SC-MP is compared to legacy DCF protocol defined in IEEE 802.11 through extensive simulation to evaluate the network performance. The key result is that SC-MP is able to improve the performance compared to DCF for metrics that include; network throughput, network capacity, and network delay.

Furthermore, the thesis gives a detailed mathematical analysis of SC-MP using Markov modelling and semi-Markov modelling. Effective capacity is derived using three-state semi-Markov modelling for the proposed SC-MP. Analytical results are validated through the simulation results. In addition, an optimal queue scheduling and resource allocation problem with QoS guaranteed between the licensed and unlicensed band is formulated to minimize the bandwidth of licensed spectrum and maximize the aggregated effective capacity of a three-tier network. The results proved that the proposed SC-MP can perform better compared with the state of art.

ACKNOWLEDGEMENTS

I would like to thank God for making the Ph.D journey easy for me. There are so many people that helped me, professionally and personally, through my Ph.D journey. There were many moments that I could not see a light at the end of the tunnel, but I am glad that the support and encouragement that I received along the have brought me to this point. I would like to acknowledge all the people who contributed to this journey.

First, I would like to thank my principal supervisor, Dr. Mehran Abolhasan, who had given me an opportunity to join as a post-graduate researcher. He has always been supportive and a source of encouragement and have an endless source of new ideas to investigate. I would like to thank Wei Ni, for giving me new ideas and helping me through the analytical part and giving me time whenever it was required. I would like to thank David Smith, for all the guidelines he has given me through my Ph.D. tenure. Also, I would like to thank Daniel Franklin, for tirelessly providing guidelines in technical writing.

I would like to thank my parents and my family, especially my kids, who have been supportive throughout my Ph.D. journey. I am thankful for all the sacrifices you have willingly made over the past few years to support me in so many endeavours.

In the end, I would like to thank my colleagues for the interesting discussions, technical support and motivation.

Chapter 1

Introduction

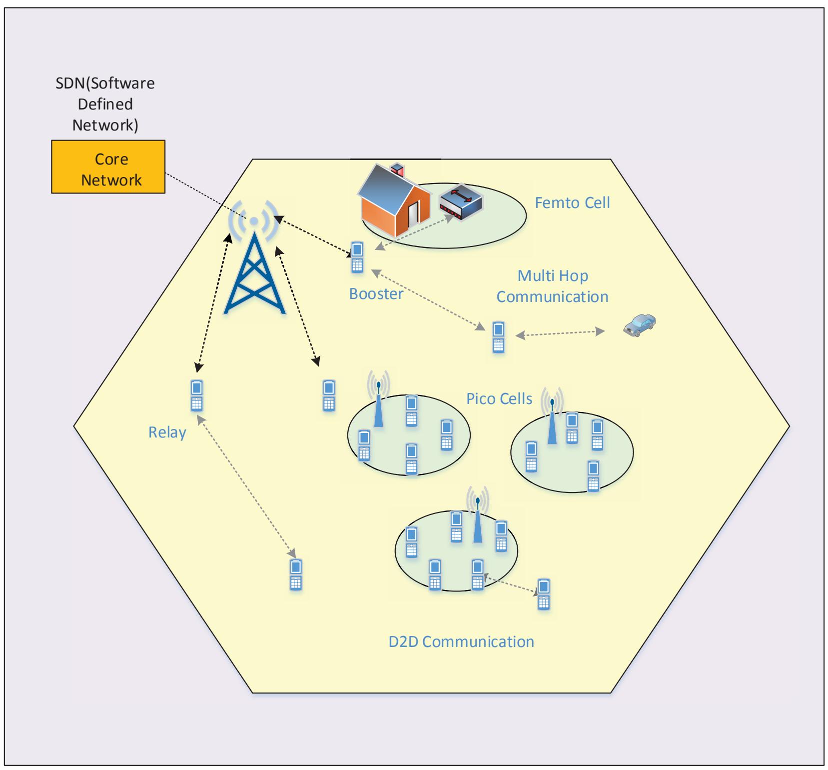


Figure 1.1: Heterogenous Wireless Communication Network Towards 5G Networks

Nowadays, the world is dominated by modern science and technology, and the internet is playing a major role in everyone's life. During the past few decades, the volume of mobile data traffic has increased at a rapid pace due to popular multimedia applications on smartphones and tablets. Besides the high data volume problem, the user experience is also a major challenge. It will be a challenge for current networks to give good Quality-Of-Service in a dense environment where users are in close proximity to each other. Currently, wireless communication systems revolve around 3G and 4G networks. To meet the requirements of future technology and enhance the communication networks, worldwide technologists are looking forward to the next generation wireless systems, 5G.

5G is expected to substantially improve wireless communications and enable a wide range of machine-type communication applications, such as e-Health, vehicular safety, and industrial automation.

The Third Generation Partnership Project (3GPP) is a collaborative project aimed at developing global specifications for the current and future generations of wireless networks. 3GPP caters for the majority of the telecommunication wireless networks in the world. According to 3GPP, the 5G network will produce drastic changes by introducing new dimensions, such as an entirely new air interface and a new transport layer based on ICN (information-centric networking). 5G will propose an improvement in already developed techniques and one of them is D2D communication, which is already introduced in LTE-A (LTE advanced). 5G will also influence the core infrastructure by enabling software defined networks (SDN), packet layering and encryption flexibility.

The question that arises in our minds is why do we need 5G, and what additional benefits it can provide to the wireless industry?

5G will emerge in an attempt to address the following challenges:

- Higher Data Rate
- Capacity enhancement

- Lower E2E latency
- Massive Device connectivity
- Cost reduction
- Consistent QoS provisioning
- Wearable device with artificial intelligence (AI)
- Ability to connect the user to different Wireless Access Technologies like 2.5G, 3G, 4G, Wi-Fi and WPAN (Wireless Personal Area Network)

Due to the substantial benefits, a considerable number of researchers and industries are working toward to making 5G a reality. Ericsson, Alcatel-Lucent, Huawei, DOCOMO, Nokia and European Telecom companies launched 5GPPP (5G Infrastructure Public Private Partnership) in 2014. 5GPPP created METIS, whose primary function is to help launch 5G by 2020 [1]. Since 5G is an emerging technology, a significant number of papers is published in the various area of 5G. One of the important ideas of designing 5G cellular architecture is to seperate the outdoor and indoor scenarios to prevent the penetration loss through building walls [2]. In paper [2–4], the authors propose 5G outdoor network architecture in which the BS are deployed outdoor in the middle of the cell. Outdoor BS are equipped with large antenna arrays that can benefit massive MIMO technologies [2]. In paper [5], the author suggests an indoor architecture and in such an architecture the user communicate with indoor wireless AP with large antenna arrays installed outside the buildings [2]. Papers [6–11], explain a detailed overview of 5G networks, in addition to its features and challenges.

1.1 Research Areas in 5G Networks

Some important research areas of 5G are discussed below:

- *Massive Multiple Input and Multiple Output (MIMO):* Massive MIMO contain multiple transmitters and receivers to multiply the capacity of radio links. It is a technique for sending and receiving more than one data signal with the same radio channel simultaneously via multipath propagation [12]. Massive MIMO is deployed to overcome interference. It enables direct communication with beamforming and also increases the capacity of the cell. Paper [13], proposes the hybrid architecture with beamforming solution for 5G. An extensive survey of cooperative MIMO channel models is presented in [14].

- *Cognitive Radio Technology:* In Cognitive radio, multiple radio technologies share the same spectrum, hence allowing the cellular networks to effectively utilize the under-utilized frequency bands without causing interference. Since 5G will interconnect with all wireless technologies, the cognitive radio will adapt to these wireless technologies and will work with all of them. The author in the paper [15], gives an overview survey of cognitive radio including, the fundamental requirements, architecture, applications, and its challenges.

- *Millimetre Wave Frequency (mmWave):* Millimeter frequency band (20-60 GHz) will be implemented to support a data access speed of 10 Gbit/s. The connection comprises short wireless links at the end of the fiber wired connection, hence creating a converged fiber-wireless network. The author in [16], proposes a mmWave communication survey towards 5G, which includes, characteristics, challenges, and applications of mmWave communication.

- *Relaying:*. Multi-hop relay communication has gained global attention and will be a promising technology for the next generation wireless networks. To improve the coverage and throughput, it is estimated that large-scale relaying will be deployed in 5G [17]. Multi-hop cooperative networks have the capability to increase the capacity, density and to reduce energy consumption by bringing the RAN closer to the end user [18]. The author in [19] proposes two-way relaying in 5G with multiple antennas at the nodes.

- *Full Duplex:* Other favorable techniques in 5G include Full-duplex. Full duplex can double the spectrum efficiency by simultaneous transmission and reception on the same frequency and time resource. It can also reduce E2E packet delay and improve network efficiency [17].

- *Software Defined Networks (SDN):* SDN and virtualization is an innovative feature towards 5G as they can enable flexible and automated deployment. To gain higher speed and massive connectivity, it is important that decisions made at the core side should be faster, structured and congestion controlled. Morever, SDN concept enables us to adapt the operation of the back-haul network to the needs of the radio access network [18]. This can be achieved by using software defined controllers that appear to have a global view of the network and is a single logical switch to applications and policy engines. SDN provides better inter-cell management, enforcement of Quality of Service (QoS), firewall policies, and seamless user mobility across different technologies. Researchers have proposed different SDN architectures towards next generation in papers [20–22].

- *Heterogeneous Networks (HetNets):* HetNets refers to the deployment of different radio networks in a single cell that has different transmission power, radio access technology (RAT), and are supported by a different type of back-haul links. Deploying smaller cells (Pico/Femto/Micro/Relay Nodes) in a macro-cell will not only increase the coverage but will also enhance the capacity and Quality of Service (QoS). Additionally, HetNets will play a vital role in enabling spatial and frequency reuse. In [23–25], resource allocation schemes are proposed in heterogeneous networks, based on QoS. Papers [26–

29], propose ideas to offload traffic from a macro-cell to a smaller cell using Common Radio Resource Management (CRRM). A comprehensive survey of HetNets has been presented in [30].

- *Device to Device Communication (D2D):* Due to the massive connectivity of devices in 5G, D2D communication will play a vibrant role towards 5G networks. D2D communication is gaining immense popularity nowadays, as it can provide high throughput, extended network coverage, efficient spectral usage, improved energy efficiency, delay, and fairness. D2D communication enables users to communicate with each other with minimum involvement from the Base station (BS). D2D communication offers decentralized location-based services that enable efficient, flexible, and secure applications, including social network applications. D2D communication is still an unexploited research area, that requires extensive survey efforts from industry as well as from the academia. D2D communication can be complex as it has many challenges to be addressed before it can be fully adopted, such as interference management, resource allocation, Quality of Service (QoS), communication session set-up, upgrading of network and network cost.

Heterogeneous networks and Device-to-Device (D2D) networks are the focus of this thesis and will be discussed in detail in Chapter 2.

1.2 Thesis Statement

Device-to-Device communication is one of the important features towards the improvement of 5G networks that can offload traffic in a heterogeneous network using licensed or unlicensed spectrum. Introducing D2D communication in a heterogeneous network can bring new challenges, including decision-making criteria for scheduling problems, radio resource management, coexistence with other techniques and technologies. All these issues should be addressed in order to integrate D2D communication in heterogeneous networks.

D2D communication will allow direct packet transmission between peers with little or no involvement from a cellular base station (BS). Efficient resource allocation is one of the major challenges for D2D communication either in cellular or ad-hoc networks. Researchers have provided multiple solutions for resource allocation in cellular networks and proposed different MAC protocols in ad-hoc networks but no work has been proposed yet that can suggest an efficient MAC protocol to offload traffic in heterogeneous networks using both cellular and ad-hoc networks.

1.3 Objectives and Overview of Thesis

The objective of this thesis is to develop an efficient MAC protocol in a three-tier heterogeneous network to offload traffic from cellular networks to wireless local area networks (WLAN) based on IEEE 802.11 MAC. A Scalable MAC Protocol (SC-MP) is proposed for D2D communication based on the IEEE 802.11 PCF access mechanism. A novel SDN-based mechanism for PCF based on the best Signal to Noise Ratio (SNR) polling scheme is introduced in contrast to a round robin scheme. The throughput of the network is further increased by sharing multimedia through D2D communication between WLAN users. Our proposed architecture is built on the LWA technology with modification in the resource allocation scheme based on IEEE 802.11 PCF to access a WLAN channel with an additional increase in throughput by D2D communication. The protocol entitled, Scalable MAC Protocol (SC-MP), increases the network capacity while decreasing the network delay and giving fair channel access to users by using a polling coordinated function (PCF).

A second major objective is to develop a new analytical model for SC-MP that is novel in the depth of analysis offered in terms of effective capacity and saturation throughput. A hypothetical system is set-up to measure the effective capacity of Scalable MAC Protocol

(SC-MP), under statistical QoS limitations using a three-state semi-Markovian model and further analyzing the throughput using a four-state traditional Markov chain.

Furthermore, the optimal joint queue scheduling and resource allocation problem is formulated with the QoS guarantee between licensed and unlicensed band to minimize the bandwidth of licensed band. A new iterative algorithm is proposed to convexify the problem as a series of block coordinated descent (BCD) and Difference of Convex functions (D.C) program.

1.4 Thesis Outline and Contributions

This section comprises the outline of the thesis and summarizes the main contributions.

- In Chapter 2, an extensive review of relevant literature is presented, including MAC protocols for both infrastructure and ad-hoc D2D networks. Furthermore, the features and challenges of a D2D MAC protocol - in particular, several over-looked and under-explored areas such as fairness, scalability and interoperability - are discussed.

- In Chapter 3, a detailed overview and comparison of LTE technologies operating in an unlicensed 5GHz band that includes; carrier Wi-Fi, LTE-U, LTE-LAA, LWA, and MuLTEfire are discussed. A detailed simulation is performed for LTE-LAA, LWA and MuLTEfire technologies in the presence of a Wi-Fi hotspot and results are compared. Furthermore, the performance of these technologies is investigated by applying D2D communication over them. It is concluded from the results that MuLTEfire can increase the throughput of users drastically but the network saturates quickly. For a scalable network, it will be best to use D2D with LWA as it increases the throughput and the capacity of the network.

- In Chapter 4, an innovative three-tier network is proposed for a dense environment based on LWA technology, consisting of Macro BS, WLAN BS, and D2D communication. A new centralized Scalable MAC Protocol (SC-MP) is proposed for D2D communication between WLAN users, based on the IEEE 802.11 Point Coordination Function (PCF) access mechanism. The proposed scheme will allocate WLAN resources to users in a centralized and efficient manner based on a novel PCF strategy, which will develop a centralized control in a distributive manner. A polling scheme in PCF will be based on the basis of best SNR in contrast to a round robin scheme. Simulation results show that the proposed MAC scheme can increase the network capacity while decreasing the network delay and perform better than the legacy distributed coordination function (DCF) defined in IEEE 802.11.

- In Chapter 5, An analytical three state semi-Markovian model is proposed for SC-MP based on PCF access mechanism with D2D communication. A new closed-form expression is theoretically derived to evaluate the effective capacity of SC-MP against the QoS and instantaneous transmission rate for random diverse traffic; voice and video. The three-state semi-Markovian model is modified to a traditional four-state Markovian model using a discrete Markov chain to analyze the saturation throughput having persistent traffic for voice and video/multimedia. Furthermore, the behaviour of D2D communication is observed in the model. The analytical results are validated using simulation tool Matlab.

- In Chapter 6, an optimal queue scheduling and resource allocation problem is investigated under various statistical delay constraints of three-tier network based on LWA technology with a modification in resource allocation scheme based on IEEE 802.11 PCF to access WLAN channels and further offload multimedia files through D2D communication. The optimal joint queue scheduling and resource allocation problem is formulated with the QoS guarantee between licensed and unlicensed bands to minimize the bandwidth of the licensed band. A new iterative algorithm is proposed to convexify the problem as a series of block coordinated descent (BCD) and difference of convex functions (D.C) program. A simulation is performed using two scenarios of the proposed schemes; in the first scenario, the voice traffic uses a licensed band whereas multimedia traffic uses a unlicensed band and in the second scenario half of the voice traffic goes through the licensed band and another half of the traffic goes to the unlicensed band along with the multimedia traffic. Our result proved that both scenarios perform better than SMS and scenario 2 outperforms scenario 1.

- In Chapter 7, the contributions and key results of the thesis are summarised. Directions for possible future research are also discussed.

1.5 Related Publications

The publications related to the contribution of the Thesis are as follows:

- Performance comparison and high-level simulation of LTE solutions operating in an unlicensed 5GHz band; LTE-Unlicensed (LTE-U), LTE-License Assisted Access (LTE-LAA), LTE WiFi Link Aggregation (LWA), and MuLTEfire and further applying D2D communication in these technologies (Chapter 3): Bushra Ismaiel, Mehran Abolhasan, David Smith, Wei Ni and Daniel Franklin, "A Survey and Comparison Of Device-To-Device Architecture Using LTE unlicensed Band", in *IEEE Vehicular Technology Conference (VTC)*, June 2017 [31].

- An innovative three-tier 5G architecture is proposed for D2D communication, which will offload cellular traffic to WLAN in a dense environment using a Scalable MAC Protocol (SC-MP). Simulation results show that the proposed MAC scheme can increase the capacity of the network and perform better relative to the legacy Distributed

Coordination Function (DCF) defined in IEEE 802.11 (Chapter 4): Bushra Ismaiel, Mehran Abolhasan, David Smith, Wei Ni and Daniel Franklin, "Scalable MAC protocol for D2D communication for future 5G networks", in *Consumer Communications and Networking Conference (CCNC)*, Jan 2017 [32].

- A three-state semi-Markovian model is proposed to derive a closed-form expression of effective capacity in terms of transmission rate and quality-of-service (QoS) of SC-MP. Further, SC-MP is analytically modeled using a four-state traditional Markov model to derive the saturation throughput. The analytical results are validated through simulations, hence proving the appropriateness of the model (Chapter 5): Bushra Ismaiel, Mehran Abolhasan, Wei Ni, David Smith and Daniel Franklin, "Analysis of Effective Capacity and Throughput of Polling Based Heterogeneous Networks", in *IEEE Transaction on Vehicular Technology*, May 2018 [33].

- Resource allocation and optimal joint queuing scheduling problems are formulated with the QoS guarantee between licensed and unlicensed bands to minimize the bandwidth of the licensed band using SC-MP. A new iterative algorithm is proposed to convexify the problem as series of sub-problems based on Block Coordinate Descent (BCD) and difference of two convex functions (D.C) program. The simulation results show that the scheme performs better than the existing scheme (Chapter 6): Bushra Ismaiel, Mehran Abolhasan, Wei Ni, David Smith and Daniel Franklin, "PCF-Based LTE Wi-Fi Aggregation for Coordinating and Offloading the Cellular Traffic to D2D Network", to be appear in *IEEE Transaction on Vehicular Technology*, September 2018 [34].

- A survey paper with an extensive literature review of both infrastructure-assisted cellular D2D and distributed/ad-hoc techniques are studied, and their scalability is compared. Furthermore, the features and challenges of D2D MAC protocol - in particular, several over-looked and under-explored areas such as fairness, scalability, and interoperability are discussed (Chapter 2): Bushra Ismaiel, Mehran Abolhasan, David Smith, Wei Ni and Daniel Franklin, "MAC Protocols for Device-to-Device Communication in 5G Networks-A Survey", submitted to *IEEE Access*, July 2018 [35].

Chapter 2

Literature Review

2.1 Introduction

This thesis explores the performance gain of heterogeneous networks using IEEE 802.11 MAC layer with D2D communication. This chapter includes general discussion of architecture of heterogeneous and Device-to-Device networks. Moreover, resource allocation and MAC protocols of infrastructure assisted and ad-hoc networks are discussed in detail. The key topics of this chapter includes:

- Features of heterogeneous networks and comparison with traditional cellular network;
- Different resource allocation schemes in heterogeneous networks with centralized control;
- Features of device-to-device communication and importance towards 5G networks;
- Resource allocation schemes in infrastructure assisted D2D communication;
- MAC protocols in distributed/ad-hoc networks;and
- Comparison of MAC protocols based with their advantages and disadvantages.

2.2 Heterogeneous Networks

Currently, most of the wireless networks are deployed as homogeneous networks. The Homogeneous cellular network is a network in which there are planned base stations with the users connected to it. Base stations have the same transmit power levels, antenna pattern, receiver noise floors and back-haul connectivity to the network. Macro base stations are carefully planned to overcome the interference and minimize the coverage holes.

Heterogeneous Network	**Traditional Cellular Network**
Connects to BS to provide highest data rate	Connects to the strongest signal strength BS
BS are placed opportunistically and locations are better modelled as random process. Smaller cells are nested in macro cell	Hexagonal grid is a best model for BS locations. BS have distinct coverage area
Outage/Coverage probability in terms of rate/ area spectral efficiency	Outage/coverage probability in terms of SINR/spectral efficiency
Downlink and uplink should not necessarily use the same the same BS and can have different SINRs	Same BS uplink and downlink have approximately same SINR. Best downlink and uplink BS
Interference is hard due to irregular back-haul and sheer number of BS	Employ fractional frequency reuse to tolerate very poor cell edge rate
Hand-offs and dropped call are more frequent in terms of high mobility. High overhead.	Hand-offs to stronger BS when poor coverage area

Table 2.1: Heterogeneous Networks Vs Traditional Cellular Networks

Due to the rapid growth of mobile data traffic, it is a challenege for network operators to fulfil the requirement of the users using limited resources in licensed spectrum. Heterogeneous network can increase the capacity while keeping the same infrastructure and deploying smaller cells. Heterogeneous networks are an essential part for the next generation wireless network. Heterogeneous networks are utilizing a diverse set of base stations can be deployed

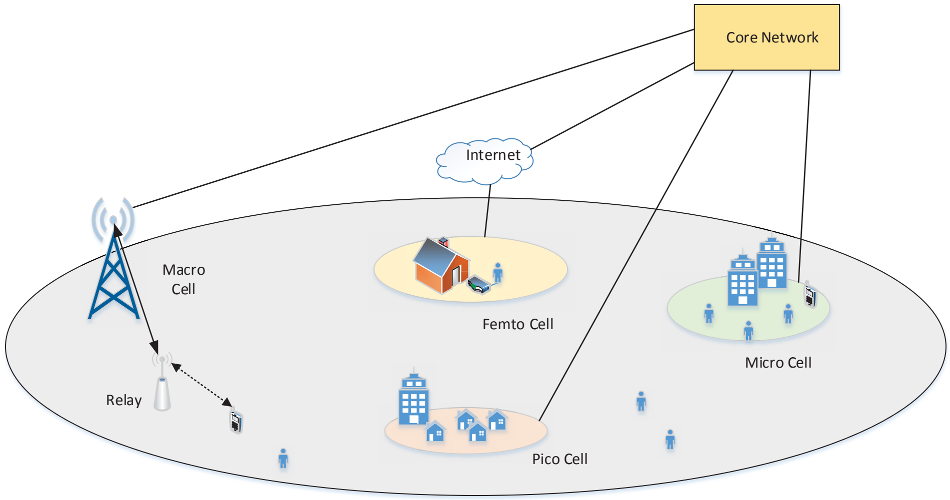


Figure 2.1: Heterogeneous Network Architecture towards 5G

to increase the spectral efficiency of the network.The Table. 2.1 shows a comparison between heterogeneous network and traditional cellular network [30].

A typical heterogeneous network is shown in Figure. 2.1, where a macro base station is overlayed with Pico/Femto cells. Normally the small cells are unplanned and deployed to eliminate the coverage holes present in a homogeneous network. The idea of Heterogeneous network is to overlay low power and low cost devices on coverage holes and increase capacity.The macro cells have large coverage, larger capacity and high transmit power. Small cells include Femto, pico and micro cells which are normally used to extend the coverage or increase the capacity as they have a smaller capacity, smaller coverage area and small transmit power. Properties of different cells are shown in the Table. 2.2, [36].

There are many advantages in heterogeneous networks that are making them a contender for next generation wireless networks [36]:

- Improved spectral efficiency

- Numerous small cells can be used in one macro cell to improve the performance and

Characteristics	**Femto**	**Pico**	**Micro**	**Macro**
Indoor/Outdoor	Indoor	Indoor/ Outdoor	Outdoor	Outdoor
Number of Users	4-16	32-100	200	200-1000+
Maximum of Output Power	20-100mW	250mW	2-10W	40-100W
Maximum Cell Radius	10-50m	200m	2km	10-40km
Bandwidth	10Mhz	10-20Mhz	20-40Mhz	60-75Mhz
Technology	3G/4G/Wi-Fi	3G/4G/Wi-Fi	3G/4G/Wi-Fi	3G/4G
Back-haul	DSL/Cable/ Fiber	Microwave/ mm	Fiber/ microwave	Fiber/ Microwave

Table 2.2: Properties of Different Cells

capacity of the network

- It works on the topology in which Macro cell can operate in licensed band whereas the small cells can operate in licensed/unlicensed band.
- Macrocell is connected to small cell through back-haul or sometime with a booster as well.
- Self organizing network (SON) can be used in HetNets to mitigate the interference and optimize the network.
- Using Wi-Fi (unlicensed band) in multiple tiers can increase the capacity of the network
- HetNets can be used to offload traffic from macro cells to small cells.
- Besides capacity HetNets can decrease the cost as well as smaller cells are much cheaper.
- Deployment of low power devices are easier and efficient.

2.2.1 Resource Allocation in Heterogeneous Networks

In [37] dynamic resource allocation problem for heterogeneous services in cognitive radios with imperfect channel sensing is proposed. Cognitive radio networks consist of primary and secondary users using OFDMA. In each time slot, a secondary channel can sense and utilize the idle channel. Once the channel is sensed by the secondary user, it sends the channel information to the secondary BS which makes the decision regarding whether the channel is idle or not. The major disadvantage of this is that they cannot be applied to all small cells such as Femto cells.

Asynchronous split phase protocol [38] for the heterogeneous network have been proposed. No tight synchronization is required and nodes can hop independently to make efficient use of channel resources. The WiFlex platform has been proposed, and this supports communication among heterogeneous devices with different physical capabilities. Fairness and priority access are also taken into account so one type of device does not starve another. The MAC protocol is divided into three phases. Initially devices in the common control channel go through an observe and review phase, which eliminates the hidden node problem while implementing as implements priority where a higher priority device may override prior reservations. The disadvantage of this protocol is the low efficiency under low load, and idle slots on the data channel, which cause inefficient bandwidth utilization. The result show that the total throughput is higher when no fairness algorithm is used. A tradeoff between fairness and throughput exists.

2.2.1.1 Resource Allocation in Dense Heterogeneous Networks

In [23] load offloading in two tiers is proposed by joint resource partitioning to improve the rate of cell edge users in co-channel heterogeneous network. A resource partitioning is done in such a way that the macro cell stops transmitting for certain time/frequency and let the smaller cells schedule the users of the corresponding resources, this protects from interference as well. Rate distribution was derived in a heterogeneous network taking in account joint resource partitioning and limited bandwidth back-hauls. However, no optimal partition is analytically given.

In [24] a decentralized algorithm is proposed for resource allocation in a heterogeneous network with mobile terminals having dual interfaces. It allows the users to use licensed band (LTE) or unlicensed band (Wi-Fi). The algorithm assigns resources on the basis of call arrivals, call departures, and service requests. It will help to improve the call blocking, sufficient resources to be provided to each user and reduce the overhead, hence offloading traffic between different RAT in an efficient manner. Cellular/Wi-Fi interworking requires communication management through asynchronous radio access technology (RAT) and also interface modifications is also required which make difficult to achieve user service continuity and also resource allocation is hard to achieve.

In [25] proposed a novel approach to offload traffic in a heterogeneous network using Software defined network (SDN). Partial real time services are offloaded using load balancing and also taking in account QoS and network conditions. A new load balancing method is proposed to improve the equilibrium and stability of the network. The offload manager calculates the traffic flow of all the users in the Wi-Fi offloading zone and if Wi-Fi can transfer all the required information of that user within the time delay it will be offloaded to Wi-Fi otherwise partial offload will be applied by the offload manager which is situated in SDN.

The results show that this scheme saves primary resources and decreases the threshold miss probability.

2.2.1.2 Resource Allocation In HetNets Based on QoS

It is difficult to support QoS in heterogeneous networks as many of the users may use these devices. To implement QoS tight coordination is required and especially using unlicensed bands in different tiers may cause a security problem as well. Some papers have proposed

QoS in HetNets which has both their as and disadvantages.

In [39], QoS has been enhanced using Wi-Fi and LTE, small cells that allow for tighter

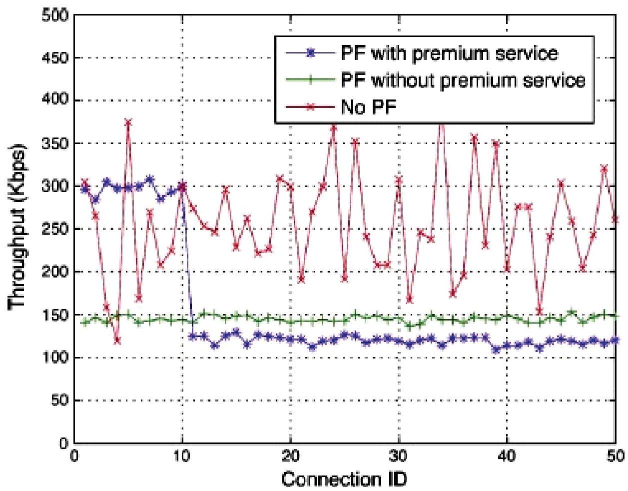


Figure 2.2: Throughput Fairness Test

multi-radio coordination. QoS-aware scheduling algorithms have been designed for on time throughput optimization and measurement of the number of packets delivered to a user prior to the delayed deadline. The Cross RAT coordination scheme is also proposed to decrease the delay and further enhance QoS. This scheme particularly seeks to increase the number of users receiving data throughput in excess of their targeted timely throughput. Users that have received more than the required bits at any given time are not assigned any resources. The Scheduler only finds the users who have received less bit in the given time and hence increase throughput using minimum resources. Figure. 2.2, shows a graph between throughput and number of connections, which gives a comparison between different schemes. The blue line represents the proportional fair (PF) scheduling with premium service, the green line represents the proportional fair (PF) scheduling with no premium service and the red line represent no proportional fair (PF). Initially PF without premium service gives better throughput but as the number of connection increases the throughput drops down rapidly [39].

In [40] different schemes for QoS are compared in HetNets and fault tolerant architecture is also presented to support QoS under multiple heterogeneous networks. The first scheme is the first serve bandwidth allocation scheme where equal priority is employed. In the proposed scheme first user is served and allocated a bandwidth. The network accepts the call as long as it satisfies the users bandwidth requirement. In the second scheme, lowest priority is given to the user which can access most wireless networks. In the third scheme when allocating resources, low QoS users will be given higher priority over high QoS users. The fourth scheme is based on the users QoS history, if a user has always been given high priority by all wireless networks, the user will be given high priority. The last scheme is based on the mobility of the users, the greater the mobility, the higher the priority. Although the last scheme would provide the best results theoretically, it is much more complex than the other schemes [40].
The Table. 2.3 presents the description of the schemes.

2.2.1.3 Common Radio Resource Management (CRRM)

A major challenge in a heterogenous network is the Radio Resource Management (RRM). The common radio resource management is to jointly manage radio resources in overlapped radio access technologies (RAT) in an optimization-based mindset. It is a type of a heterogeneous network in which capacity can be increased by directing some of the traffic to smaller cells or Wi-Fi. Interaction between CRRM and RRM has two main functions; one is information reporting in which RRM report the relevant information to CRRM (available QoS, max bit rate for a given service, average buffer delay, cell load, received power level, etc) and the second function is RRM decision support function which describes the way of

Description of QoS Scheme	Comments
FCFS (equal priority for users)	QoS renegotiation, voluntary downgrade and prioritized update can be employed to improve the overall performance
Priority based on the amount of resources requested	This is likely to reduce call blocking for users with significantly lower QoS requirements
Priority based on the number of networks a user can access at the location	This may result in better placement of users to different wireless networks
Priority based on the QoS history	Likely to involve some additional overhead, but a user with previous record of satisfactory QoS may be more open to a temporarily lower QoS
Priority based on mobility pattern (class)	A faster user is likely to free up the network resources in a shorter time than a slower user
Priority based on two or more factors (number of networks, QoS required, QoS history, mobility class etc)	Many possible combinations exist and much more work is necessary to evaluate the effectiveness of the different combinations of factors used to determine priorities

Table 2.3: A Comparison of Differnt Resource Allocation Schemes for Heterogeneous Networks

interaction between RRM and CRRM based on the decisions [26] .

The advantages of CRRM are:

- Load sharing for efficient usage of resources
- Interference distribution to provide higher spectral efficiency
- Improved QoS management.

Researchers have looked into CRRM and provided different schemes to improve the network performance. Currently, most of the researchers are looking into low to intermediate inter-

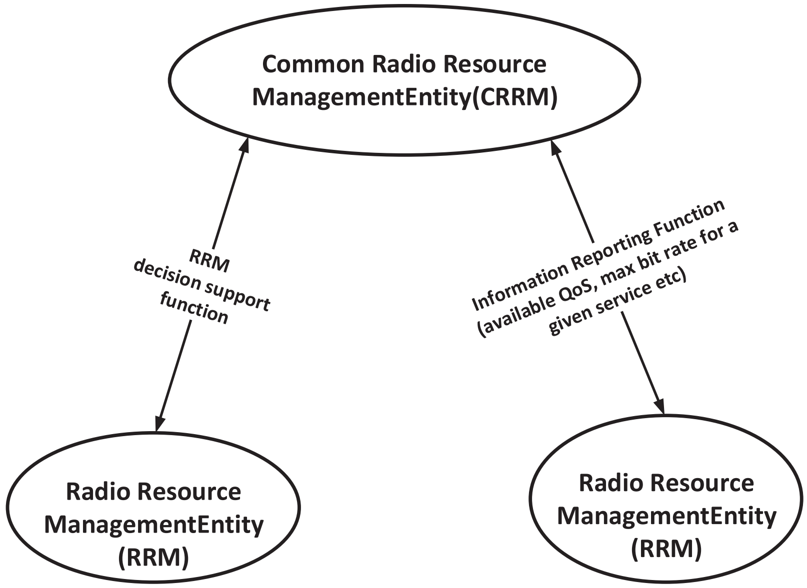


Figure 2.3: Common Radio Resource Management Model

action degree models [26]. Applying CRRM, users can be allocated to most suitable RAT to improve system performance and QoS in terms of load balancing and guaranteed service. WLAN first scheme [41] has been proposed, in which voice and data traffic attempt to connect to WLAN to increase the capacity of BS. The author proves that overall resource utilization can be increased when admission regions for voice and data services in cellular and WLAN are properly configured. The mean data packet service rate is obtained for each vector in the admission region. The access point of WLAN is equipped with the admission control module and decides whether to accept or reject an incoming call based on the numbers of ongoing calls and the admission region. When there are large numbers of voice calls in WLAN, the overall performance degrades.

In [27], a fuzzy neural scheme is proposed to manage common radio resources. It proposes three main functions: RAT and cell selection, bit rate allocation and admission control. Admission control in two steps. In the first step, a combination of three cells is being selected around the different RAT. In the second step, the most suitable RAT is selected and grants bit rate to each user. This scheme determines the necessary bit rate at the assigned RAT, however, it does not address the issue of intra-Rat radio resources assignment.

In [28], the paper proposes the benefits of changing the load based handover thresholds. It improves the load balancing algorithm by proposing an adaptive load threshold rather than a fixed one. In the adaptive threshold algorithm, the load balancing threshold is adjusted depending on a load of its inter-RAT neighboring cells. The higher the load in the neighboring cells, higher threshold is set. Results show that the adaptive threshold algorithm performs better than the fixed threshold algorithm. This scheme is proposed for real-time services, however, does not capture the service dimension problem.

In [42], a scheme is proposed to offload traffic from the Macro cell to WLAN on the basis of mobility. A slow user is connected to the WLAN using CSMA and a fast user will connect to the BS. The sub-channels are allocated to center part and the edge part of the cell. If a slow user is rejected by WLAN it will try to connect to the center part of the cell to BS only if the sub-channel is available. If this is not the case, the user will be connected to the sub-channel of the edge part to the BS. QoS is also taken into account by setting the throughput of WLAN to a certain threshold. Congestion increases if WLAN area increases, which results in a high call block hence degrading network performance.

In [43], discusses different strategies for initial RAT selection. It proposes a policy based algorithm for initial RAT selection. In VG the voice users are connected to the GERAN and other service users are connected to UTRAN, in VU policy, it is vice versa. IN policy have been proposed in which indoor uses are allocated to the GERAN and outdoor users are allocated to the UTRAN. If there is no capacity in the GERAN the user is directed to the UTRAN. Simulation results show that when the voice users are high IN*VG policy performs better than VG*IN. IN*VG also achieve better load balancing as well. However, this scheme only take into account the technical aspect and not the economical aspect.

In [29] proposes three different schemes in combining WWAN/WLAN network. In the first scheme "Always WWAN" users are connected to WWAN as default which decreases the vertical handoffs (VHO) whereas which have a capacity issue. The second scheme "Always WLAN" users are default connected to WLAN which will increase the VHO for high mobility users. The third scheme decides RAT based on different parameters if a user remains in the hotspot area will be connected to WLAN whereas if it will leave the area will be connected to WWAN. The third scheme performs better than the rest two in reference to blocking/dropping probability, but is more complex and require lots of information.

In [44], a new CRRM scheme is proposed to offload traffic from cellular to WLAN on the basis of mobility, load, and service. Whenever a new user arrives, it will try to connect to cellular BS, and if no channel is available, it will run the offloading algorithm. In the offloading algorithm, when there are the resources available in WLAN some of the cellular users in the hotspot region will connect to WLAN and the new users will connect to cellular BS. If the new user is in the overlapping area, the mobility of the user is considered. If the users mobility is above the threshold value, it will connect to cellular BS. If the users mobility is below the threshold and is a non-real time, the service will connect to WLAN. If the users mobility is less than the threshold and is a real time service then cellular BS is selected on the basis of load threshold. The disadvantage of this is, if too many much traffic is on WLAN, blocking probability is high.

2.2.2 Challenges in Heterogeneous Networks

There are few challenges in HetNets from traditional networks [30]:

- Performance metrics
- Network topology
- Mobility

- Interference Management
- Downlink uplink relationship
- Back-haul bottleneck

D2D is gaining immense popularity, as it can provide high throughput, extended network coverage, efficient spectral usage, improved energy efficiency and delay, and green communications through reduced power consumption. Furthermore, a detailed literature review of D2D communication and its mac protocol will be discussed in heterogeneous networks.

2.3 Device to Device Communication (D2D)

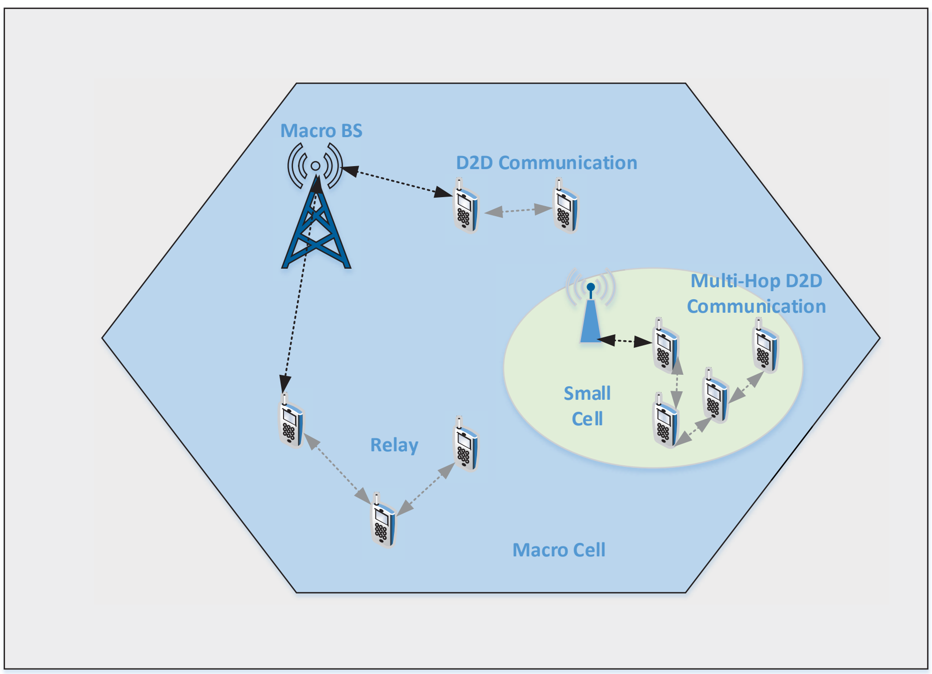


Figure 2.4: Device To Device Networks

Wireless communication is advancing towards 5G heterogeneous networks and one of the integral features of 5G is Device-to-Device communication (D2D). With the massive increase in technology usage and the number of the connected devices, it is predicted that these will continue to increase in the near future, which will increase the traffic volume as well and D2D can play a significant role to fulfill these requirements. D2D communication will facilitate the interoperability between the public safety and commercial networks for example constructing a hybrid system consisting both the cellular and ad-hoc links.

D2D communication was part of 4G/LTE-A (longterm evolution advanced); however, issues with using D2D communication in LTE-A persist, including authorization, authentication, and direct communication between devices that are under subscriptions with different operators [5]. D2D will be an integral part of the next generation of networks, that is, 5G (fifth generation). D2D communication refers to technology that enables devices to communicate without using the network infrastructure.

- *Features of D2D Networks* D2D communication has distinct features that make it attractive for 5G networks. It will be useful for offloading the traffic from the Macro cell to smaller cells. D2D communication will increase spectral efficiency and resources will be saved. Sharing of spectrum resources between cellular and D2D users will result in a capacity gain. In the case of faulty/damaged cellular network, D2D communication will help cellular communication terminals to develop an ad-hoc network. Multi-hop D2D communication can be used for peer-to-peer communication or even access to cellular networks, this can expand the number of wireless applications [45]. D2D communication will also decrease End-to-End latency as the time delay will decrease due to the direct communication link between D2D users. Another important feature of D2D communication is the green environment, it will help to reduce the power consumption, the battery can last for more time as the distances between the devices will be less as compare to the distance between the device and BS. D2D communication can establish a group communication, where the number of devices may vary depending on application and communication approach. Also, using the D2D proximity of equipment will provide high bit rates due to the close proximity and potentially favorable propagation conditions. D2D technology based on nearby user discovery will enhance user experience in location based services, social and commercial activities and data sharing.

- *Types of D2D Networks* There are four types of D2D communication [46]
  - *Device Relay Operator Control (DR-OC):* Communicates with BS through another device acting as a relay where there is poor coverage at the edge of the cell.
  - *Direct D2D communication with operator controlled link establishment (DC-OC):* Two devices communicating with each other without BS involvement, apart from the link which is established by the BS.
  - *Device Relay Device Controlled link establishment (DR-DC):* Operator is not used in link establishment, but the source and destination communicate with each other using relays.
  - *Direct D2D communication with device controlled link establishment:* The source and destination can communicate with each other without operator involvement. Resources should be used in such a way to minimize the interference with other devices in the same or a different radio network.
- *Challenges of D2D Networks:* The first step to establish D2D communication is the peer discovery of the two users; the device should be aware of other devices to which it needs to communicate. Peer discovery depends on the distance between the two devices and it should be short enough in order to support the reliable connection. User

mobility varies the distance between the devices, hence causing a major reason for D2D communication failure. The other major challenge is reusing of cellular frequencies which cause interference between cellular and D2D user. Interference management and allocation of resources should be done intelligently, as it is the crucial part that can have a significant impact on the network performance. Relay selection in the network design is also a challenge in D2D network, as the number of relays in the network should be optimal. Synchronization is also a challenge in scenarios like outdoor coverage or multi-hop D2D networks. Optimum power consumption in also another obstacle in D2D assisted networks, as the energy of the user devices are limited [47]. D2D communication operating in an unlicensed band can be at high-security risk. Security is another area which needs to be addressed in D2D networks.

- *Application of D2D Networks:* D2D communication can be implemented in a number of applications including local services, emergency communication, and IoT things. Local services are mainly used for the social applications like finding nearby users to share data or play games. It will also help local advertising service based on proximity that can target users to maximize its benefits for example nearby shops can send discounts and promotions to users which are in their vicinity. Local services can offload traffic from cellular to WLAN, by sharing the same videos between users, hence creating a hot spot.

  An impromptu network can be setup using D2D in a short time as a replacement for damaged communication network and infrastructure during the earthquake and other natural calamities. Multi-hop D2D can establish an ad-hoc network to ensure the smooth wireless connection between the terminals.

  One of the most important application for D2D communication is IoT; a typical example is a vehicle to vehicle communications (V2V). If a vehicle is at high speed it can warn other nearby vehicles in the D2D mode before slowing down or changing lanes.

Similarly, M2M is also another application that communicates between machines without human interaction. Home appliances and other items embedded with electronics, software, sensors, actuators, and connectivity which enables these things to connect and exchange data are all included in IoT.

Other applications for D2D communication may include; multi-user MIMO enhancement, cooperative relaying and virtual MIMO. Nowadays IPTV is gaining a lot of popularity and it can be watched using D2D communication.

- *Spectrum Sharing:* D2D communication was proposed to optimise spectrum efficiency and resource utilisation. The spectrum sharing approaches in D2D and cellular networks are defined as either outband or inband communication.

  Inband D2D Communication: D2D links in inband communication share the pri-

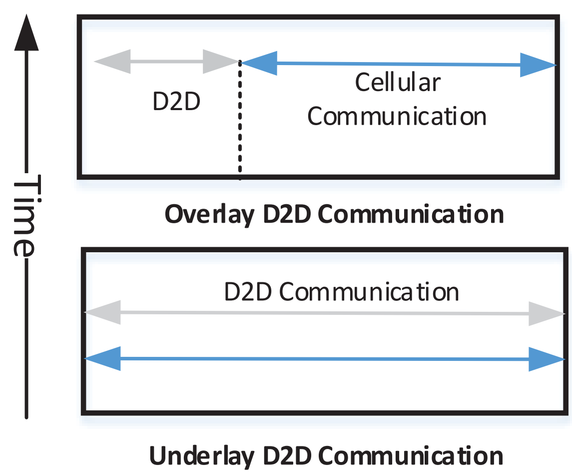


Figure 2.5: Inband Spectrum Sharing in D2D Communication Networks

mary licensed spectrum that is used in regular cellular communications. There are two types of inband D2D communication: overlay D2D communication and underlay

D2D communication. In overlay D2D communication, a portion of licensed spectrum is dedicated to D2D communication, and the other portion is used for cellular communication. In underlay D2D communication, no portion of licensed spectrum is dedicated to D2D communication. The motivation behind using inband communication is the tight control and high security due to the centralised radio management. However, inband D2D communication introduces additional challenges and complexity in interference management.

Outband D2D Communication: D2D links in outband communication use unlicensed

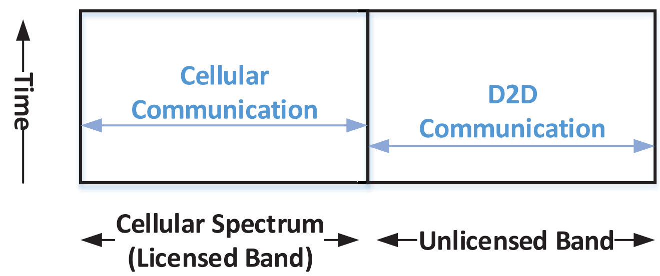


Figure 2.6: Outband Spectrum Sharing in D2D Communication Networks

spectrum to eliminate the interference issue between D2D communication and cellular links. This method requires an additional interface and other wireless technologies, such as Wi-Fi, Bluetooth or ZigBee. Outband D2D communication helps to eliminate the interference between D2D and cellular users, but interference may still be present due to other electronic devices (Bluetooth, Wi-Fi). Cellular devices with two interfaces can use outband D2D communication [48].

- *Categorization of D2D Networks:* D2D networks can be categorized into two groups: infrastructure assisted D2D networks and ad-hoc networks.

  Infrastructure assisted D2D is generally proposed for single hop communication. Infrastructure-assisted D2D communication is a cellular-based service with centralised control. Infrastructure assisted D2D can perform better than ad-hoc networks because D2D operates in a cellular spectrum (licensed band) as well in an unlicensed band under centralized control, whereas Wi-Fi and Bluetooth operate in an unlicensed band with no centralized control, which can become a major problem as proximity based services proliferate. Wi-Fi and Bluetooth require manual pairing which can cause a drawback for proximity services. Since both Wi-Fi and Bluetooth use unlicensed bands, the security features are not suitable for public safety services. One of the key challenges in infrastructure-assisted D2D communication is spectrum sharing, whereas some of the advantages are that it is highly scalable and provides good security. One of the major aspects of MAC protocol design in infrastructure-assisted D2D communication is resource allocation.

ad-hoc networks are mostly distributed wireless network operating in an unlicensed band [49]. MAC protocols in ad hoc network can be decentralised or centralised based on the technique adopted to design the MAC protocol. Ad hoc networks can be very beneficial in applications such as military communications on the battle field [50, 51]. However ad-hoc networks are not scalable risks. Ad hoc networks can be classified into various types of networks: wireless mesh networks (WMN), wireless sensor networks (WSN), mobile ad hoc networks (MANET), and vehicular ad hoc networks (VANET) [52].

2.3.1 MAC Protocols in D2D Networks

The main objectives of MAC protocols are to efficiently utilise the transmission medium and maximise its reuse [53]. Resource allocation is a vital research area in D2D networks, due to its crucial impact on the system performance. In 5G, D2D communication can be used as a multi-hop or relay, therefore, radio resource management and scheduling is more complex and difficult as compared to the traditional cellular network. The selection of the D2D resources is considered to be a key factor for guaranteeing fair, reliable, and interference-free spectrum sharing between cellular and D2D communications, as well as among D2D communications [54].

In recent years, wireless MAC protocols have experienced a number of structural evolutions, as shown in Table. 2.4. In this section, we will review the papers that employ resource

Evolution 1 (1980-1999)	**Evolution 2 (2000-2008)**	**Evolution 3 (2009-2010)**		**Evolution 4 (2011-Present)**
CSMA/TDMA Based Protocols	**Multi-Channel / Multi-Radio MAC Protocols**	**Intelligent MAC protocol Techniques**		**D2D MAC Protocols in LTE/LTE-A network**
		Cooperative / Hybrid MAC Protocols	**Cognitive Radio MAC Protocols**	
MACA [55]	DBTMA [56]	Distributive cooperative MAC protocol for multi-hop networks [57]	CRAHNs [58]	Efficient resource allocation for D2D communication underlaying LTE network [59]
MACAW [60]	MAC protocol with directional antenna [61]	CRBAR [62]	AMRCC [63]	A decentralised spectrum management using D2D and cellular links [64]
MACA-BI [65]	DCA [66]	VC-MAC [67]	DCR-MAC [68]	A stochastic geometry analysis of D2D overlaying multi-channel downlink cellular networks [69]
FPRP [70]	S-MAC [71]	CoRe-MAC [72]	DH-MAC [73]	
PAMAS [74]	RBAR [75]	A hybrid reservation MAC protocol [76]		

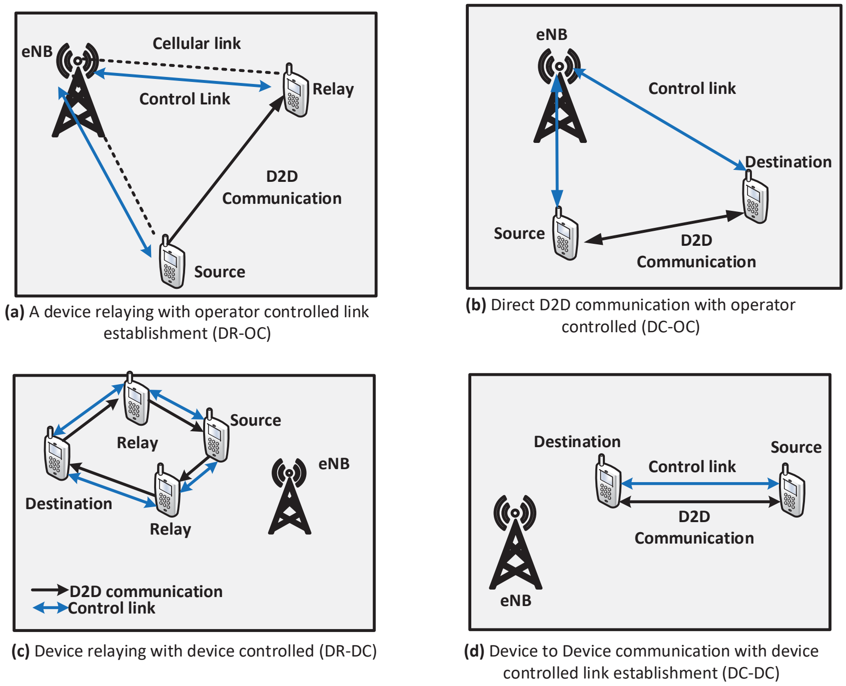
Table 2.4: Evolution in MAC in Wireless Networks

allocation in infrastructure assisted D2D networks and ad-hoc networks along with their advantages and disadvantages.

2.3.2 Resource Allocation in Infrastructure Assisted D2D Networks

Two important features in infrastructure-assisted D2D communication are resource allocation and establishing a link between two users. Different scenarios of D2D communication are shown in Figure 2.7.

In Figure. 2.7 (a), a device relays its messages through another relay using operator-controlled (DR-OC) link establishment, which is suitable when coverage is poor. In Figure. 2.7 (b), two devices communicate through operator-controlled (DC-OC) links can share the content without the need of eNB, as the eNB only establishes the control link. In Figure.



(a) A device relaying with operator controlled link establishment (DR-OC)

(b) Direct D2D communication with operator controlled (DC-OC)

(c) Device relaying with device controlled (DR-DC)

(d) Device to Device communication with device controlled link establishment (DC-DC)

Figure 2.7: D2D communication scenerios underlaying LTE-A Networks

2.7 (c), a device relays via device-controlled (DR-DC) links, where the eNB does not control communication link. In this scenario, more than one device can be used as a relay. In Figure. 2.7 (d), direct D2D communication is established with a device-controlled link (DC-DC). In this scenario, two devices can communicate without the allocation of resources through the eNB. The devices periodically broadcast their identity to other devices.

Infrastructure-assisted D2D communication is cellular based, and the MAC protocols work mainly on TDMA-based strategies. The primary function of MAC is to efficiently utilise the access medium between the nodes and maximise spatial reuse [77]. Radio resource management and scheduling is complex in cellular D2D communication compared to traditional cellular networks, as D2D communication can be used as multi-hop or relay extension. Therefore, the selection of D2D resources is considered to be a key factor for guaranteeing fair, reliable, and interference-free spectrum sharing between cellular and D2D communications [54, 78].

In [59], an optimal resource allocation method is proposed using an underlay system to improve the data rate of D2D communication and the spectral efficiency by reusing more than one cellular user resource by one D2D pair. A D2D user can reuse the resources of a cellular user by guaranteeing a minimum transmission data rate to the cellular user at a given SINR. The main complexity of this method is that resources are allocated based on the channel state information (CSI) of all the links, which is difficult to predict and also requires tight interference control.

In [79], a new bidirectional spectrum sharing technique is proposed to improve the sum rate of D2D users and cellular users using overlay D2D communications. In this protocol, the D2D users can communicate bidirectionally with each other while one of the D2D users assists the two-way communication between the cellular BS and the cellular user. This protocol can further improve the performance by applying relay selection, which is not taken into consideration.

In [80], a distributed multislot multichannel MAC protocol is proposed using outband D2D communication. The protocol allows for concurrent data broadcasting on different channels. D2D synchronisation is implemented in the application of an OFDMA system with multiple resource blocks. D2D transmitters transfer content for the data channels during the contention phase and check their channel list. If there is a control signal on the control channel, the channel ID is later removed from the list. There is no feedback mechanism in this protocol, and half-duplex radio is cost effective.

In [81], a two-stage resource allocation approach is presented to accommodate as many users as possible in a limited spectrum using full-duplex radio and D2D data offloading. In the first stage, the scheduler of the full-duplex-enabled BS allocates wireless subchannels to each pair of uplink and downlink half-duplex UEs. In the second stage, D2D communication is used to address the inter-node interference by replacing fullduplex cellular links with D2D links. This scheme can increase the capacity using limited spectrum resources by utilizing the same resource for uplink and downlink users. Nevertheless, full-duplex schemes are less energy efficient than half-duplex schemes.

In [82], authors focus on maximising the throughput by guaranteeing the quality of service (QoS) through channel selection and power control in heterogeneous network. Two classes of application has been analysed in it; file sharing and streaming. The channel gain information is gathered from the BS and algorithm checks for the optimal one. This scheme provides better resource utilization as compared to other state-of-art allocation scheme but the only disadvantage is the increase in overhead.

In [83], matching theory is exploited for resource allocation. The objective is to match the resources and the users according to a set of quality of service (QoS) metrics. These QoS metrics are classified into three categories: one player can match one player from the opposite set (one-to-one match), one player can match to multiple players of the other set (many-to-one match), and at least one player of the two sets can match to more than one player of the other set (many-to-many matches). Each cellular and D2D user decides their preference in consideration of the channel conditions, transmission power, and QoS metrics. The pair cannot be matched if these criteria are not met. A D2D user can change its preferences to increase performance. In this work, the selection of desirable matching is a critical design issue, and the optimality of a stable solution may not be guaranteed. Additionally, supplementary signalling may also be required during a deferred acceptance algorithm.

In [84], the operator controls resource allocation and data transmission to provide better user experience as compared to traditional free D2D communication. The operator controls access authentication, resource allocation, connection control and lawful interception of communication information. It is a centralised resource allocation scheme in which eNB informs the D2D UEs of the resources available for data transmission using control signalling which gives quick and accurate user discovery. However, it may cause high signal overhead at the BS as compared to light BS control. In light BS control, the BS only periodically broadcasts the set of resources used for transmitting and receiving discovery beacons and users who want to participate in D2D communications can send or listen to the discovery beacon within the discovery resources.

In [85], a scheme is proposed to maximise the throughput and guarantee the quality of service (QoS) of both cellular users (CU) and D2D pairs. A centralised resource allocation scheme is presented using a convex approximation method, and the distributed resource allocation scheme is presented using a Stackelberg game. In the distributed scheme, the BS is the leader and decides the interference price for D2D communication in the uplink to maximise its own profit, and the uplink transmission from the CUs to the BS is protected through the pricing. One complexity in this scheme is the pricing of the resources, which must be done in a novel way for practical implementation

In [86], a new scheme is proposed in which D2D users opportunistically select their transmission mode: direct transmission between D2D users (direct one-hop transmission) or indirect transmission through the BS (indirect two-hop transmission). A stochastic optimization problem is formulated to improve the sum rate of the system while satisfying the QoS of each user. An optimal subchannel scheduling algorithm is applied to provide opportunistic subchannel scheduling and transmission mode selection for D2D users considering the time-varying channel condition of each wireless link and the QoS requirement of each user. Pairing is a key technique to achieve high reuse gains.

In [87],an auction-based distributed resource allocation scheme is proposed for a multitier D2D cellular network. The scheme is proposed to decrease the intracell interference and increase the data rate of the user. However, specific improvement results compared to other approaches are not evident.

In [64], a decentralised spectrum management consisting of both cellular and D2D links is proposed. The Stackelberg game is applied for resource allocation, and the SPPP algorithm is used to identify the optimal price. The interference tolerance level is a predefined parameter, and the BS adjusts the prices to control total interference. D2D communication acquires the CSI of the links from the BS. The CSI is updated depending on user mobility. In this scheme, the BS broadcasts the price and measures the aggregate interference at this price. If the interference is greater than the tolerance level, the cost is increased; however, if this is not the case, the price is decreased. The total rates of D2D communication and cellular links are shown using different algorithms. SPPP and a bisection algorithm provide better D2D/cellular rates than other algorithms

In [88], an efficient resource allocation method is proposed by using a greedy heuristic algorithm. In this scheme, eNodeB controls D2D connections, and the D2D shared radio resource allocation problem is resolved as an MINLP problem. Then, a greedy heuristic algorithm is applied to minimise the interference and improve the gain by using the information provided by the primary scheduler. In this scheme, the cooperation between the cellular users and D2D pairs is not considered, and the interference channel gain alone is not optimal for pairing the cellular user and D2D pair.

In [89], a distributed joint spectrum sharing scheme is proposed to optimise the throughput of D2D users and minimise the interference between cellular users and D2D pairs. Resource allocation is solved through a Stackelberg game. This scheme overcomes the complexity and resolves the interference problem between the cellular user 7 and D2D pair, but only the QoS of cellular users is taken into account.

In [69], a new scheme based on stochastic geometry is proposed. In this scheme, UE at the cell edge will connect to a BS in a two-hop manner, where the D2D links act as the relay. This scheme is used to increase the coverage of a cell. The received signal strength threshold is set to any UE with a value less than the threshold. This enables connection to the BS through the D2D relay (2 hops at the most). Channels are reserved for D2D communication, and if a UE has to communicate through a relay, the BS assigns a channel. In this case, the complexity lies in calculating the threshold of the received signal strength. The threshold should be an optimum value, as an excessively large value will result in congestion in the BS and an excessively small value can lead to high D2D relays.

In [90], a new scheme is proposed to increase the throughput of video files in the cellular network using a clustering and caching strategy. A BS controls D2D communication, and the BS maintains records of all the 6 videos downloaded by the user in its memory. If any user needs to download the same file, the BS directs the user to the nearest device that has the desirable data, which is then transmitted through the D2D link. The results show that this method can improve throughput for high user density in a cell. In this scheme, strategies to request distribution and to bring files into cache memory are of great importance.

In [91], a new scheme using half- and full-duplex channel sharing between a cellular user and the D2D link is proposed using a clustering strategy. The half-duplex scheme is divided into two stages. In the first stage, cluster members remain silent, and the BS transmits the data. By contrast, in the second stage, the BS remains silent and the cluster head transmits data to the cluster members. In the full-duplex scheme, the cluster head transmits and receives in the same frequency band. The drawback of using the half-duplex scheme is that it degrades the reuse gain of the cellular spectrum resource, whereas in full-duplex mode, the transmit powers at the cluster head and at the BS should be well coordinated to eliminate the co-channel interference.

In [92], opportunistic scheduling is proposed with D2D communication using a clustering strategy to improve the throughput in contrast to round robin schedulers, and fairness is provided to the users. The cluster member with the highest CQI (channel quality indicator) becomes the cluster head. The cluster head is responsible for forwarding the cellular traffic of its clients (cluster members of the same cluster) to the BS. The coalition game theory approach is applied to analyse the cluster formation mechanism. The complexity lies in the cluster head life because in the case of stable conditions, the cluster head will discharge quickly compared to the cluster members

In [93], a cognitive-based two-phase solution is proposed for fair resource allocation between cellular UEs and D2D UEs. In the first phase, fair subchannel and rate allocations are performed for the cellular DL and UL flows with max-min fairness. In the second phase, resource allocation for D2D flows with rate protection is performed. This approach does not optimally maximise the spectral efficiency.

In [94], a dynamic common control channel MAC protocol (DCCC-MAC) for cognitive radio networks is proposed. DCCC-MAC eliminates the requirement of a dedicated channel for control information exchange. The common control channel (CCC) is dynamically selected by a set of secondary users (SUs) using a support vector machine (SVM)-based learning technique. The protocol include four main phases: in the first phase, SUs sense the channel; in the second phase, SUs select the CCC; in the third, phase data are transmitted; and in the last phase, beaconing occurs. The proposed protocol minimises the control overhead by finding the CCC efficiently using SVM-based machine learning. The author in this paper does not identify the difference between this proposed scheme and other classical CR MAC protocols.

In [95], a synchronous distributed opportunistic scheduling protocol under fairness constraints (DO-Fast) is proposed. DO-Fast uses a round robin strategy to integrate the opportunistic scheduling. In this algorithm, at the beginning of each frame, D2D links broadcast and order their channel state indicators (CSI). D2D links are divided into two groups: high CSI order and low CSI order. In each traffic slot of the frame, the D2D links in each group select their priorities randomly within their group. The complexity lies in selecting the optimal group, as the throughput and delay can increase if the low CSI order group is selected.

In [96], a new scheme, FlashLinQ, is proposed to extend the managed services by cellular providers by deploying an ad hoc network. FlashLinQ is a distributed resource allocation scheme based on OFDM (orthogonal frequency division multiplexing) synchronous

MAC/PHY architecture for D2D communications. The objective is to schedule a channel-state-aware maximal-independent set at any given time slot based on the current traffic and channel condition, and the scheduling algorithm leads to spatial throughput gains over an ad hoc system. This resource allocation scheme can be used in fully controlled D2D communication networks and in loosely controlled D2D communication.

A comparison summary of all the MAC protocols in infrastructure-assisted D2D networks, the problem they address and their drawbacks is shown in Table. 2.5.

Proposed Scheme	**Problem Addresses**	**Disadvantages**
Optimal Resource Allocation [59]	Increase the data rate and spectral efficiency	Channel state information for all links is hard to predict. Tight interference control is required
Matching Theory for Resource Allocation [83]	Improve the Quality of Service (QoS)	Selection of desirable matching is critical issue. Supplementary signalling is also required during deferred acceptance algorithm
Operator control Resource Allocation [84]	Improve the Quality of Service (QoS)	Overhead is increased
Two stage Resource Allocation [81] Scheme	Increase the capacity	Less energy efficient
Stochastic Geometry for Resource Allocation [69]	Increase the coverage	Complexity lies in calculating the signal strength threshold. Too large will result in congestion and too low will lead to high D2D relays
Distributed Multi-slot Multi-channel MAC Protocol [80]	Increase the data rate	No feedback mechanism. Increase the cost due to half-duplex radio.
Clustering and Caching Scheme [90]	Increase the throughput of Video files	Complexity lies in requesting and bringing the files in the cache memory.
Stackelberg Game for Resource Allocation [85]	Improve Quality Of Service (QoS) and maximise the throughput	Complexity lies in defining the price for interference
Auction based Distributed Resource Allocation Scheme [87]	Increase the data rate and decrease the intra-cell interference	Comparison with other techniques are not clear
Dynamic Common Control Channel MAC Protocol for Cognitive Radio [94]	Decrease the control overhead	No comparison is shown with other similar cognitive radio MAC Protocols
Opportunistic scheduling using clustering strategy [92]	Increases throughput, energy efficient and provide fairness	Cluster head charging drains quickly in stable conditions
Channel selection and Power control for Resource Allocation [82]	Increase the throughput and improve Qulaity of Service (QoS)	Overhead is increased
Distributed Opportunistic scheduling protocol under Fairness constraints (DO-Fast) [95]	Provides fairness	Delay increases if low CSI group has more chance to transmit
Efficient Resource Allocation Using Greedy Heuristic Algorithm [88]	Increase gain and minimises Interference	Only interference channel gain is not optimal to pair the cellular user and D2D pair.
Distributed Joint Spectrum Sharing for Resource Allocation [89]	Increase throughput and minimise Interference	Only QoS of cellular users are taken into account
Opportunistic sub-channel selection in D2D [86]	Improve the sum-rate and QoS	Pairing is the key technique for high reuse gains
Bi-directional Spectrum Sharing for Resource Allocation [81]	Improve the sum-rate	Relay selection should have further adopted to improve the performance

Table 2.5: Summary Table for MAC Protocols Of Infrastructure Assisted D2D Networks

2.3.3 Ad-hoc Networks

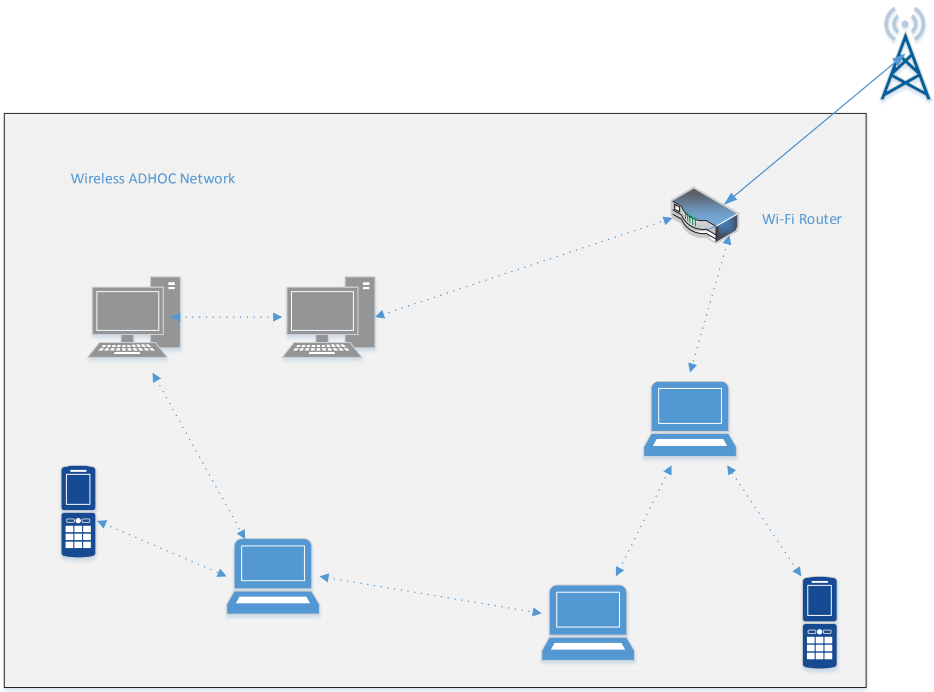


Figure 2.8: Ad-hoc Networks

Ad-hoc networks are decentralised wireless networks operating in an unlicensed band that do not rely on pre-existing networks. A wireless distributed/ad-hoc network is shown in Fig. 6. Ad-hoc networks are a type of multi-hop wireless network that can communicate with each other through a common channel. Ad-hoc networks also have limitations that include small wireless transmission range, high packet loss in dense environments, limited bandwidth and power constraints.

A comparison between ad-hoc and Infrastructure assisted D2D network is shown in Table. 2.6.

Ad-hoc networks can be classified into a number of different networks: Wireless Mesh Network (WMN), Wireless Sensor Networks (WSN), Mobile Ad-hoc Networks (MANET) and Vehicular Ad-hoc Networks (VANET).

Infrastructure Assisted D2D Networks	Ad-hoc Networks
Operate in licensed and unlicensed band	Operate in unlicensed band
Centralized Control	No centralized control
In licensed band there are less security risks	Operates in unlicensed band high security risks
Mostly used in one or two hops as a relay	Works in multiple hops
Adventageous for proximity services and public safety	Not suitable for proximity services
Recommended for highly dense environment	Not suitable for highly dense environment

Table 2.6: Infrastructure Assisted D2D Networks Vs Ad-hoc Networks

Wireless Sensor Networks (WSN): Wireless sensor networks are a type of ad-hoc networks in which the sensors are spatially distributed to monitor physical or environmental conditions and to pass their data through the network to the main location.

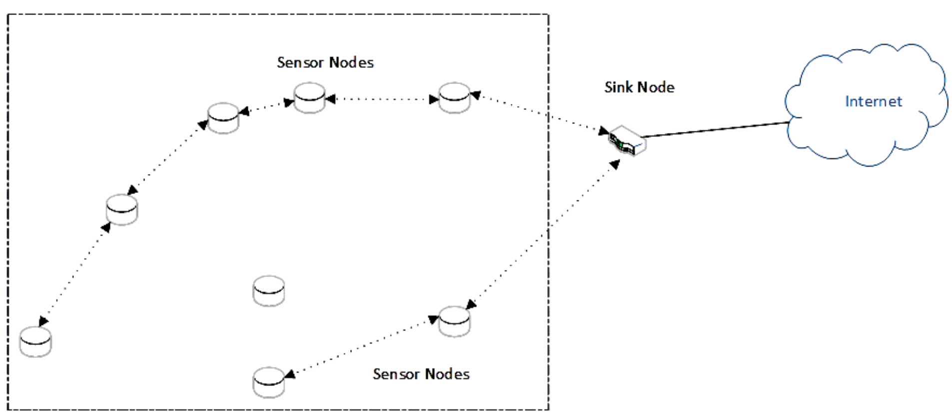


Figure 2.9: Wireless Sensor Network

Wireless Mesh Networks (WMW): Wireless mesh network is a type of ad-hoc network in which radio nodes are connected in a mesh topology. They consist of mesh clients, routers and gateways.

Mobile Ad-hoc Networks (MANET): Mobile ad-hoc network is infrastructure less of mobile

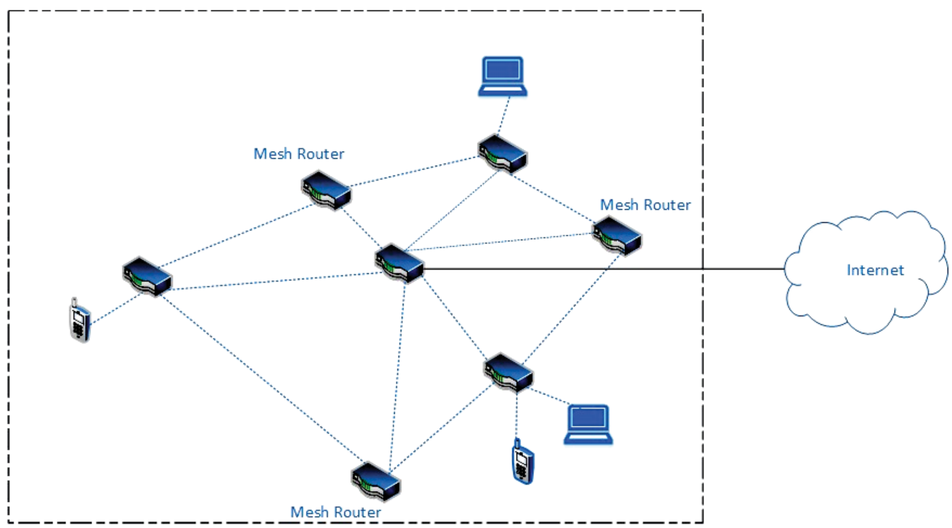


Figure 2.10: Wireless Mesh Network

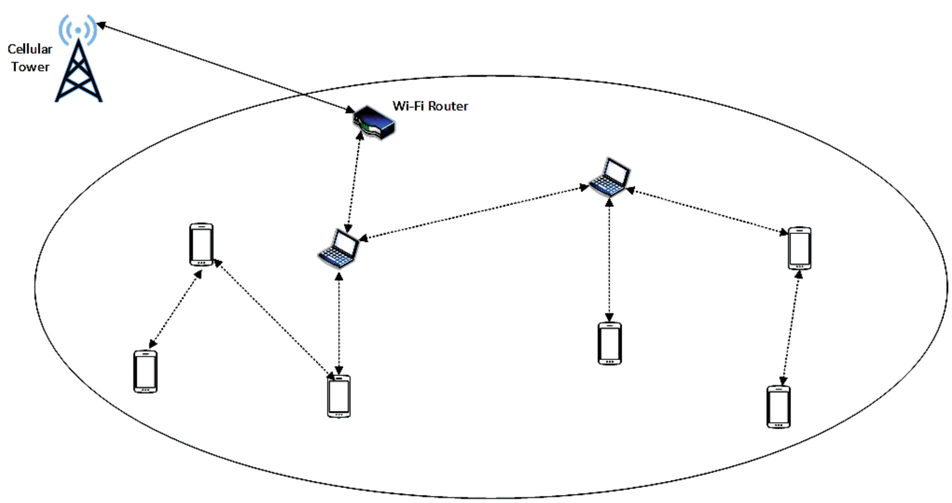


Figure 2.11: Mobile Ad-hoc Networks

devices. They are free to move anywhere frequently, which changes their links with other devices.

Vehicular Ad-hoc Networks (VANET): Vehicular ad-hoc networks were created on the principle of mobile ad-hoc networks. They provide wireless communication services between vehicles and vehicle to roadside infrastructure.

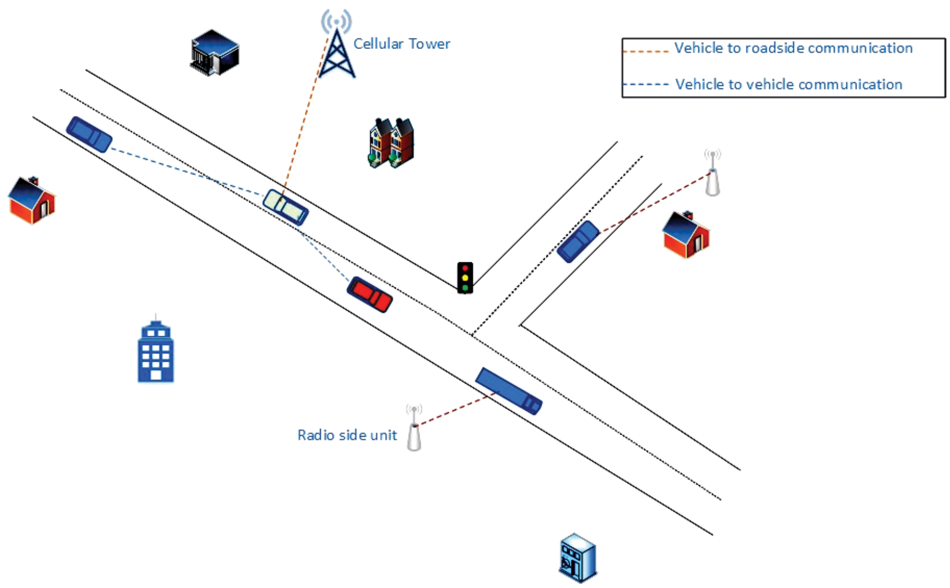


Figure 2.12: Vehicular Ad-hoc Networks

2.3.4 MAC Protocols in Ad-hoc Networks

Certain measures should be taken into account when designing an ad hoc MAC protocol, including efficient bandwidth utilisation, minimum control overhead, few delays and minimal packet retransmission [97–99]. Furthermore, energy consumption should be minimised, and the MAC protocol must be scalable to large networks [97, 100].

IEEE 802.11 has already standardised the WLAN MAC Protocol in distributed coordination function (DCF), point coordination function (PCF) and hybrid coordination function (HCF).

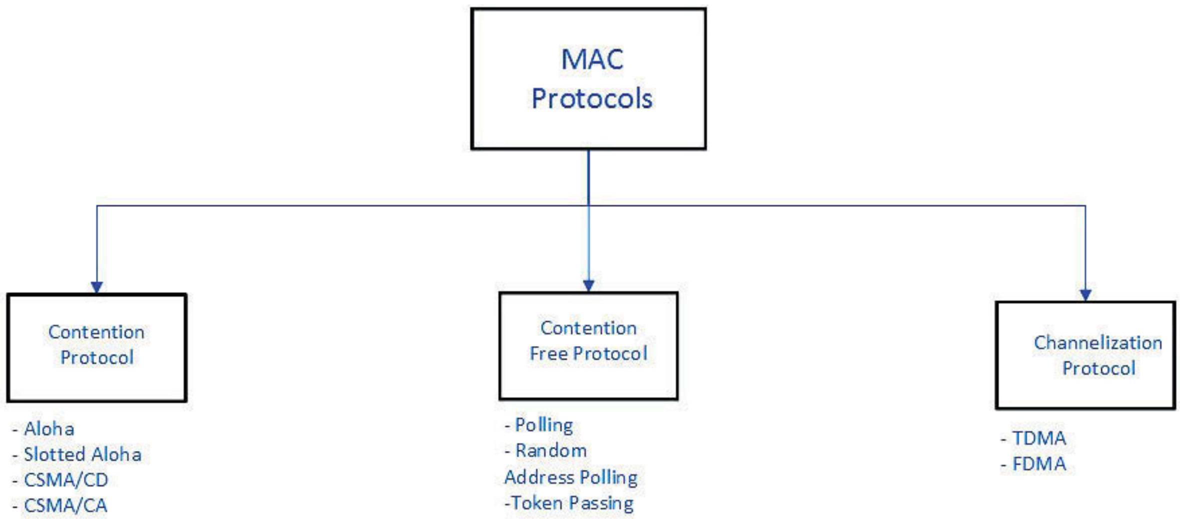


Figure 2.13: Types of MAC Protocols in Ad-hoc Network

Besides IEEE 802.11-based MAC standards, a number of MAC protocol variations and improvements have been proposed. These protocols are categorised into contention protocols, contention-free protocols and channelization protocols. The different types of MAC protocols are shown in Figure. 2.13.

2.3.4.1 Contention Protocol

The fundamental access method for the contention protocol is DCF, which works on the principle of carrier sense multiple access collision avoidance (CSMA/CA). CSMA/CA supports asynchronous data transfer on a best effort basis. CSMA/CA uses carrier sense functionality and a binary exponential backoff (BEB) mechanism. Each window, and the backoff timer decreases for each idle time slot (TS). The node starts transmitting packets when the timer expires or reaches zero. If the transmission fails, the size of the CW doubles. When a channel is idle for the DCF interframe space (DIFS) period, the winning node sends an RTS request to the receiver, and after receiving the RTS request, the receiver sends a clear-to-send (CTS) message to the sender. Once the sender receives CTS, it sends the packet, which allows the receiver to send an ACK message [101–103]. RTS-CTS Mechanism is shown in the Figure. 2.14. To date, the majority of the literature on contention protocols focuses on improving

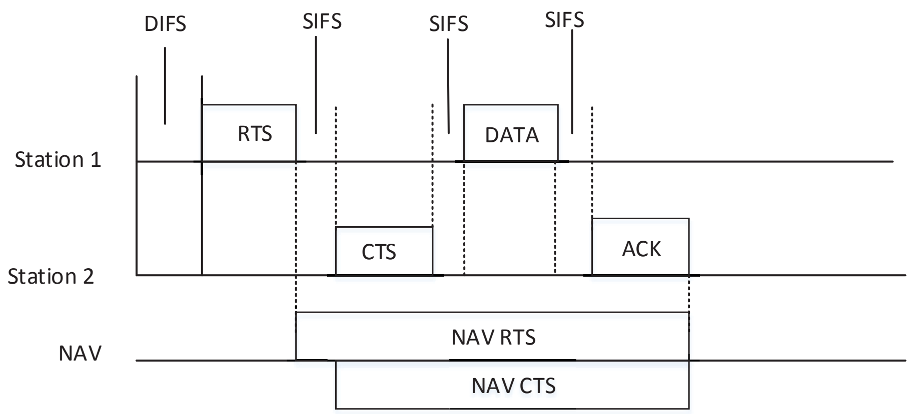


Figure 2.14: RTS-CTS Mechanism

the hidden and exposed node problem. The contention based MAC schemes can be classified into; sender initiated, receiver initiated, synchronized protocol, multiple channel protocol, Quality of Service (QoS) protocols and clustering protocols.

A *Sender initiated protocols*

The multiple access with collision avoidance (MACA) protocol [55] represents an RTS/CTS method without the use of an ACK message. The sender sends an RTS to the receiver. When the neighbouring node hears the RTS message, it stop its transmission. Once the RTS is received, a CTS message is sent back to the sender to initiate data transmission. If the node transmits a packet, the node uses a binary exponential backoff (BEB) algorithm to backoff a random interval of time before trying. The limitation of this protocol is that it does not use the ACK message and, 10 hence, does not provide any acknowledgement of data transmission at the data link layer. If a

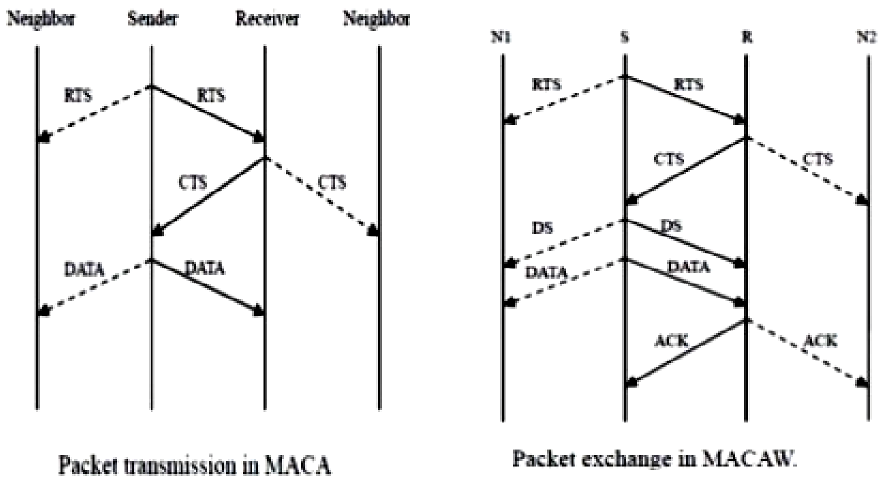


Figure 2.15: MACA Vs MACAW

transmission fails for any reason, re-transmission must be performed by the transport layer, which increases the potential delay.

MACA for Wireless (MACAW) [60] is a modification of MACA. After sending the packets, an ACK message is sent by the receiver to the sender. All the neighbouring nodes overhearing the ACK message know that the packet has been delivered, and the transmission is completed.

Floor acquisition multiple access (FAMA) [104] is also a modification of MACA in which the sending node sends an RTS using either non-persistent packet sensing (NPS) or non-persistent carrier sensing (NCS). The receiver responds with a CTS packet that contains the address of the sender. Stations overhearing this CTS packet recognise the station that has acquired the channel. The CTS packets are repeated long enough for the benefit of any hidden sender.

These protocols show good performance in light load conditions; however, under heavy load conditions, there is a high probability of collision, which can lower the throughput of the network.

Multiple Access with Reduced Handshake (MARCH) [105]is a multi-hop protocol that reduces the number of control messages. RTS/CTS is exchanged during the first hop, and for all the subsequent hops, only CTS messages are used. In MARCH, the MAC layer maintains a table comprising information of the routes and the forwarding nodes in the routes. Due to the low control overhead, collisions are reduced in the heavy load condition. The results show that it can perform better than MACA during heavy load. Many assumptions are proposed in this protocol, including the notions that all the nodes are equidistant from each other and that the transmit power of all ther nodes is the same, even though in reality, this may not always be the case.

B *Receiver Initiated Protocols*

Receiver-initiated protocols attempt to reduce the control message overhead. MACA by invitation (MACA-BI) [65] is a receiver-initiated protocol that reduces the number of control packet exchanges. Instead of a sender waiting to gain access to the channel, MACA-BI requires a receiver to request the sender to send the data by using a ready-toreceive (RTR) packet instead of the RTS and CTS packets. Therefore, this protocol is a two-way exchange (RTRDATA) as opposed to the three-way exchange (RTSCTS- DATA) of MACA. MACA-BI requires the traffic information of the neighbouring nodes, which is present in the DATA message.

Receiver initiated multiple access (RIMA) [106]is a new packet arrival prediction method. RIMA is an improved modification of MACA-BI in which mobile hosts (MH) are assumed to have the same packet arrival rate. When an MH receives a packet, the neighbouring MH also receives a packet and sends an RTR packet to invite the neighbouring MH to start transmitting. These protocols can reduce the control overhead and also show better performance than sender-initiated protocols when traffic is light. The main drawback is the correct prediction of the packet arrival time by the receiver, which is hard to achieve.

A rate adaptive MAC protocol for multi-hop wireless networks (RBAR) [75] presents a rate adaptation mechanism in the receiver instead of a sender. Channel quality estimation is calculated on the receiver side, and rate selection is performed per packet basis during the RTS/CTS exchange, before packet transmission. This scheme cannot be implemented in real wireless networks because it is not compatible with 802.11, as both the data and control packet formats require modification.

C *Synchronized*

A distributed medium access control protocol for synchronised networks (SYN-MAC) [107] proposes a binary countdown scheme to solve unfairness and the hidden terminal problem. Time synchronisation is achieved through GPS receivers. Each sender creates a binary number, and the one with the highest binary number wins access to the channel. This scheme includes three steps: contention resolution, hidden station elimination (used to resolve the hidden station problem and data transmission), and acknowledgement message. SYN-MAC performs well in terms of collision probability, packet delay tolerance, and efficiency.

Z-MAC [108] combines the strengths of TDMA and CSMA while offsetting their weaknesses in a sensor network. The main feature of Z-MAC is its adaptability to the level of contention in the network under low contention. Z-MAC behaves like CSMA and occasionally like TDMA under high contention. Before a node transmits during a slot, a carrier sensing mechanism is implemented and a packet is transmitted when the channel is clear. However, the owner of the slot always has priority over non-owners in accessing the channel. The priority is implemented by adjusting the initial contention window size in such a way that the owners are always given an earlier opportunity to transmit than non-owners. The goal is to reduce collision since owners are given the initial opportunity to transmit. When a slot is not in use by its owners, non-owners can borrow the slot. This priority scheme has the effect of implicitly switching between CSMA and TDMA depending on the level of contention. When the interference and synchronisation become less accurate, performance degrades to CSMA.

D *Multi channel protocols*

One of the earliest papers in this category is Busy Tone Multiple Access Protocol (BTMA) [109] a centralised multichannel protocol that proposes using a separate busy channel to solve the hidden node problem. The transmitting channel is divided into two parts: the busy tone channel and the data channel. A node senses the channel, and if there is no busy tone, it will transmit the busy tone and start sending data. When a node is transmitting, no other nodes within two hops can transmit data.

Dual Busy Tone Multiple Access Protocol (DBTMA) [56] is an extension of BTMA. The transmission channel is 11 split into two channels: the data channel and the control channel (RTS, CTS and busy tone). The two busy tone channels are at different frequencies and are meant for transmission and reception. This scheme uses a distributed approach to solve the hidden terminal problem, in contrast to the centralised approach utilised by BTMA. This scheme is not suitable for many users in small areas, as at any time, only one user can access the data channel.

In [110],a full-duplex spectrum sensing (FD-SS) MAC protocol is proposed for multi-channel non-time-slotted cognitive radio networks, where primary users (PU) randomly access and leave the licensed channels. The FDSS scheme is developed to guarantee high-throughput transmission for primary users (PUs) and high channel utilization for secondary users (SUs) by timely sensing the PUs reactivation during the same time when the SUs are transmitting their signals. However, this scheme is not applicable to hierarchical spectrum access, where PUs should have higher spectrum access priority than SUs.

A distributed multichannel MAC protocol [111] recommends a slow and fast hopping sequence with a dual radio interface to improve network capacity. Transmission is performed on one interface with fast hopping, and reception is performed in another interface with slow hopping. Four components control the protocol:

- Hopping control generates hopping sequence for both slow and fast hopping.
- Discovery sends hello packets to discover the nodes.
- Rendezvous is performed when nodes have slow hopping channel information and ready to send data.
- Broadcast support broadcast packet on both interfaces.

A parallel rendezvous approach requires special coordination between sender-receiver pairs, so they use the same channel to make transmission arrangements. This protocol requires loose synchronisation in the network.

E *QoS MAC Protocols*

Quality of service (QoS) is important and necessary for real-time services such as voice and video. QoS in ad hoc networks is challenging as it means to provide a guarantee and assurance about the level of service given to an application. QoS is different for various applications as it depends on the specific requirements of an application. Some applications are delay sensitive and may require delay guarantees (like voice). Some applications require that the packet should flow at a certain minimum bandwidth, and bandwidth is the QoS parameter [112]. Other applications require reliability, that is, packets are delivered reliably from the source to the destination [112]. Several challenges exist in provisioning QoS in ad hoc networks. Due to dynamically changing topology, QoS may suffer. MAC protocols in ad hoc networks are distributed, which can complicate QoS provisioning. Furthermore, the hidden node problem can degrade the QoS. Ad hoc networks, such as MANET and WSN, have limited bandwidth and battery life, which can also cause adverse effects on the provisioning of QoS. Ad hoc networks are less secure than infrastructure-assisted networks as they implement network access control and have security challenges [112]. QoS is easy to manage in centralised and reservation-based MAC protocols. However, many protocols use a distributed contention-based approach.

In [113], a distributed multi-constrained QoS scheduling algorithm is proposed for delay/throughput-sensitive applications to guarantee constrained QoS. Medium access with delay and throughput constraints priority is given to urgent packets using co-ordinated multi-hop scheduling to maintain end-to-end QoS targets. This approach piggybacks scheduling information onto RTS/DATA packets and uses this information in existing IEEE 802.11 priority backoff schemes to determine the exact schedule. This scheme decreases the end-to-end (ETE) packet delay and achieves low QoS outage probability compared to other protocols. The limitation is that the nodes must monitor all transmitted packets to extract the scheduling information.

Multi-hop access collision avoidance with piggyback reservation (MACA/PR) [114] is an extension of FAMA that proposes guaranteed bandwidth for real-time services through reservations. The RTS/CTS message is used during the first packet to set up the reservations; the remaining packets do not require this message. An ACK message serves to renew the reservation. If ACK is not received due to link failure, retransmission occurs from the network layer. Reserving a significant portion of the channel for a single transmission is not suitable as it may block the nodes from transmitting for a long time

Dynamic Multi-Channel Assignment MAC Protocol (DCA) [66] is designed to increase the throughput and utilisation. DCA implements an on-demand method to assign channels to the host, and the number of channels is independent of topology. No clock synchronisation is required, and few control messages are exchanged to assign channel medium access. Bandwidth is divided into control channels (resolve contention and assign data channels to the mobile hosts) and data channels (send data packets and ACK). Each mobile host maintains a channel usage list (CUL) containing a list of neighbours and a free channel list (FCL). The receiving node matches the CUL with the FCL and makes the sender aware of the free data channels. The spectral efficiency decreases for separate dedicated control radios and channels.

F *Clustering protocols*

Contention protocols can employ clustering strategy to improve network performance. Distributed Queuing Medium Access Control Protocol (DQMAN) [115] consists of a dynamic clustering mechanism and is integrated with a near-optimum distributed queuing medium access control protocol. When a node seizes the channel, it becomes the temporary cluster head of a spontaneous cluster and coordinates peer-to-peer communications between cluster members. Nodes within the range of the cluster head become cluster members. Nodes implicitly belong to a cluster as long as they receive and can decode the control packets from the cluster head. This protocol shows that the clustering strategy performs better than the standard IEEE 802.11 DCF.

There are many challenges in contention protocols: lack of centralised control, network topology changes with mobility, a collision cannot be detected entirely leading to channel inefficiency, unfairness also occurs as nodes with a smaller contention window always win, and quality of service (QoS) needs to be provisioned careful when designing the contention MAC protocol. Table V compares the contention protocols and their proposed schemes, advantages and disadvantages.

Type of Protocol	Classification of Protocols	Proposed Scheme	Problem Addressed	Advantages/Disadvantages
Contention Protocols	Sender Initiated	MACA	Hidden/Expose node problem and Power control	Advantage: No Data collision. Disadvantage: Increases packet delay in case of transmission failure as there is no ACK and decreases in bandwidth utilization
		MACAW	Hidden/Expose node problem	Advantage: Improve back-off mechanism and ensure fair allocation of resources in transmitters. Disadvantages: Does not fully solve hidden node problem and also has longer step exchange between the nodes.
		FAMA [104]	Minimize node Collision	Advantage: Reduces hidden/expose node problem. Disadvantage: Not recommended for heavy load of traffic due to increase of collision.
		MARCH [105]	Hidden/Expose node problem	Advantage: Reduces collision in heavy load and also reduces the handshakes in multi-hop network. High throughput and low control overhead. Disadvantage:Access to routing information is required.
	Receiver Initiated	MACA-BI [65]	Hidden/Expose node problem	Advantage: Reduce control overhead and have better performance than sender initiated under light load. Disadvantage: Difficult to predict packet arrival time.
		RIMA [106]	Hidden node problem	Advantage: Reduce control overhead and have better performance than sender initiated under light load. Disadvantage: Difficult to predict packet arrival time.
		RBAR [75]	Channel quality estimation and Hidden/Expose node problem	Advantage: Estimation is more accurate and can be implemented in IEEE 802.11. Disadvantage: More overheads and routing protocol prefers long reliable links.
	Clustering	DQMAN [115]	Hidden/Expose node problem	Advantages:Performs better in a large network and is efficient

Type of Protocol	Classification of Protocols	Proposed Scheme	Problem	Advantages/Disadvantages
Contention Protocols	Synchronized	SYN-MAC [107]	Unfairness and hidden node problem	Advantages: Perform better in terms of collision probability, packet delay tolerance and efficiency. Disadvantage: Does not use channel efficiently.
		Z-MAC [108]	Hidden node problem/Congestion	Advantages: Reduces Hidden Terminal problem. Less processing and resources are required. Disadvantages: Signalling overhead increased due to TDMA and cannot mitigate the effect of funnelling events to a sink choke point.
	Multi-Channel	BTMA [109]	Hidden/Expose node problem	Advantages: Very simple protocol.Collision probability is very low. Disadvantage: Bandwidth utilization is low
		DBTMA [56]	Hidden/Expose node problem	Advantage: Completely eliminated the hidden node problem. Disadvantage: Not suitable for many users in smaller area as one user can access data channel at a time
		[51]	Network capacity	Advantage: Loose synchronization is required. Disadvantage:Tight coordination is required between sender receiver pairs
	QoS	MACA/PR [114]	Hidden/Expose node problem	Advantage: Global synchronization is not required. Disadvantage: Random empty slots are not utilized efficiently
		DCA [66]	Power control	Advantage: No clock synchronization required and low control overhead. Disadvantage: Decrease in spectral efficiency when few channels are available and also increases the cost.
		[52]	QoS (priority traffic)	Disadvantage: Requires all the nodes to have scheduling information and does not provide backward compatibility

Table 2.7: Comparison Summary of Contention Protocols in Distributed/Ad-hoc Wireless Networks

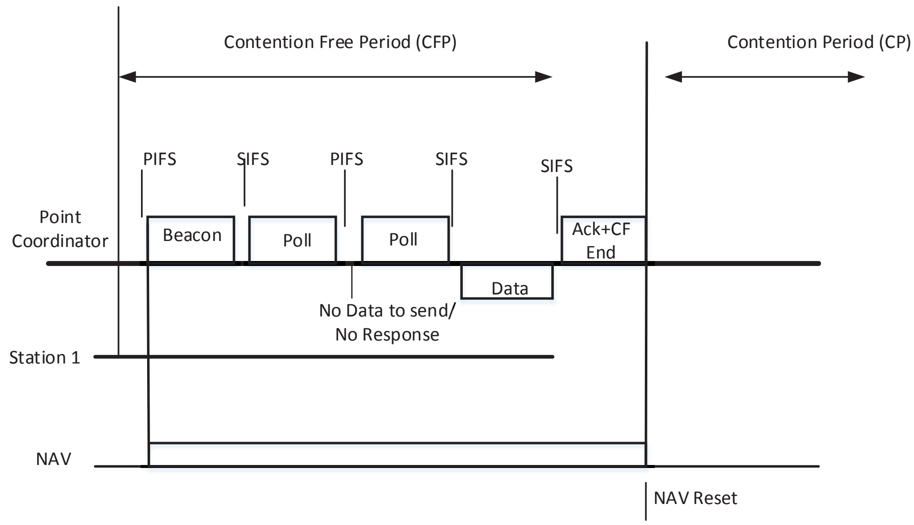


Figure 2.16: PCF Operation Structure During CFP

2.3.4.2 Contention Free MAC Protocols

To overcome the challenges of contention protocols, the contention-free framework has been explored by researchers, and several strategies have been proposed. The fundamental access method for the contention-free protocol is PCF. PCF lies above DCF and has higher priority than DCF. PCF has a point coordinator (PC) that is located at an access point (AP). The point coordinator (PC) creates a polling list and polls the node accordingly. Nodes that have data to transmit do so when they are polled.

During the contention-free period (CFP), the PC delivers frames to the station and polls the previously registered stations in the polling lists. The PC ensures that the interval between two polling stations is not more than the PCF interframe spacing (PIFS) [116–118]. PCF frame transmission is shown in Figure. 2.16.

The previous literature on contention-free MAC protocols can be categorised into polling MAC protocols, TDMA-based polling protocols and quality of service protocols.

A *Polling MAC Protocols:*

Distributed Deficit Round Robin (DDRR) MAC protocol [119]proposes a distributed fair queuing scheme with a deficit counter (DC) to each station in the polling list. This scheme was proposed to allocate bandwidth fairly to voice and video with different packet sizes. Only nodes with positive DC are polled, and for every poll, the DC is decremented by the length of the packet. The node will no longer be polled when the DC is negative. The disadvantage is that during heavy loads, access delay is not guaranteed.

Cyclic Shift and Station Removal (CSSR) MAC protocol [120] uses a discrete-time Markov chain and improves the performance when silence detection is used at wireless terminals. The AP polls only active stations and removes them from the polling list if they are silent by sending a null frame. After some time, the idle station is returned to the active polling list. The disadvantage is that voice packets can be lost if talk-spurt starts when the node is idle.

Earliest Deadline First (EDF) MAC protocol [121] caters to the different delay requirements of real-time services in heterogeneous networks. It supports real-time and non-real-time services in WLAN with dynamic timedivision duplexed (D-TDD) transmission. The packet with the earliest deadline is polled first. This scheme works well in a heterogeneous environment, but the limitation is that the proposed admission control works for constant bit rate traffic only.

Simultaneous Transmit Response Polling (STRP) MAC protocol [122] enables simultaneous polling and transmission of information packets using the capture effect. Nodes are divided into logical rings as active rings and idle rings. Any node that has data to transmit will enter into the active ring and will be polled in a round robin fashion. The capture effect can help the point coordinator (PC) to receive and interpret a strong signal regarding a weak signal. This scheme can increase the delay when transferring from the active ring to the idle ring.

Two Step Multi-polling (TS-MP) MAC protocol [123] proposes a TXOP allocation with two polling frames to reduce the overhead. The first frame gathers information (pending frames, transmission rates), and the second frame contains the polling scheme based on the first frame. CBR (constant bit rate) and VBR (variable bit rate) real-time traffic are taken into account. This scheme achieves better performance than CP-MP (contention period multi-polling) MAC protocols, but the status collection process increases the delay.

Relay-Enabled Point Coordination (rPCF) MAC protocol [124] minimises the control overhead by improving the performance. This scheme transmits data using a relay by taking into account the channel condition of each link. If the channel conditions do not satisfy the requirements, the relay is used to transmit data. Each node maintains a table (containing the data rate and sender ID) and updates the AP if any changes occur in the table. Similarly, the AP maintains a table containing the data rate and the sender and receiver IDs. The AP piggybacks the selected transmission rate to the node to update it, and the AP informs the sender if it has to transmit the data via one or two hops. The modification is made in the MAC header. The results have been compared with PCF-based ARF using NS-2, indicating that the proposed scheme performs better.

GreenPoll MAC protocol [125] is an energy saving protocol that turns off the radio once the data have been exchanged. The contention-free period is split into two virtual phases. In the first phase, nodes with no data to transmit go to sleep, whereas nodes that need to transmit data remain awake. Once the awake stations transmit data, they go to sleep until the cycle is completed or the NAV timer expires. The last node never goes to sleep, so it must remain awake during the entire cycle, which can increase the overall energy consumption of the network.

B *TDMA based Polling MAC protocol*

Isochronous coordination function (ICF) MAC protocol [126] reduces the control overhead and provides fair polling for uplink and downlink voice traffic. ICF emulates a dynamic TDMA-like mechanism, reduces the polling overhead and provides fair polling for uplink and downlink voice traffic. ICF suggests a TDMA-like time slot assignment that is broadcast through a super ICF-poll frame at the start of each ICF cycle. The ICF super poll frame, which is a TDMA-like time slot, is transmitted, thereby substantially reducing the polling overhead. A Markov model is used for voice traffic in the proposed protocol, which is compared with other polling-based protocols. However, the protocol does not provide the explicit contention-free medium that is required for voice traffic.

In [71], energy is saved via the awake and sleep periods implemented using a clustering strategy. Nodes are polled by the cluster head based in round robin order. Each node with data to transmit is assigned a time slot (TS). Nodes that do not have data to transmit go into sleep mode for a certain amount of time. The proposed model is compared with S-MAC using NS-2. This scheme saves energy and has less delay than the SMAC protocol.

Polling and reservation access scheme (PRAS) MAC protocol [127] is a centralised protocol using both reservation and polling techniques to assign transmission slots. This protocol is proposed for priority scheme ATM service using a piggyback mechanism. The piggybacking overhead should be minimised, and the polling period should be efficiently adjusted to reduce bandwidth waste.

Priority ELF MAC protocol [128] proposes wireless scheduling based on an effort-outcome disconnection and effort-limited fairness (ELF) strategy using dynamic weight adjustments. The ELF guarantees that all flows with an error rate below a threshold receive their expected service. All flows experiencing an error rate below a per-flow threshold receive their expected service. During the PCF duration, the point coordinator (PC) checks the station with the highest type of service (TOS) and allows the highest TOS station to transmit or receive data first. The point coordinator (PC) also checks the values of the counters, which record some frames transmitted by each station in a particular period, and whether each counter is larger than one. If not, the PC will not transmit data or poll frame to a station even though the PC have data to send in their queues or frames. In this protocol, it is impractical to guarantee identical throughputs to each user over short time scales.

The combination of reservation and polling multiple access with distributed scheduling (CRPMA/DS) MAC protocol [129] proposes a distributed scheduling architecture for ATM services. Different frequencies are assigned for uplink and downlink as the protocol works on a TDMAFDD principle. The BS allocates the bandwidth to users, and users decide what kind of traffic to transmit. This scheme reduces collisions by applying the contention resolution algorithm and also serves data traffic with a high transfer rate. The disadvantage in reserving the bandwidth is that it is wasted if there is no traffic. In [130], the authors propose non pre-emptive priority scheme to transfer voice packets efficiently. The AP dynamically creates the polling list by assigning a priority to each voice packet based on two parameters: packet transfer rate and jitter. When no voice packet is transmitted, the protocol changes to DCF mode to transfer data packets.

Polling MAC protocols also face some challenges, including unpredictable beacon delays, unknown transmission duration of polled stations and lack of suitability for small networks with few users, as the delay will increase without improving the network performance. A comparison summary for contention free MAC protocols is shown in Table.2.8.

Type of Protocol	Classification of Protocols	Proposed Scheme	Problem	Advantages/Disadvantages
Contention Free Protocols	Polling	CSSR [120]	Scheduling Polling list	Advantage:Reduce wastage of polling frames. Disadvantage:Voice packet are lost if talk-spurt is on in an idle mode. Takes longer time to connect to the node.
		DDRR [119]	Fair Bandwidth allocation	Advantage: Suitable for heterogeneous network with real time traffic. Disadvantage: Does not measure access delay for heavy load networks.
		EDF [121]	Access Delay for real time traffic	Advantage: Suitable for heterogeneous real time traffic. Disadvantage: Admission control only work for constant bit rate traffic
		STRP [122]	Schedule Polling	Advantage: Taking in account capture phenomena which increases the efficiency. Disadvantage: Increase delay.
		TS-MP [123]	Scheduling Polling list	Advantage: Shows better performance. Disadvantage:Overhead is increased.
		rPCF [124]	Reduce Transmission delay and improve Throughput	Advantage: Improve system performance. Disadvantage: Overhead is increased
		GreenPoll [125]	Energy Consumption	Advantage: Conserve energy of the user and Wlan. Disadvantage: Performance degrades as last node does not sleep and conserve energy.
		[64]	Scheduling Polling list	Advantage: Improved adaptive polling discipline. Disadvantage:High polling rate can cause high collision and too law may result in high delay
	TDMA based Polling	ICF [126]	Scheduling Polling list	Advantage: Reduces polling overhead and provides fair polling. Disadvantage:Does not provide explicit contention free environment which is required for voice traffic.
		[65]	Energy Consumption	Advantage: Less delay and energy is conserved
		PRAS [127]	QoS	Disadvantage: Bandwidth is wasted if not utilized efficiently and piggybacking overhead should be reduced
	QoS Polling	Priority ELF [128]	QoS	Disadvantage: Not practical to guarantee identical throughputs to each user over short period of time
		CRPMA/DS [129]	QoS	Advantage: Reduces collision. Disadvantage: Reserved bandwidth is wasted if there is no traffic

Table 2.8: Comparison Summary of Contention Free Protocols in Distributed/Ad-hoc Wireless Networks

2.3.4.3 Channelization MAC Protocol

The research on time division medium access (TDMA)-based MAC spans decades. TDMA uses a pre-computed schedule to transmit packets and can solve the hidden terminal problem without additional message overhead. The TDMA technique can provide low latency and guaranteed bandwidth [131, 132].

In [133], the authors propose a MAC protocol that minimises the end-to-end delay and improves the bandwidth utilisation using QoS parameters. A shorter TDMA frame is proposed to reduce the delay, and scheduling is performed in a distributive manner. The delay-aware link scheduling problem is solved in two parts. A TDMA algorithm is developed to find conflict-free TDMA class, and a polynomial algorithm is developed to find the transmission order to minimise the endto- end delay. This protocol does not take into account burst traffic, and delays are neglected.

Adaptive MAC Protocol for Heterogeneous networks (AMPH) [134] is proposed with QoS support to provide high channel utilisation with hybrid and adaptive behaviour for good data delivery of heterogeneous traffic. Nodes can transmit in any time slot; each node is assigned a TS in such a way that two nodes within two hops cannot have the same TS. Priority is given to the owners of the slot and to higher priority traffic. Lower priority packets are not considered and can be lost.

Five Phase Reservation Protocol (FPRP) [70] proposes a distributed MAC protocol for securing temporary channel access before transmission using inband signalling. Each slot is divided into information and reservation slots. Each node that needs to reserve the information slot should contend for it during the reservation slot. The reservation slot has five phases. The first four phases are used for reservation to eliminate the hidden node problem, and the fifth phase reuses the same slot efficiently within two hops. A node that reserves an information slot can transmit with a low chance of collision during the slot. In this protocol, the primary complexity lies in establishing reservations that do not conflict within twohop neighbours. Additionally, due to inband signalling, this protocol suffers from unused time slots when signals collide because of their randomness.

The Node Activation Multiple Access (NAMA) [135] proposes a distributed election algorithm to achieve collision-free transmission. For each slot, one transmitter is selected in a two-hop neighbourhood. Collision-free data are received by the onehop neighbour. Energy conservation is not taken into account in NAMA.

Traffic Adaptive Medium Access (TRAMA) [136] is a modification of NAMA that increases the utilisation of energy in an efficient manner. TRAMA uses a distributed election scheme based on the traffic information at each node to determine which nodes can transmit in a particular time slot. Nodes that do not have data to transmit are switched into idle mode. The disadvantage of this method is that the delays are high compared to contention-based protocols because the sleep time is high.

Distributed RAND (DRAND) [137] proposes a channel reuse scheduling algorithm that avoids collisions within a twohop neighbourhood. Once a slot is assigned, each node reuses its slot periodically in each frame. A node allocated to a slot is the owner, and other nodes are non-owners. Each slot can have more than one owner as DRAND allows any two nodes beyond a two-hop neighbourhood to own the same slot. This scheme does not require any synchronisation and can adapt to local topology changes. DRAND is most advantageous when there is less mobility. The disadvantage is that it does not provide any method to change or repair schedules.

TDMA-based MAC protocols also face some challenges: tight synchronisation is required, and they create interference at a frequency that is directly related to the time slot length. Table. 2.9 compares the channelization MAC protocols

Type of Protocol	**Classification of Protocols**	**Proposed Scheme**	**Problem**	**Advantages/Disadvantages**
Channelization Protocol	TDMA	AMPH [134]	QoS	Advantage: Provides fair data delivery for higher priority packets. Disadvantage: Lower priority packets are lost.
		[69]	End to End delay	Advantage: Frame size has been reduced and scheduling is distributive. Disadvantage: Not recommended for bursty traffic and delay is not taken into account.
		FPRP [70]	Scheduling	Advantages: Hidden terminal problem is handled efficiently. Suitable for scalable and dynamic networks. Disadvantages: It suffers from unused time slots when signals collide because of their randomness. Also the overhead is increased
		NAMA [135]	Scheduling	Advantage: Receive collision free data. Disadvantage: No energy conservation is taken into account
		TRAMA [136]	Throughput and Energy Efficiency	Advantages: Reduces collision and increase energy savings. Disadvantage: Length of the random access interval affects the overall duty cycle and energy savings
		DRAND [137]	Scheduling	Advantages: No synchronization required. Easily adaptable to the topology change. Disadvantages:Schedules can not be modified.

Table 2.9: Channelization Protocols Comparison Summary in Distributed/Ad-hoc Wireless Networks

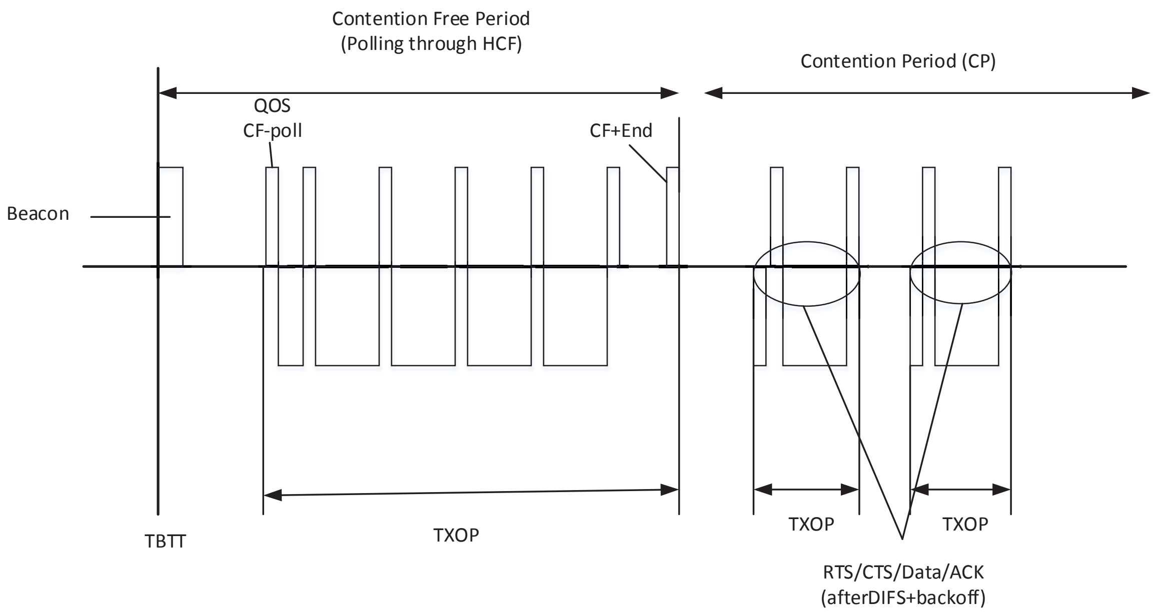


Figure 2.17: HCF Mechanism

2.3.4.4 Hybrid Coordination Function

The hybrid coordination function (HCF) is an improved amendment of DCF that defines a set of quality of service (QoS) enhancements for wireless LAN applications through modifications to the MAC layer [138]. HCF enables transmission and reception in both the CFP and CP. HCF consists of two parts: enhance distributed coordination function (EDCF) for contention-based access and controlled access (HCCA) for contention-free access. Both HCCA and EDCF define traffic classes (TC) and traffic streams (TS), for example, internet browsing can be given lower priority and voice over WLAN can be given higher priority. Figure. 2.17 shows the mechanism for HCF.

Minimal work has been performed using the hybrid coordination function (HCF). Further attention is needed, as it can be used to design a better MAC protocol where QoS is required.

A *Hybrid MAC Protocols*

Adaptive Enhanced Distributed Coordination Function (AEDCF) [139]is built on top of the EDCF. The contention window size is adjusted by taking into account the network and application requirements. AEDCF does not reset the contention window size to a minimum but takes into account the collision rate in each station. This scheme outperforms EDCF by reducing collision rates for highly loaded networks.

Adaptive Fair Enhanced Distributed Coordination Function (AFEDCF) [140]extends EDCF to further improve network performance. The contention window size increases during the deferring period when the channel is busy, and an adaptive backoff mechanism is used when the channel is idle. AFEDCF reduces the idle time slots through the adaptive backoff mechanism by taking in account the channel load. This method increases the quality, especially for multimedia applications.

In [141], throughput and channel access delay are improved by applying a delay-aware distributed dynamic adaptation for the contention window. This scheme is known as D2D, and the medium access protocol used is D2D channel access (D2DCA). The protocol is a modified version of EDCA, where only the backoff algorithm for collision avoidance is modified. The D2D mechanism takes into account the delay deviation ratio and the channel busyness ratio. The idea of D2D is that if a station experiences adequate channel access delay, the station should not increase the channel access delay of other stations by blindly resetting the CW to its minimum value. The D2D scheme is adaptive to changes in congestion level. The results show that the scheme outperforms EDCA.

Hybrid MAC protocols also face some challenges. The admission decision should be made at different active stations rather than the admitting station, increasing the flow throughput beyond a certain threshold results in en-queuing delays, the protocols are characterised by high randomness, and fairness is difficult to guarantee [142].

Type of Protocol	**Classification of Protocols**	**Proposed Scheme**	**Problem**	**Advantages/Disadvantages**
HCF	QoS	AEDCF	QoS	Advantage: Collision rate is reduced for highly loaded network. Disadvantage: Not suitable for intro AC QoS differentiation
		AFEDCF	QoS	Advantage: Reduces idle TS. Disadvantage: Not suitable for intro AC QoS differentiation
		D2DCA	QoS	Advantage: Performs better than EDCA.

Table 2.10: Summary Comparison of Hybrid MAC Protocols in Distributed/Ad-hoc Wireless Networks

2.4 Conclusion

A comprehensive literature review was undertaken explaining the two important features of 5G networks that includes; Device-to-Device communication and heterogeneous networks. Resource allocation methods for D2D communication are explained in detail in infrastructure assisted network. Different types of MAC protocol for ad-hoc networks are discussed with their advantages and disadvantages.

The research focuses on resource allocation of D2D communication in heterogeneous networks. The research proposes a Scalable MAC Protocol (SC-MP) based on point coordination function (PCF), that offloads the traffic from licensed band to unlicensed band. The proposed MAC protocol increases the network capacity and network performance as compared to traditional DCF MAC protocol. The importance of PCF access mechanism is that operates in centralized manner and highly suitable for dense environment as it avoids collision. Hence, SC-MP creates a centralized control in a distributive manner. The proposed scheme can be applied to to IEEE 802.11 devices without requiring significant modification to the physical layer.

Chapter 3

Performance Investigation of D2D Communication over LTE Solutions Operating in Unlicensed Band

3.1 Introduction

The usage of unlicensed spectrum has provided an excellent opportunity for the mobile operators to meet the data requirement of mobile users. Mobile operators can expand the network capacity and provide a better quality of service (QoS) by using LTE in unlicensed 5GHz band at a lower cost [143].

Device to Device (D2D) communication an integral part of the future 5G standard, aims to improve the capacity, spectral efficiency, and coverage by enabling direct communication between the devices with minimal involvement from the base station (BS) in infrastructure assisted cellular system. D2D communication has been a central topic of research for quite some time and it has been shown that D2D communication increases the network throughput while minimizing the interference and power usage [144]. D2D communication can be very beneficial for the mobile operators when they combine D2D with LTE technologies operating in unlicensed 5GHz band.

This chapter evaluates a detail overview of LTE technologies operating in an unlicensed 5GHz band that includes; carrier Wi-Fi, LTE-U, LTE-LAA, LWA, and MuLTEfire. A detail overview of these technologies are discussed that includes; architecture, working mechanism, co-existence with Wi-Fi and overall cost. A comparison is done between these technologies and further their advantages and disadvantages are concluded.

Furthermore, this chapter also present a detail simulation for LTE-LAA, LWA and MuLTE-fire technologies in the presence of Wi-Fi hotspot and compare their results. The performance of D2D communication over these strategies are also evaluated.

3.2 Background Of LTE Operating in Unlicensed Band

The concept of LTE in the unlicensed spectrum was introduced first time in the TV white space [145]. In this scheme frequency hopping and time hopping was used by LTE small cells in TV white space band to reduce interference between other devices in the band. Currently, the crucial issue for LTE networks to exploit 5GHz unlicensed spectrum is the coexistence problem with Wi-Fi technology [146]. LTE adopts scheduling based access technology in licensed band, whereas Wi-Fi works on contention-based access mechanism and carrier sensing technology. The LTE unlicensed proposal has been a big concern for the Wi-Fi vendors and service providers, which have used 5GHz band for a long time and want to continue to do so without being unfairly penalized by the introduction of LTE [143].

Recently new types of LTE-Wi-Fi aggregation solutions have been proposed at the radio link, as well at the TCP link [147]. One is carrier Wi-Fi, in which network operators can deploy their own Wi-Fi to offload traffic and reduce congestion [148]. LTE unlicensed (LTE-U), is the first version of LTE unlicensed which was initially proposed by Ericsson and Qualcomm in 2013, based on 3GPP release 10-12 [143]. LTE-U works on the mechanism of carrier sensing adaptive transmission (CSAT) and can only be used in China, USA, South Korea and India [149]. For worldwide deployment, LTE unlicensed needs to deploy Listen-Before-Talk (LBT) mechanism. LTE-Licensed Assisted Access (LTE-LAA) has been standardised by 3GPP in release 13 for downlink operation, which uses the mechanism of LBT and is the modified version of LTE-U [150]. It is believed that if LTE-LAA and Wi-Fi use the same spectrum, LTE-LAA will have more chance to dominate the spectrum and Wi-Fi devices will keep deferring to LTE-LAA transmissions [147]. An alternative solution was proposed to LTE-U/LTE-LAA, called as LTE Wi-Fi link aggregation (LWA). A part of LWA is released by 3GPP in release 13 [151]. LWA increases the capacity by offloading some of the LTE traffic through Wi-Fi using CSMA protocol. Currently, Qualcomm proposed another LTE based technology known as Multifire that solely operates in an unlicensed band and does not require an anchor in the licensed spectrum. [152].

A detail architecture and comparison of LTE solutions in unlicensed band is explained in the next section.

3.3 State Of Art: LTE Networks Operating in Unlicensed Bands

In this section, the different solutions are investigated provided for LTE operating in an unlicensed 5GHz band. Their architecture, working mechanism, cost and co-existence problem with Wi-Fi are discussed. Furthermore, a comparison is shown for these technologies with their advantages and disadvantages.

3.3.1 Carrier Wi-Fi

Carrier Wi-Fi is considered to be the first step for mobile operators to utilize the unlicensed bands. Network operators can deploy their own Wi-Fi access points (AP) to further offload traffic, increase coverage and reduce congestion.

Carrier Wi-Fi can lead to ineffectiveness of network management and low spectrum efficiency as it adopts different access and management mechanisms from LTE networks [148].

3.3.2 LTE-Unlicensed (LTE-U)

LTE-U uses the same LTE channels as LTE CA but unlike LTE CA LTE-U operates in unlicensed band at 5GHz [147]. In LTE-U, only LTE channel can work as the primary channel whereas the unlicensed channel can work as the secondary channel. LTE-U works on the mechanism of carrier sensing adaptive transmission (CSAT), where the channel is sensed by the LTE devices for a longer period of time (up to 200ms) and based on the activities define an ON and OFF duration for a duty cycle [148]. LTE devices can transmit when the duty cycle is in ON mode and remains silent when the duty cycle is in OFF mode. In the OFF mode, LTE-U allows the Wi-Fi users to access the network using unlicensed band.

LTE-U requires a new installation of 5 GHz LTE-enabled hardware on the user device and small cells. LTE-U will only be implemented in the regions where regulation does not require LBT, such as China, Korea, India, and USA [149]. The biggest issue of LTE-U is its intrusion to Wi-Fi, as it can block the access to Wi-Fi users by giving priority to LTE-U users.

3.3.3 LTE-Licensed Assisted Access (LTE-LAA)

LTE-LAA is the modified version of LTE-U and adopts LBT mechanism for the coexistence of LTE-LAA and Wi-Fi, as shown in Figure. 6.1 (a). In LBT, LTE device first listens to

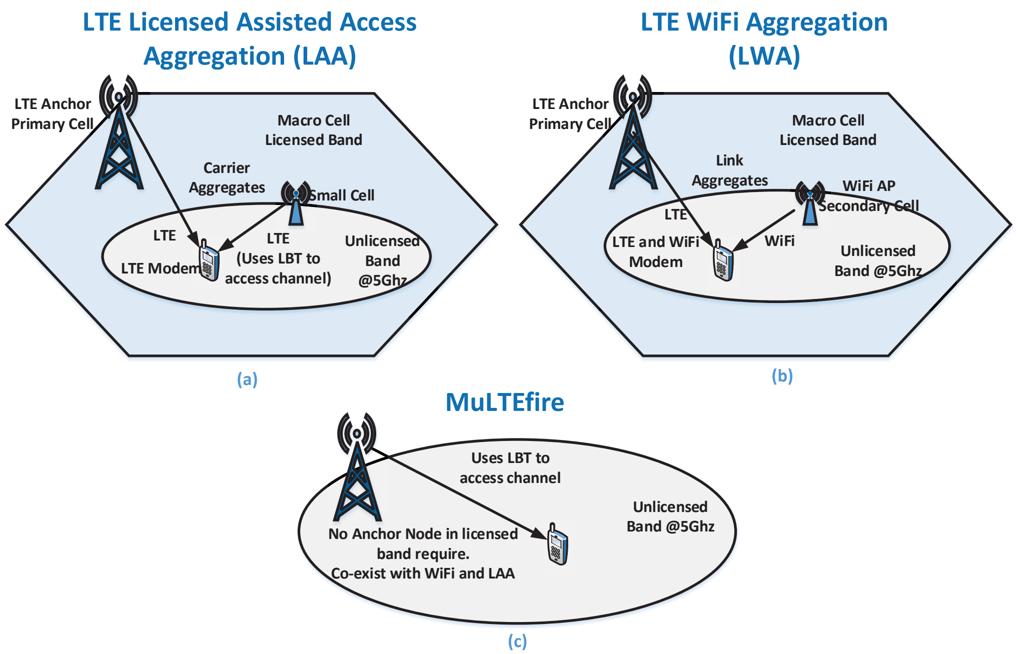


Figure 3.1: LTE in unlicensed bands: LWA, LAA, MuLTEfire

check for ongoing transmission. If the channel is clear it will transmit for a period of time and then back-off to re-check the channel availability, but if the channel is busy it will not transmit and keep on listening unless the channel is available [148]. The LTE device examine another channel if the current channel is busy after several attempts [148].

Two options for LBT schemes have been suggested by the European Union; Frame-Based Equipment (FBE) and Load-Based Equipment (LBE) [147]. In FBE, the transmit/receive structure is not directly demand-driven but has fixed timing. After every frame, clear channel assessment (CCA) is checked if the channel is free the data is transmitted otherwise it has to wait for another frame period. In LBE LBT, CCA is performed whenever there is data to transmit. If the channel is available data is transmitted but if the channel is busy it will retransmit the data after the back-off time during extended CCA (eCCA) [147].

LTE-LAA is developed with the single global solution framework and is replacing the current terminology LTE-U [153]. LTE-LAA is the version of LTE in an unlicensed band that 3GPP standardizes in Release 13 and is set to become a global standard as it strives to meet regulatory requirements worldwide [149]. As LTE-LAA is the modified version of LTE-U, it also requires an additional installation of 5GHz LTE-enabled hardware.

The biggest concern for LTE-LAA is the co-existence with Wi-Fi technology. Qualcomm and Ericsson tested and claimed that LTE-LAA operates with little interference with Wi-Fi and can increase the system capacity with faster mobile broadband speed [154]. Whereas, some of the companies like Google are not in favour of LAA as it can block the access to Wi-Fi users and will degrade the overall network throughput of Wi-Fi by giving priority to LTE-LAA users [155].

3.3.4 LTE Wi-Fi Link Aggregation (LWA)

LWA emerged as an alternative technology to LTE-U and LTE-LAA and is capable of leveraging existing Wi-Fi access points (AP) to improve the network performance [148]. LWA architecture is shown in Figure. 6.1 (b).

For the transmission of LTE traffic, LWA uses unlicensed band similar to LTE-U and LTE-LAA but the transmission is done through Wi-Fi. Unlike LTE-U/LAA, LWA does not require a new LBT based protocol, but LWA uses the current Wi-Fi protocol to transmit the LTE traffic. LWA base station performs scheduling of packets at PDCP layer and transmits some over LTE and other on Wi-Fi after encapsulating in Wi-Fi frames. All the packets received from LTE and Wi-Fi are then aggregated at PDCP layer of LWA UE. LWA base stations can improve the LTE performance by managing the radio resources according to the load and RF conditions [147]. During this transmission, Wi-Fi APs can use LTE core network for authentication, security, billing, etc, without a dedicated Gateway (GW) and without disturbing the native Wi-Fi APs.

LWA uses LTE on LTE band and Wi-Fi on Wi-Fi band which is not the case in LTE-U/LTE-LAA. LWA requires the deployment of small cells and other Wi-Fi APs near it can get a software upgrade to support LWA. The Wi-Fi APs can also support non-LWA traffic by using separate service set identifier (SSID) [148]. Some researchers believe that LWA will perform better in terms of co-existence than LTE-U/LTE-LAA in the presence of Wi-Fi as it offloads the traffic using the same protocol as Wi-Fi, hence sharing the spectrum fairly [147]. Some researchers also believe that LWA is a solution that leverage's the existing Wi-Fi APs and does not impact on an unlicensed band, hence improving the network performance [147]. Also Wi-Fi companies are much in favour of LWA as it boosts the LTE user data without affecting pre-existing Wi-Fi users. Alcatel-Lucent is collaborating with Qualcomm to demonstrate LWA at Mobile World Congress [156].

3.3.5 MuLTEfire

MuLTEfire is based on 3GPP LTE-LAA and LTE-eLAA (LTE-enhanced licensed Assisted Access). Figure. 6.1 (c) shows the architecture for MuLTEfire. MuLTEfire performs like LTE-LAA and LTE-eLAA only the difference is that it operates in an unlicensed band without a licensed anchor node.

According to Qualcomm, MuLTEFire will benefit mobile operators with new deployment opportunities for offloading their mobile network traffic [152]. The primary goal for MuLTEfire is to provide a seamless user experience in a hyper-dense network by combining the benefits of LTE technology with the simplicity of Wi-Fi like deployments. It is a solution that may be attractive to network operators and cable operators that lack licensed spectrum, although this mode has not been discussed in 3GPP [149].

A comparison among the different LTE technologies operating in an unlicensed band; Carrier Wi-Fi, LTE-U, LTE-LAA, LWA, and MuLTEfire is shown in Table. 3.1. It is easier to deploy carrier Wi-Fi as the network operator can easily deploy their own Wi-Fi to offload traffic but the spectrum efficiency cannot be same as of LTE as Wi-Fi and LTE works on different mechanism. LTE-U/LTE-LAA carrier aggregates the licensed and unlicensed band to further improve the performance, whereas LWA aggregates the link of licensed and unlicensed band and MuLTEfire solely operates in unlicensed band. The major disadvantage of LTE-U, LTE-LAA and MuLTEfire is the contention problem with Wi-Fi as compared to LWA. LTE-LAA and MuLTEfire are still in the process of getting standardised.

	Carrier Wi-Fi	**LTE-U**	**LTE-LAA**	**LWA**	**MuLTEfire**
Features	Deploy operator own Wi-Fi	LTE operates in unlicensed 5GHz band using CSAT	LTE operates in unlicensed 5GHz band using LBT	LTE operates in unlicensed 5GHz band aggregated with carrier Wi-Fi	LTE only operates in unlicensed 5GHz band using LBT
Advantage	Easy to deploy	Unified management mechansim	Unified management mechanism. Global standard.	Easy to implement. Quick for commercialization	Unified management mechansim in unlicensed band
Disadvantage	Lack the performance benefits from LTE in unlicensed band	Contention problem with Wi-Fi. Applicable where CSAT mechanism is applied. No global standardisation	Contention problem with Wi-Fi. Commercialization will take longer time	For better QoS latency of the link between the LTE eNB and Wi-Fi AP should be kept low.	Contention Problem with Wi-Fi. Need to replace or upgrade old Wi-Fi AP to get LTE performance.
Access Network Cost	Medium (New Wi-Fi AP)	High (New LTE enabled cell with LTE-U)	High (New LTE enabled cell with LTE-LAA)	Medium (New small Cell LWA aware Wi-Fi AP)	High (New LTE enabled Cell)
New Hardware Support	No	UE and eNB (5GHz)	UE and eNB (5GHz)	No	UE and eNB (5GHz)
LTE-Wi-Fi Coexistence Problem	No	Contention Problem	Contention Problem	No	Contention Problem
Global Problem	No	Global Harmonization	Required	No	Global Harmonization Required
Standardisation	3GPP	LTE-U forum/3GPP	3GPP	3GPP	3GPP

Table 3.1: Comparison between Carrier Wi-Fi, LTE-U, LTE-LAA, LWA and MuLTEfire

3.4 Simulation

In the simulation, a single LTE macro cell is considered with small cell deployed in it. This small cell is LTE-LAA/LWA/MuLTEfire, taking one at a time for simulation. A single Wi-Fi AP is also deployed in the same LTE macro cell whose coverage area overlaps with the same small cell; LTE-LAA/LWA/MuLTEfire. The small cell (LTE-LAA/LWA/MuLTEfire) and Wi-Fi share the same 20 MHz unlicensed sub-band within the 5 GHz band. LTE system with a maximum capacity of 100 Mbps and WLAN with the capacity of 54 Mbps. The ITU-UMI model is used to generate the channels between the users and the Wi-Fi BS. The number of Wi-Fi users are fixed and are associated with Wi-Fi AP. It is assumed that the number of LTE-LAA/LWA/MuLTEfire users are randomly distributed around the small cell in a 250 m × 250 m area. Hand-overs are not considered here. Total network throughput is the aggregated throughput of the users operating in the licensed band as well as in the unlicensed band and the offered load is the total number of users in the licensed and unlicensed band.

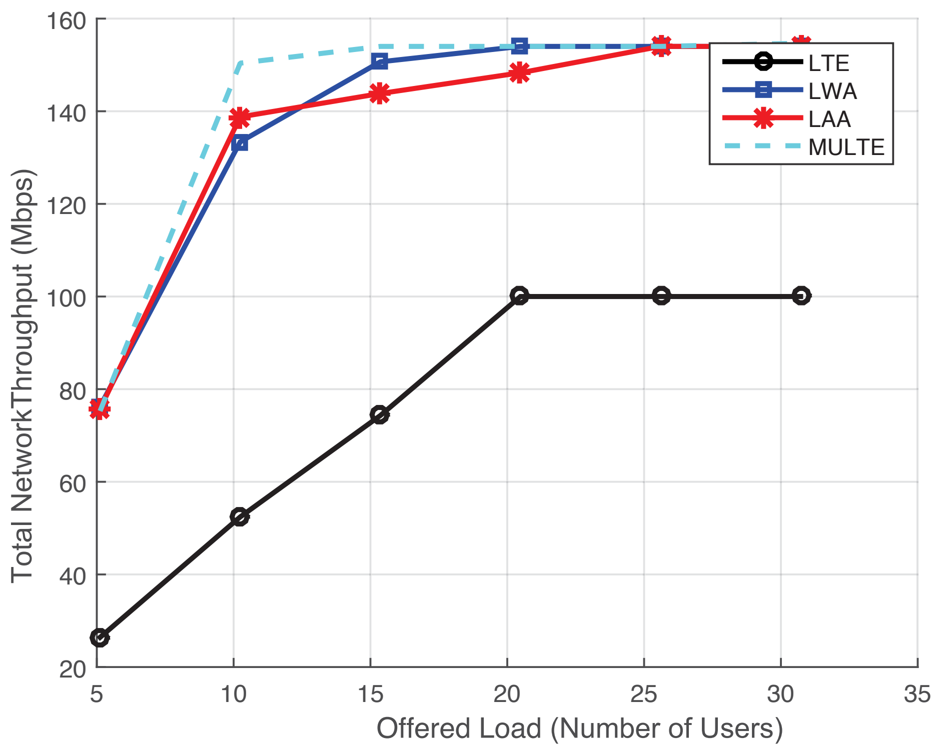


Figure 3.2: Total Network Throughput vs Offered Load

Total network throughput graph versus offered load is shown in Figure 3.2. LTE users only operates in licensed band and the throughput saturates quickly at around 20 users, whereas

MuLTEfire operates in unlicensed or shared spectrum and has the highest throughput and has more chance to capture the channel and transmit the data. MuLTEfire has the highest throughput but saturates quickly as can be seen from the graph and might not be suitable for a scalable network as it can not accommodate more users. Whereas it can be observed from the graph that LWA can provide higher throughput when the users increases as compared to LTE-LAA. LTE-LAA and MuLTEfire can be a better solution for small loads as the throughput is higher whereas for higher load LWA can perform better.

Figure. 3.3 shows the effect of Wi-Fi users in the presence of a small cell; LTE-LAA, LWA, and MuLTEfire. The throughput of Wi-Fi users degrades in the presence of LTE-LAA small cell, as preference is given to LTE-LAA users over Wi-Fi users. This is because LTE has a continuously as well as a periodically transmitting protocol to transfer a variety of control and reference signals whereas Wi-Fi is designed to coexist with other technologies through random backoff and channel sensing. Due to this Wi-Fi users will have little chance to sense a clear channel and transmit. Similarly, in case of MuLTEfire, the Wi-Fi throughput degrades further as it is LTE that is operating solely in an unlicensed band. Whereas, in case of LWA the Wi-Fi throughput is does not degrade much as there is no contention problem between the Wi-Fi and LWA users and can share the spectrum fairly.

Figure. 3.4 illustrates the delay graph vs the number of users. It can be seen from the graph that LTE-LAA has higher delay than LWA. Due to the mechanism of LBT in LTE-LAA, it can increase the delay when contending for the channel in the presence of Wi-Fi. Whereas MuLTEfire has minimum delay as it solely operates in unlicensed band.

To further extend the idea D2D communication is applied in the given scenario. The behaviour of D2D communication is investigated when applied to LTE-LAA, LWA and MuLTEfire. It is observed from the Figure. 3.5 that D2D communication in MuLTEfire can increase the throughput drastically as it operates only in unlicensed band and has more chance for the users to communicate through D2D communication. But it can also be observed that

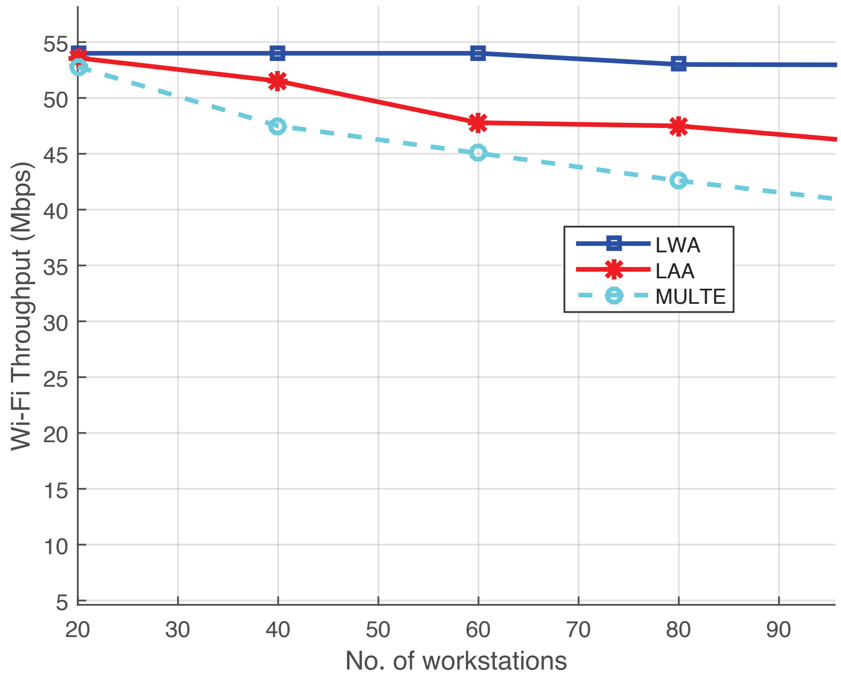


Figure 3.3: Effect of Wi-Fi Throughput vs Number of Workstations

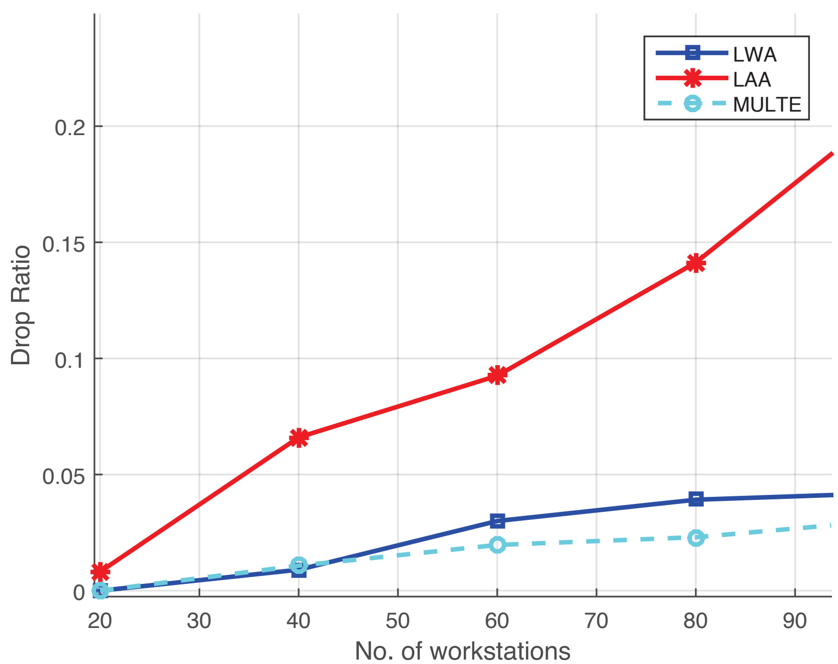


Figure 3.4: Delay vs Number of workstation

network saturates quickly when using D2D communication in MuLTEfire. Whereas, if we

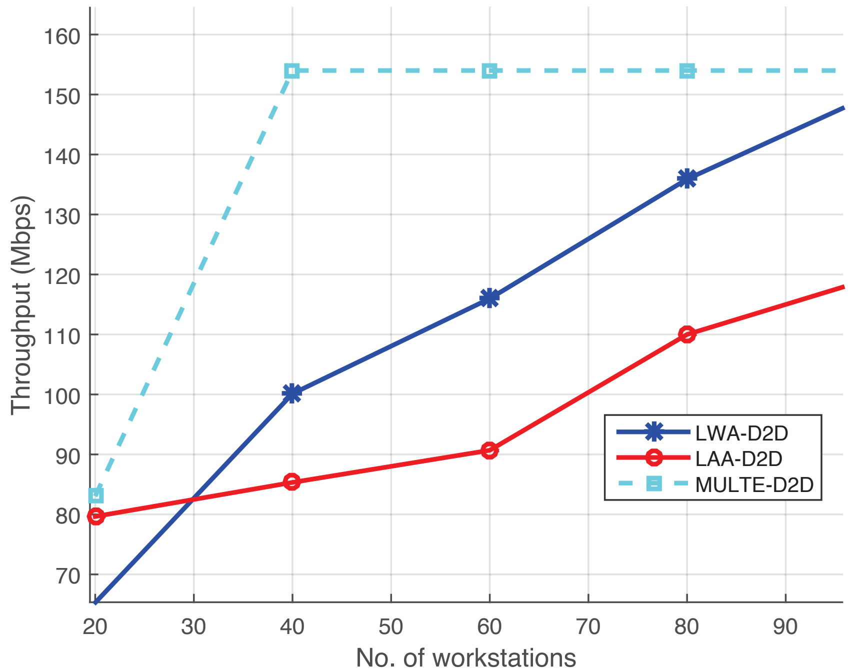


Figure 3.5: Network Throughput with D2D communication

need to have a scalable network it will be more beneficial to apply D2D communication using LTE-LAA and LWA, as more users can be accommodated with higher throughput.

3.5 Conclusion

This chapter presents a detailed overview of the solutions for LTE operating in unlicensed 5GHz band, along with their features and comparison. Simulation is also performed for LTE-LAA, LWA, and MuLTEfire in the presence of Wi-Fi hot spot. It is concluded from the simulation results that MuLTEfire can provide higher throughput but limits overall scalability of the network coexistence. Consequently, Wi-Fi users have a higher impact in the presence of MuLTEfire and can degrade the performance of Wi-Fi throughput. LTE-LAA has higher throughput for smaller load but the throughput decreases as the load increases and also Wi-Fi throughput degrades. Whereas, LWA can provide better throughput than LTE-LAA for a large network and has minimal effect on the Wi-Fi users as there is no contention between the users. Similarly, when D2D communication is investigated in these techniques MuLTEfire can increase the throughput but cannot increase the users whereas LWA and LTE-LAA can accommodate more users, where LWA showing better throughput. LWA gives better results when a Wi-Fi AP is deployed near it and is more cost-effective and easy to implement as compared to LTE-LAA. For future challenges, it will be very important that all the technologies operating in 5GHz band should have equal control to access the medium to satisfy fairness.

A new Scalable Mac Protocol (SC-MP) is proposed in Chapter 4 for offloading traffic from licensed band to unlicensed band in three tier architecture based on LWA technology.

Chapter 4

Scalable MAC Protocol For D2D Communication For Future 5G Networks

4.1 Introduction

As more users and devices become connected, one of the leading challenges facing mobile operators is to meet the demand of secure and high speed data to the customers with consistent quality. To meet this demand deployment of smaller cells has gained significant importance, as it will not only increase capacity but will also enhance the coverage of the network. Wireless local area network (WLAN), as an additional small cell, is gaining immense popularity, as it enables access to local services and internet at a lower cost for infrastructure by utilizing unlicensed band [144].

Joint deployment of cellular networks and WLAN can combine the unique strength of Wi-Fi with cellular network and can offer the subscriber a true compelling experience. It will increase the overall capacity of the heterogeneous networks with minimum interference. Recently, different LTE-Wi-Fi aggregation solutions have been introduced to boost the capacity as mentioned in Chapter 3.

As mentioned in Chapter 3, allowing D2D communication in LTE-Wi-Fi aggregation solutions can be very beneficial with D2D operating in unlicensed band for Wi-Fi transmissions, to further offload traffic and enhance coverage while minimizing cross-tier and cross-technology interference.

Nowadays, popular YouTube videos and movies are downloaded by sending a request to BS, such requests waste precious spectral resources. Storage space is the fastest growing quantity in communication devices. Smart-phones and tablets, that have 10-64 GB of storage are under-utilized [90]. D2D communication can be very beneficial for sharing of various types of contents. Service providers can take advantage of D2D functionality by offloading traffic in an area where traffic is increased for a certain period of time, for example, during an exhibition, fair, or stadium by allowing direct transmission between two devices without choking the network.

The aim of this chapter is to propose an idea which can revolutionize an architecture for next generation and can accomplish the capacity issue in a dense environment without degrading the network performance. A three tier 5G architecture is proposed where the cellular traffic will be offloaded from Macro Base Station (MBS) to Wireless Local Area Network Base Station (WBS). A Scalable MAC (SC-MP) protocol for D2D communication based on IEEE 802.11 PCF access mechanism is proposed. Furthermore, a novel SDN-based mechanism for PCF based on best Signal to Noise Ratio (SNR) polling scheme is proposed. The throughput of the network is further increased by sharing multimedia through D2D communication between WLAN users. The proposed architecture builds on the LWA technology with modification in resource allocation scheme based on IEEE 802.11 PCF to access WLAN channel with an additional increase in throughput by D2D communication. The simulation results show that the proposed scheme can perform better in a dense network when compared IEEE

802.11 Distributed coordinated function (DCF) access mechanism and LTE-U.

4.2 Background

Few papers have been published relating to offloading excess traffic from cellular to WLAN networks [23,25], especially in indoor environments [157], but this area requires extensive attention in research, especially for outdoor environment with further improvement using D2D communication. Resource Allocation is one of the vital features that needs to be addressed efficiently and effectively to improve the performance of the network when offloading traffic using D2D communication in heterogeneous network.

4.3 System Model

4.3.1 Proposed Three Tier Architecture

Often, during peak times, Macro Base Station cannot accommodate all users, either QoS is degraded due to interference or the call cannot get through due to network saturation. Deploying a temporary WLAN Base Station (WBS) in these problematic areas will be very beneficial as it can accommodate more users without degrading the QoS. It is easier to deploy WLAN because of its cost effectiveness and relatively low difficulty to deployment. Figure. 6.1 presents a system model that proposes a three tier 5G heterogeneous architecture for a dense environment. Long Term Evolution (LTE) technology is used as a baseline in the proposed architecture for the cellular network.

A cell consists of one LTE MBS that is overlayed by one WBS. Real-time services; voice and video are taken into consideration. All users operate in multi-mode terminals that can function in both radio technologies: LTE and WLAN. In the proposed scenario, Quality of service (QoS) is taken into account and users are differentiated in two categories, pre-

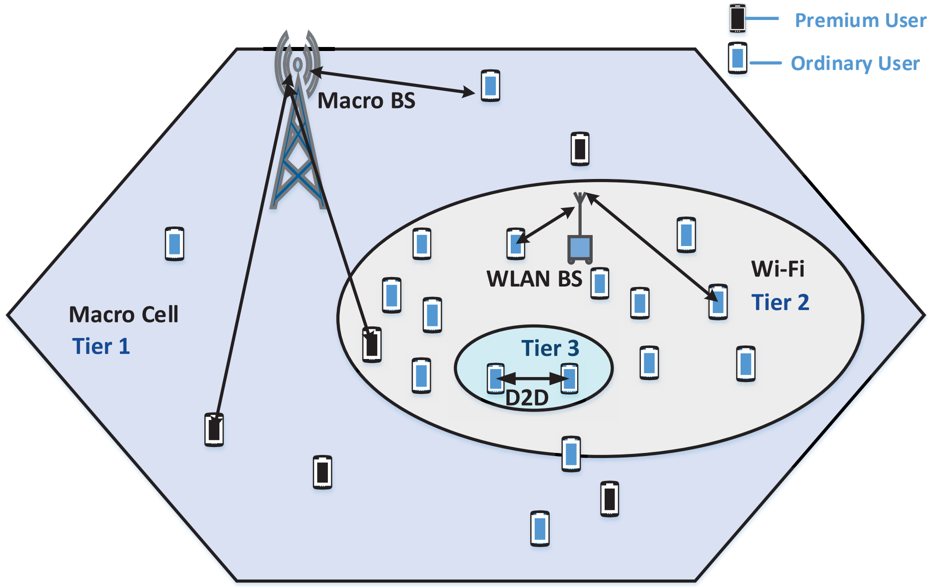


Figure 4.1: Proposed three tier Future 5G Architecture For Dense Environment

mium/corporate and normal users. Signal to interference Noise ratio (SINR) is a major parameter for users in the downlink to access WLAN. Premium users are always given priority over normal users and, as a result, they are always connected to LTE MBS. LTE MBS will calculate the SINR of each premium user and will accommodate the best premium users with the best SINR and will dedicate a certain amount of bandwidth (BW). The premium users can use the dedicated bandwidth for voice or video applications. Normal users are offloaded to WLAN based on the SNR, to further improve the network performance.

It is assumed that the network is synchronized and WBS is aware of the geographical location of users. If a user moves from one location to another, the location of that user are updated by WBS. As this architecture is proposed for 5G, it is also assumed that Software Defined Network (SDN) is deployed at the core end. To gain higher speed and massive connectivity, it is important that decisions made at the core side be rapid, structured and congestion controlled. Software defined controllers that appear to have a global view of the network will manage the network services quicker and more efficiently. Currently, SDN and routing protocol are out of the scope of our paper but for further understanding paper [158] can be referred.

Resource allocation in WLAN has to take place efficiently and intelligently so that most normal users can be accommodated without saturating the network. To achieve this, we propose a new MAC protocol called SC-MP which is described in the following section.

4.3.2 Scalable MAC Protocol (SC-MP)

Resource allocation schemes need to be devised so that they can balance the traffic load and minimize the interference simultaneously, hence allowing for a balance between these two objectives to be achieved.

SC-MP combines the advantages of PCF and Time Division Multiple Access (TDMA-based) services for transporting voice and video packets efficiently. PCF works efficiently under high network load and reduces contention for a large number of users. TDMA partitions time into fixed slots and users transmit data in their assigned slots, hence avoiding collision and packet loss. TDMA based protocols are more energy efficient and the energy consumed is proportional to the length of the transmission cycle while the latency is proportional to the size of the network [71]. As PCF works over polling based scheme, instead of using simple round robin scheme, we introduce a best SNR polling scheme, as it will minimize the delay by taking in account the channel conditions and can connect the users with the best coverage first. In this protocol, each user is assigned a time slot (TS) once its polled by the point coordinator, resulting in dynamic slot assignments on a frame-by-frame basis. WBS will be acting as a point coordinator (PC) for the users associated with WLAN. WBS has to first sense the channel idle for PCF Inter-frame Space (PIFS) period and then a beacon frame is transmitted at the start of the contention-free period by WBS. As WBS is aware of the location of users, it calculates the distance of the user from WBS and further calculates

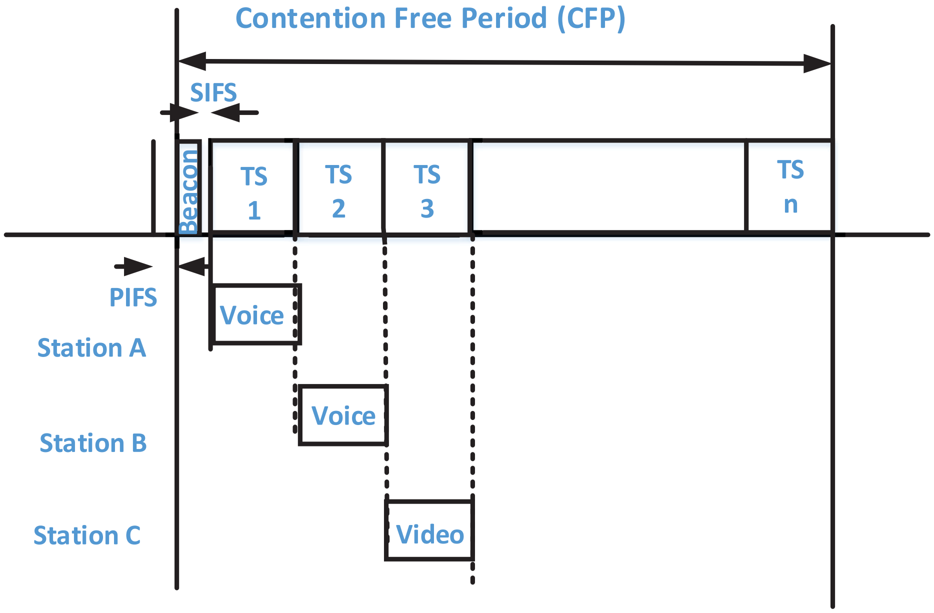


Figure 4.2: Polling scheme based on Best SNR in SC-MP

the SNR of each user in its vicinity and will create the polling list with descending SNR. In this scheme, WBS must poll all stations regardless of whether a station has a pending frame or not. Each frame begins with the polling of the first user which has the best SNR and this order continues with better SNRs taking priority over. If the first user needs to transmit a packet, a TS is assigned to the user by WBS immediately. Once the transmission is over, WBS polls the second user on the polling list and so on. If the user does not need to transmit data, a null frame is sent by the user and no TS is assigned to it and the next user will be polled. If all the users in the polling list are not polled and the cycle ends, in the next cycle WBS will resume polling from where the polling list ended. This will provide fairness between the normal users. The advantage of creating a polling list on the basis of SNR is that the best coverage or closest user is always served first which will cause less delay and maximize the throughput.

Figure. 4.2 shows the polling mechanism for SC-MP based on best SNR, where station A has the best SNR followed by station B and station C. Station A is polled and a TS is assigned by the WBS for voice call. Similarly station C is polled and assigned a TS to download a video.

SC-MP will not only allocate resources in an efficient manner but also provide fairness to all users. Since each station is scheduled to transmit at a definite time slot no collision occurs in SC-MP, this will, in turn, maximize the throughput.

4.3.3 Device to Device Communication (D2D) in SC-MP

The SC-MP scheme is further extended by introducing D2D communication for caching popular multimedia files, in particular the video files, to mobile devices and WLAN BS.

When a user is polled by WBS and needs to download a video, WBS check in its cache memory, whether or not it has already been downloaded. If the video is not downloaded, WBS assigns a time slot to the user to download the video, as shown in Figure. 4.2 for station C. Once the video is downloaded, it will be stored in the cache memory of WBS with the location information of the user. If a new user request for the already downloaded video, WBS will check for the video and location information of the user in reference to that video. WBS will check the distance between the downloaded user and the new user. If the distance is below the threshold value, WBS will allow D2D communication between the two users by assigning a free channel for a specific time period. If a video is not requested, by any other user for a certain time, WBS deletes the video from the cache memory to save storage space. By allowing D2D communication in WLAN, it can decrease the delay especially for the normal users at the edge of the cell, which in return will improve the network performance. Also D2D communication can further increase the capacity of the network by saving WLAN resources in efficient manner.

A flow chart for the proposed scheme is shown in Figure. 4.3.

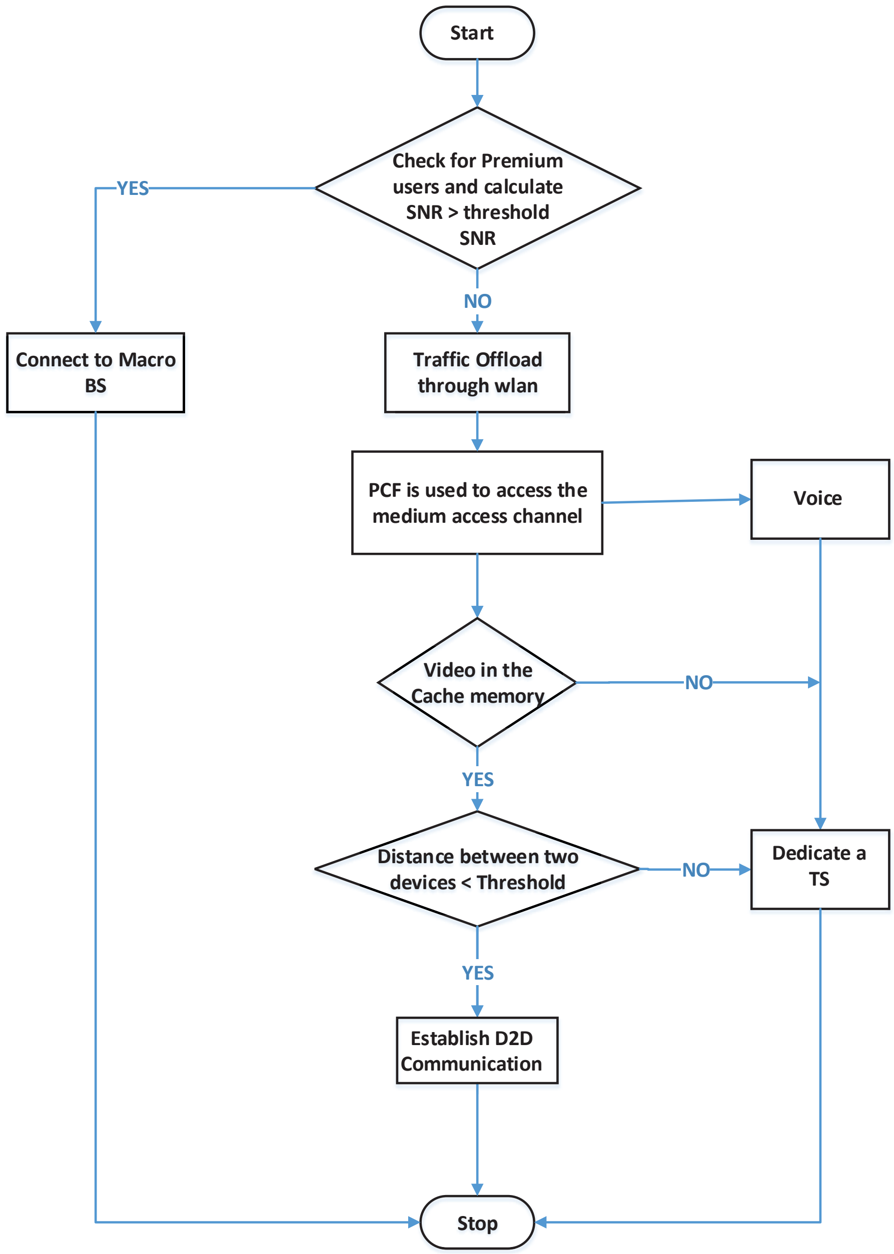


Figure 4.3: Flow Chart of Proposed Scheme

4.4 Simulation Result

A single hexagonal cell is considered with one LTE MBS overlaid with one WBS. In the analysis, both voice and video applications are considered. LTE system with a maximum capacity of 100 Mbps and WLAN with the capacity of 54 Mbps. Users are randomly generated that varies from 20 to 150. In the simulation, 20 percent of users are assumed to be premium users with the rest being normal and 5 Mbps bandwidth is dedicated to the premium users which they can use for video or voice applications.

The network load, network throughput, and packet loss ratio is evaluated through simulation. The simulation consists of three scenarios, first scenario is the LTE system without offloading, second scenario is the LTE system offloading to WLAN using the DCF access mechanism, and the third scenario is the proposed scheme which is the LTE system with offloading to WLAN using the PCF access mechanism with D2D communication.

Figure. 4.4 illustrates the result for a total network load versus the number of users. The maximum load capacity of LTE system is 100 Mbps and load capacity for WLAN is 54 Mbps. The graph shows that LTE system without offloading reaches saturation around at 60 users, which includes the premium users as well as the normal users. It can be seen from the graph that proposed scheme performs better as compared to DCF when the traffic offloading scenario is taken into LTE system. In a large network, the DCF access mechanism can cause high collisions between the users. Further, load increases in the proposed scheme due to D2D communication between the WLAN users.

Figure. 4.5 shows the network throughput graph. LTE system without offloading has the lowest throughput whereas the proposed scheme shows the better result than LTE without offloading and LTE offloading with DCF. As the number of normal users increases, more users try to connect to WBS, hence increasing the throughput of the WLAN, since more packets are successfully received at the access point. The throughput of WLAN using the

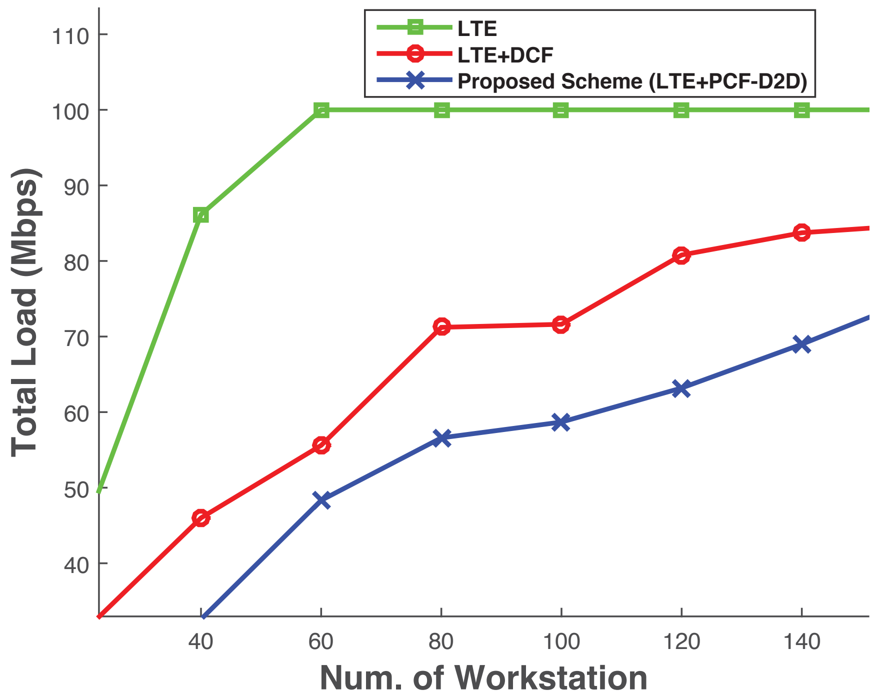


Figure 4.4: Total Network Load vs Total Number of Workstations

DCF access mechanism is less than the proposed scheme because more collision occurs with the increase of users, whereas in the proposed scheme throughput increases as there is no collision in the PCF access mechanism and two users within range can start D2D communication for content sharing.

Packet loss ratio graph has been shown in the Figure. 4.6. The proposed scheme has almost no packet loss as each active user is assigned a time slot to transmit the data and using D2D communication they can communicate without the loss of packet as the distance between the users is short. It can be seen from the graph that if DCF is used to offload traffic in WLAN, it can create high packet loss ratio due to the collision. The results show that the proposed scheme can not only increase the capacity but also can maximize the throughput of premium and normal users using D2D communication.

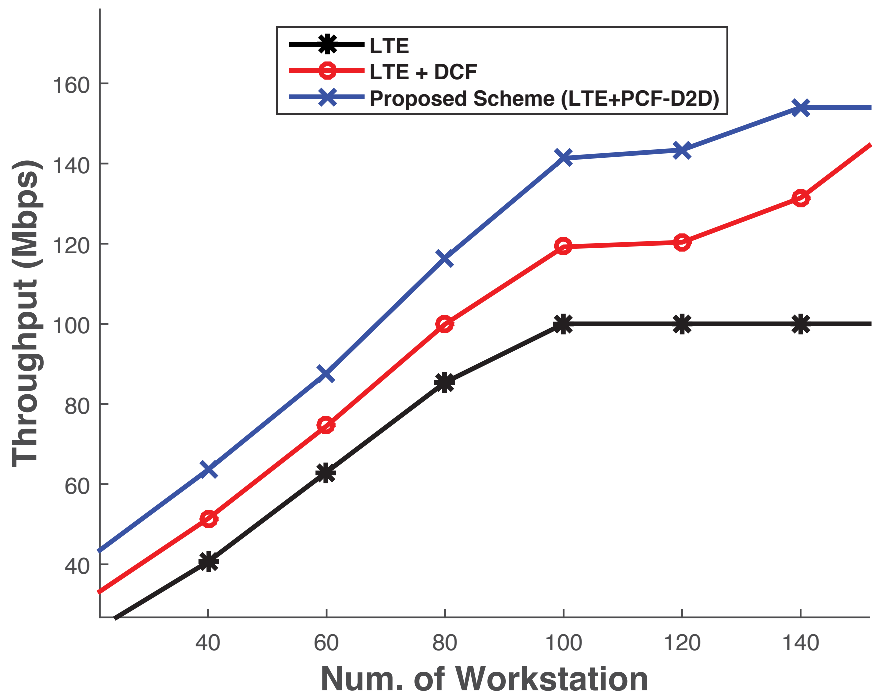


Figure 4.5: Total Network Throughput vs Total Number of Workstations

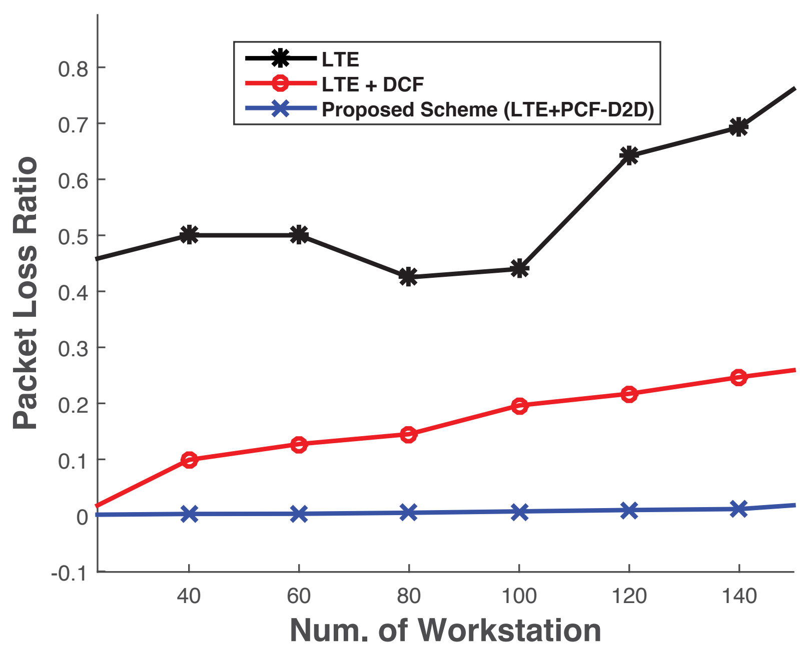


Figure 4.6: Packet Loss Ratio vs Total Number of Workstations

4.5 Applications

The proposed three tier 5G architecture and proposed Scalable MAC protocol for D2D communication has great potential to be applied in practical scenarios. It will be beneficial for events like:

- Sport matches in a stadium
- Concerts
- Award ceremonies
- Fun fairs
- School holiday programs

4.6 Conclusion

The collective use of smart-phones and data-hungry applications have evolved in significant growth in demand for wireless data services in recent years. To meet this requirement, one of the solution is to offload cellular traffic from a Macro cell to a smaller cell, particularly WLAN, as it operates in the unlicensed band.

In this chapter, a novel three tier 5G architecture is proposed with a technique to offload traffic from LTE to WLAN. Furthermore, a scalable MAC protocol for D2D communication is proposed based on PCF access mechanism. The best SINR polling scheme is used for PCF which can take in consideration the channel condition of the user instead on simple round robin scheme. D2D communication further increases the network throughput by sharing content files between the WLAN users and decreases the delay of the normal users.

A comprehensive series of simulation quantifies the SC-MP effectiveness, The key simulation results show that proposed architecture and scheme performs better as compared to IEEE

802.11 DCF access mechanism in terms of throughput, packet loss, and greater number of users can be accommodated when implementing D2D communication in WLAN.

This chapter identifies interesting points towards future research which includes, detailed economic comparison of energy consumption in SC-MP and detailed examination of scenarios in more complex mobility models.

Chapter 5 analytically explores the throughput and effective capacity of SC-MP using Markov chain and semi-Markov modelling.

Chapter 5

Analysis of Effective Capacity and Throughput of Polling Based Heterogeneous Networks

This chapter proposes a performance model for Scalable MAC Protocol (SC-MP) based on IEEE 802.11 point coordinated function (PCF) having diverse traffic; voice and video/multimedia. D2D communication is further applied to offload the video/multimedia traffic. In particular, the chapter establishes a three-state semi-Markovian model to derive a closed-form expression of effective capacity in terms of transmission rate and quality-of-service (QoS). Further, SC-MP is analytically modelled using four-state traditional Markov model to derive the saturation throughput. The analytical results are validated through simulations and proving the appropriateness of the model.

The major contributions and results of this chapter can be summarized as follows:

- An analytical three state semi-Markovian model is proposed for SC-MP based on PCF access mechanism with D2D communication;

- A new closed-form expression is derived to evaluate the effective capacity of SC-MP against the QoS and instantaneous transmission rate for random diverse traffic; voice and video. The three-state semi-Markovian model is modified to traditional four-state Markovian model to analyze the saturation throughput having persistent traffic for voice and video/multimedia; and

- Analytical results are validated by extensive Matlab simulations. We compare the proposed scheme SC-MP with DCF access mechanism with D2D communication supported.

The chapter is organized as follows. In Section 5.1, related works are mentioned. In Section 5.2, a hypothetical structure of SC-MP is set-up to analyze the effective capacity of the three-state semi-Markovian model. In Section 5.3, the semi-Markovian model is modified to traditional Markovian model to analyze the throughput of SC-MP. Furthermore, the analytical model is validated against a detailed simulation in Section 5.4. Conclusion is discussed in Section 5.5.

5.1 Related Work

In this section some related works are discussed. The performance characteristics of PCF has been extensively studied in [159–161]. The performance of video transmission using PCF mechanism is discussed in [162,163]. In [164], two-level of polling scheme has been discussed for real time services to analyze the delay through embedded Markov chain theory and probability generating method. In [165], polling system has been analyzed using exhausted services through embedded Markov chain theory. An analytical model for the delay in PCF for real time and sensitive traffic is presented in [166]. The method neither guarantees the QoS and low delay for sensitive traffic. In [167], an analytical model for non-ideal channel conditions is presented for throughput, expected channel access delay, and frame loss PCF.

None of the articles give a detail analytical study of using D2D communication with WLAN using PCF as access mechanism.

The effective capacity/bandwidth theory provides for an asymptotically tight linkage between source characteristics, system resources, and QoS [168, 169]. The effective bandwidth theory exemplifies the traffic details of time-varying bursty resources in a single function, known as effective bandwidth function, that can be utilized to express the minimal server capacity required to satisfy a given overflow probability related QoS limitations [170, 171]. In [172], a cooperative ARQ MAC protocol is proposed that is compatible to the legacy IEEE 802.11 DCF. It coordinates the transmissions among a set of relay nodes which act as helpers in a bidirectional communication. The maximization of the effective capacity with QoS and power constraints using DCF is proposed in [173]. Not much work has been done in the past taking into account effective capacity until recently [170, 173–176].

In [174, 175], the effective capacity is focused on the modelling of rate fluctuations at the physical layer. Whereas [170, 176], focuses on the effective capacity of DCF based MAC layer.

The articles [170, 172, 174–178], used DCF as access mechanism for WLAN and is not recommended for dense environment due to contention problem and high packet loss. Instead, this chapter will propose the saturated throughput and effective capacity theory for MAC layer using PCF scheme with diverse traffic; voice and video/multimedia.

5.2 System Model and Analysis of Effective Capacity of SC-MP Through Semi-Markov Process

The semi-Markov process is the generalization of Markov chain that augments the specification of the process by including a state holding-time [177]. The semi-Markovian process is the actual random process that evolves over time and any realisation of the process has a defined state for any given time. The entire process is not Markovian, i.e., memoryless, instead the process is Markovian only at the specified jump instants. In this section, the effective ca-

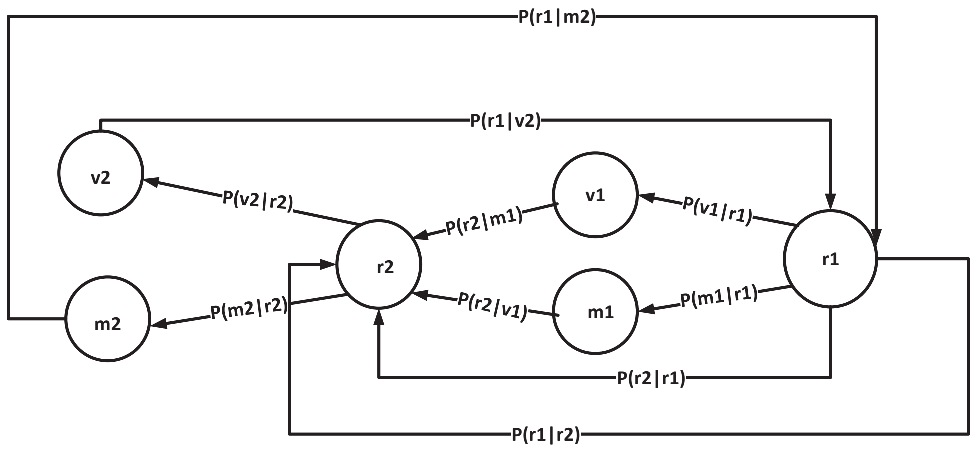


Figure 5.1: State Diagram of network using Semi-Markov Model with two users

pacity of the SC-MP is investigated for diverse random traffic; voice and video/multimedia, given the QoS exponent θ and instantaneous transmit rate R, using semi-Markovian model. The overall transmit power of Wi-Fi is P_{tot}. The power is allocated to K Wi-Fi users associated with Wi-Fi BS. The transmit power for k Wi-Fi user is P_k. B is the bandwidth and each Wi-Fi user is allocated a subband of $\frac{B}{K}$, where k = *1*,. . .,K. Similar to [176], the instantaneous transmit rate R is given as

$$R_k = \frac{B}{K} \log_2(1 + \frac{G_k P_k}{\sigma^2}) \tag{5.1}$$

where B is the total bandwidth, K is the total number of users, G_k is the channel gain of the Wi-Fi user, P_k is the transmit power of Wi-Fi user and σ^2 is the noise power. The (5.1) can be rewritten as

$$R_k = B_k \log_2(1 + \gamma_k) \tag{5.2}$$

where γ_k is the signal to noise ratio (SNR) of user k.

It is significant to mention that effective capacity is on a user basis and can be computed given the QoS requirement of a user, and the transmit power and bandwidth allocated to a user. The bandwidths can be divided unequally among the users.

The quality-of-Service (QoS), is one of the imperative parameters towards 5G technology and is very crucial to implement. The Wi-Fi users can have diverse QoS requirements including end-to-end delay compromised of queuing delay and transmission delay. A set of FIFO queues is used to buffer data traffic destined to different Wi-Fi users, one queue per user. The QoS can be can be characterized statistically by employing the QoS exponent θ_k, where $k = 1$,......,K, as given by [174, 176, 178];

$$\theta_k = -\lim_{Q_k^{th}\to\infty} \frac{\log(Pr\{Q_k(\infty) > Q_k^{th}\})}{Q_k^{th}} \tag{5.3}$$

where $Q_k(t)$ is the length of the FIFO queue at the corresponding Wi-Fi AP to buffer the downlink traffic for Wi-Fi user k at time t, Q_k^{th} is the threshold of the queue length specified for the traffic, and $PrQ_k(\infty) > Q_k^{th}$ is the buffer-overflow probability. In this sense, θ_k provides the exponential decaying rate of the probability that the threshold is exceeded.

The effective capacity for Wi-Fi user k, $k = 1, \ldots, K$, denoted by $C_k(\theta_k)$. It specifies the maximum, consistent, steady-state arrival rate at the input of the FIFO queue, as given by [174, 178];

$$C_k(\theta_k) = -\lim_{t\to\infty} \frac{1}{\theta_k t} \log(\mathbf{E}\{e^{-\theta_k S_k(t)}\}) \tag{5.4}$$

where $S_k(t)$ is the number of bits successfully delivered to Wi-Fi user n during (0,t] and $\mathbf{E}\{.\}$ denotes expectation.

K number of users is modelled with total $3K$ states with each user having 3 states. In Figure. 5.1, a semi-Markovian model is shown with 2 users having 6 states in total.

Two types of traffic is considered in the framework for the allocation of a time slot; voice and video/multimedia. Moreover, D2D communication is also applied to video/multimedia traffic for offloading the traffic. It is assumed that the duration of voice traffic in a time slot is random and the voice duration can vary for every user depending on the packets in the buffer. Similarly, for multimedia traffic either the user is allocated a time slot to transmit or the user will be allowed for D2D communication. The probability that a user has data in the buffer is P, and $1\text{-}P$ is the probability that there is no data in the buffer. The probability P can either be a voice traffic or a video/multimedia traffic.

$$P = P_V \text{ or } P = P_M \tag{5.5}$$

where P_V is the probability of voice traffic and P_M is the probability of video/multimedia traffic.

Two types of channel condition are used in modelling the SC-MP; A and B. B is the channel condition that should be satisfied for the allocation of the time slot for voice and video/multimedia traffic. Whereas, A is the channel condition that should be satisfied for D2D communication between two users. D2D communication will be allowed if the video requested by the user is already downloaded by the neighbouring user/users and is in the cache memory of the AP and satisfies the channel condition A. The probabilities are defined as, P_N= Probability that video requested is not downloaded by any neighbouring users, P_S= Probability that channel condition B is not satisfied and P_D= Probability that channel condition A is not satisfied.

Each user has 3 unique states: r_k = polling state of a Wi-Fi user, v_k = time slot allocation for a Wi-Fi user having voice traffic and satisfying the channel condition B, m_k = a time slot is allocated for video/multimedia traffic to a Wi-Fi user after satisfying the channel condition B or allow D2D communication to a Wi-Fi user if the neighbour has already downloaded the video/multimedia file and satisfy the channel condition A, where $n = 1, \ldots, K$.

The probabilities of the transition states are denoted as $P(v_k|r_k)$, $P(m_k|r_k)$, $P(r_{k+1}|v_k)$, $P(r_{k+1}|m_k)$, and $P(r_{k+1}|r_k)$.

Derivation of Semi-Markovian Model Transition Probabilities

The transition probability equations for the two users modelled in Figure. 5.1 is derived, transitioning from one state to another.

Transition Probability $P(v_1|r_1)$ & $P(v_2|r_2)$

When the $user_1$ is polled by the access point (AP) and is in polling state r_1 and has voice data in its buffer. AP will allocate a time slot TS_1 to the $user_1$ if it satisfies the channel condition B. The transition probability from polling state r_1 to v_1 for $user_1$ is given as

$$P_{A1} = P(v_1|r_1) = P_{V1}(1 - P_{S1}) \tag{5.6}$$

Similarly, the transition probability for $user_2$ is given as

$$P_{A2} = P(v_2|r_2) = P_{V2}(1 - P_{S2}) \tag{5.7}$$

where P_{V1} and P_{V2} are the probabilities of voice traffic for $user_1$ and $user_2$. $1\text{-}P_{S1}$ and $1\text{-}P_{S2}$ are the probabilities that satisfies the channel condition B for time slot allocation for $user_1$ and $user_2$. If all the users have voice traffic and satisfies the channel condition than the probability for voice is 100% i.e 1, whereas if no user has voice traffic or voice user does not satisfies the channel condition B the probability for voice will be 0.

Transition Probability $P(m_1|r_1)$ & $P(m_2|r_2)$

When the $user_1$ is polled by the access point (AP) and is in polling state r_1 and has video/multimedia traffic in its buffer. The AP will check if the video/multimedia file is already downloaded by any neighbouring user/users and is in its cache memory. If the video is not already downloaded, AP will allocate a time slot TS_2 if the $user_1$ satisfies the channel condition B. If the video/multimedia file is already downloaded, AP will calculate the channel condition A between user$_1$ and the neighbouring user/users who already have downloaded the video. D2D communication is established between the two users after satisfying channel condition A using contention period. The transition probability from r_1 to m_1 is given as

$$P_{F1} = (P_{M1}|r_1) = P_{M1}(1 - P_{S1}) + (P_{M1})(P_{S1})[1 - (P_{ND1} + ((1 - P_{ND1})P_{D1})^{N})] \tag{5.8}$$

Similarly, for user$_2$ the transition probability is given as

$$P_{F2} = (P_{M2}|r_2) = P_{M2}(1 - P_{S2}) + (P_{M2})(P_{S2})[1 - (P_{ND2} + ((1 - P_{ND2})P_{D2})^{N_n})] \tag{5.9}$$

where P_{M1} and P_{M2} are the probabilities of video/multimedia traffic for user$_1$ and user$_2$. *1*-P_{S1} and *1*-P_{S2} are the probabilities that satisfies the channel condition B for time slot allocation for user$_1$ and user$_2$, N_n is the total number of neighbours that downloaded the video, P_{ND1} and P_{ND2} are the probabilities that video requested by user$_1$ and user$_2$ is not downloaded by any neighbouring users and P_{D1} and P_{D2} are the probabilities that channel condition A is not satisfied for D2D communication for user$_1$ and user$_2$. If users do not have any video/multimedia traffic or does not satisfy channel condition B/A, the probability for video/multimedia traffic will be zero. If users do not satisfy the channel condition for D2D communication or no neighbouring users have downloaded the video/multimedia (5.8), (5.9) reduces to

$$P_{F1} = (P_{M1}|r_1) = P_{M1}(1 - P_{S1}) \tag{5.10}$$

$$P_{F2} = (P_{M2}|r_2) = P_{M2}(1 - P_{S2}) \tag{5.11}$$

Transition Probability $P(r_2|r_1)$ & $P(r_1|r_2)$

When the $user_1$ is polled by the access point (AP) and is in polling state r_1 and has no data to transmit or does not satisfy channel condition A and B, the AP will poll the next user in the polling list. The transition probability for polling state r_1 to polling state r_2 is given as

$$P_{E1} = 1 - P_{A1} - P_{F1} \tag{5.12}$$

$$P_{E2} = 1 - P_{A2} - P_{F2} \tag{5.13}$$

where P_{A1}, P_{F1}, P_{A2} and P_{F2} are derived in (5.6), (5.8), (5.7) and (5.9).

Transition Probability $P(r_2|v_1)$, $P(r_2|r_{m1})$, $P(r_1|v_2)$ & $P(r_1|m_2)$

The transition probability for all these states is 1.

Semi-Markov Modelling and Effective Capacity

The state transition probabilities generate the transition matrix $\mathbf{Q}$, using semi-Markovian model for 2 users and is given as

$$\mathbf{Q} = \begin{bmatrix} 0 & P_{A1} & P_{F1} & P_{E1} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ P_{E2} & 0 & 0 & 0 & P_{A2} & P_{F2} \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{5.14}$$

where $P_{A1},\ P_{F1},\ P_{E1},\ P_{A2},\ P_{F2},\ P_{E2}$ are already derived in (5.6), (5.8), (5.12), (5.7), (5.9) and (5.13).

$$\mathbf{\Gamma}(s,u) = \begin{bmatrix} M_1(-u) & 0 & 0 & 0 & 0 & 0 \\ 0 & M_2(Rs-u) & 0 & 0 & 0 & 0 \\ 0 & 0 & M_3(Rs-u) & 0 & 0 & 0 \\ 0 & 0 & 0 & M_4(-u) & 0 & 0 \\ 0 & 0 & 0 & 0 & M_5(Rs-u) & 0 \\ 0 & 0 & 0 & 0 & 0 & M_6(Rs-u) \end{bmatrix} \tag{5.15}$$

$$\mathbf{H}(s,u) = \begin{bmatrix} 0 & M_1(-u)P_{A1} & M_1(-u)P_{F1} & M_1(-u)P_{E1} & 0 & 0 \\ 0 & 0 & 0 & M_2(Rs-u) & 0 & 0 \\ 0 & 0 & 0 & M_3(Rs-u) & 0 & 0 \\ M_4(-u)P_{E2} & 0 & 0 & 0 & M_4(-u)P_{A2} & M_4(-u)P_{F2} \\ M_5(Rs-u) & 0 & 0 & 0 & 0 & 0 \\ M_6(Rs-u) & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{5.16}$$

The duration of both states r_1 and r_2 is T_{r_1} and T_{r_2} and is constant for both users. The duration for v_1 and v_2 is random and denoted by T_{v_1} and T_{v_2}, which depend on the number of packets in the buffer. The duration for m_1 and m_2 is random and is denoted by T_{m1} and T_{m2}, it can either transmit video/multimedia traffic through time slot or can allow D2D communication between users depending on the channel condition and neighbouring downloaded traffic.

The moment generating functions (MGF) of $T_{r1},\ T_{v1},\ T_{m1},\ T_{r2},\ T_{v2},\ and\ T_{m2}$ are $M_1(t){=}e^{tTr1}$,

$M_2(t)$=$\Sigma e^{tTv1}P_{A1}$, $M_3(t)$=$\Sigma e^{tTm1}P_{F1}$, $M_4(t)$=e^{tTr2}, $M_5(t)$=$\Sigma e^{tTv2}P_{A2}$, *and* $M_6(t)$=$\Sigma e^{tTm2}P_{F2}$.
where P_{A1}, P_{F1}, P_{A2}, P_{F2} are already derived in (5.6), (5.8), (5.7) and (5.9).

Two auxiliary variables are defined, s and u, with reference to [170,179]. A diagonal matrix is created Γ(s,u). The diagonal elements of Γ(s,u) are the MGFs of the six state semi-Markovian model and is given in (5.15).

$|\mathbf{H}(-\theta,-\theta C)-\phi(-\theta,-\theta C)\mathbf{I}|$ =

$$\begin{vmatrix} -\phi(-\theta,-\theta C) & M_1(-u)P_{A1} & M_1(-u)P_{F1} & M_1(-u)P_{E1} & 0 & 0 \\ 0 & -\phi(-\theta,-\theta C) & 0 & M_2(Rs-u) & 0 & 0 \\ 0 & 0 & -\phi(-\theta,-\theta C) & M_3(Rs-u) & 0 & 0 \\ M_4(-u)P_{E2} & 0 & 0 & -\phi(-\theta,-\theta C) & M_4(-u)P_{A2} & M_4(-u)P_{F2} \\ M_5(Rs-u) & 0 & 0 & 0 & -\phi(-\theta,-\theta C) & 0 \\ M_6(Rs-u) & 0 & 0 & 0 & 0 & -\phi(-\theta,-\theta C) \end{vmatrix} \tag{5.17}$$

$$\begin{aligned} = \phi(-\theta,-\theta C)^6 - \phi(-\theta,-\theta C)^4 M_1(\theta C)M_4(\theta C) - \phi(-\theta,-\theta C)^3 M_1(\theta C)M_4(\theta C)X- \\ \phi(-\theta,-\theta C)^2 M_1(\theta C)M_4(\theta C)Y \end{aligned} \tag{5.18}$$

The matrix **H**(s,u) is a non-negative irreducible matrix, as it cannot be arranged to upper triangular matrix, e.g, by using the Gaussian-Newton Method [176]. The spectral radius of **H**(s,u), denoted by $\phi(s,u)=\rho(\mathbf{H}(\text{s,u}))$ is a simple eigenvalue of **H**(s,u) and $\rho(.)$ represents the spectra radius.

Each permissible pair of s and u can be written as $\mathbf{H}(s,u)=\mathbf{\Gamma}(s,u)\boldsymbol{Q}$ and is given in (5.16). With reference to [179, Theorem 3.1], there exist a unique $u^*(s)$ for given $s\leq 0$, such that $\phi(s,u^*(s)) = 1$ and $\lim_{t\to\infty}\dfrac{1}{t}\log(\boldsymbol{E}\{e^{S(t)}\}) = u^*(s)$. In [179, Theorem 3.2], the effective capacity $C = \frac{u(s)}{s}$, when $\phi(s,u(s)) = 1$ and θ = -s.

As the result, the effective capacity C can be evaluated by solving $\phi(-\theta, -\theta C) = 1$ for $\theta > 0$ [170,179]. Meanwhile, $\phi(-\theta,-\theta C)$ is the eigenvalue of $(\boldsymbol{H}(-\theta,-\theta C))$, we have (6.4), where **I** is the identity matrix and |.| is the determinant of the matrix. The determinant is given in (6.5) where

$$X = P_{A1}P_{E2} + M_3(-R\theta + \theta C)P_{F1}P_{E2} + M_5(-R\theta + \theta C)P_{E1}P_{A2} + M_6(-R\theta + \theta C)P_{E1}P_{F2} \tag{5.19}$$

$$\begin{aligned} Y = {} & P_{A1}P_{A2}M_2(-R\theta + \theta C)M_5(-R\theta + \theta C) + P_{A1}P_{F2}M_2(-R\theta + \theta C)M_6(-R\theta + \theta C)P_{A1}P_{F2} \\ & + P_{A2}P_{F1}M_3(-R\theta + \theta C)M_5(-R\theta + \theta C) + P_{F2}P_{F1}M_3(-R\theta + \theta C)M_6(-R\theta + \theta C) \end{aligned} \tag{5.20}$$

The eigen value is substituted $\lambda = \phi(-\theta,-\theta C) = 1$ in (6.5). Further (6.5) is modified and propose to decouple θ between voice and video/multimedia. θ_v and θ_m representing the QoS exponent of voice and video/multimedia, respectively. The effective capacity of voice and video is denoted by $C_v\theta_v$ and $C_m\theta_m$. The effective capacity is calculated by changing the value of θ for both voice and video/multimedia.

5.3 Markov Chain and Analysis of Throughput of SC-MP

In this section the throughput of SC-MP is analysed and extend from three-state semi-Markovian model to four-state Markovian model. For comparison purpose, we develop a standard Markov model to analyze the effective capacity of the proposed SC-MP. As defined in the standard Markov model, each of the states needs to have a consistent state duration. The capacity of the Markov model can analyzed based on the steady-state probabilities of the states, as well as the time invariant durations of the states [180, 181]. We note that the state of multimedia on backhaul and the state of multimedia on D2D can have different durations depending on D2D traffic availability and channel condition. It is reasonable to set them as two separate states, as typically done in Markov modelling. In contrast, multimedia and D2D can be modelled statistically as a single multimedia state in the semi-Markov model. The semi-Markov model is able to capture time varying durations of each state. The multimedia and D2D can be probabilistically defined to characterize the duration of the single multimedia state and can merge into the state. Therefore, the semi-Markov modelling can have one less state than the Markov model which can characterize the state transitions of the proposed system.

In Markovian model the voice traffic is constant and video/multimedia state is split into two states; time slot allocation to video/multimedia traffic state and D2D communication state and is shown in Figure. 5.2. The transition probabilities for Markovian model is derived accordingly to the state diagram.

The transition probability from polling state r_1 to v_1 and r_2 to v_2 for user$_1$ and for user$_2$ is same as derived in Section 5.2 and is given as

$$P_{A1} = P(v_1|r_1) = P_{V1}(1 - P_{S1}) \tag{5.21}$$

$$P_{A2} = P(v_2|r_2) = P_{V2}(1 - P_{S2}) \tag{5.22}$$

The transition probability for allocation of time slot of video/multimedia traffic from polling

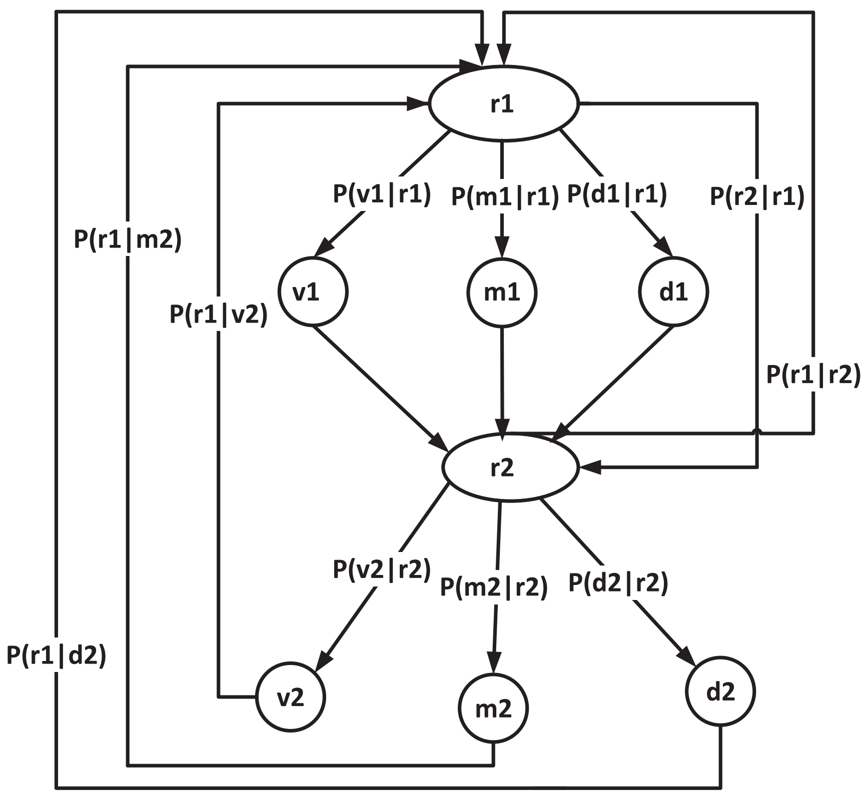


Figure 5.2: State Diagram of network using Markov Model with two users

state r_1 to m_1 for user$_1$ is given as

$$P_{B1} = P(m_1|r_1) = P_{M1}(1 - P_{S1}) \tag{5.23}$$

Similarly, the transition probability for user$_2$ is given as

$$P_{B2} = P(m_2|r_2) = P_{M2}(1 - P_{S2}) \tag{5.24}$$

The transition probability for D2D communication from polling state r_1 to d_1 for user$_1$ is given as

$$\begin{aligned} P_{C1} = P(d_1|r_1) = (P_{M1})(P_{S1})[1 - (P_{ND1} \\ +((1 - P_{ND1})P_{D1})^{N_n})] \end{aligned} \tag{5.25}$$

Similarly, the transition probability for user$_2$ is given as

$$\begin{aligned} P_{C2} = P(d_2|r_2) = (P_{M2})(P_{S2})[1 - (P_{ND2} \\ +((1 - P_{ND2})P_{D2})^{N_n})] \end{aligned} \tag{5.26}$$

The transition probability for polling state r_1 to polling state r_2 for user$_1$ is given as

$$P_{E1} = P(r_2|r_1) = 1 - P_{A1} - P_{B1} - P_{C1} \tag{5.27}$$

Similarly, the transition probability for user$_2$ is given as

$$P_{E2} = P(r_1|r_2) = 1 - P_{A2} - P_{B2} - P_{C2} \tag{5.28}$$

The transition probability for all other states are 1.

The state transition probabilities derived are used to generate a matrix, **M**, to solve the Markov process. Taking in consideration the model in Figure. 5.2, having 2 users with total 8 states and transition probabilities mentioned above, the matrix **M** is given as

$$\mathbf{M} = \begin{bmatrix} 0 & P_{A1} & P_{B1} & P_{C1} & P_{E1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ P_{E2} & 0 & 0 & 0 & 0 & P_{A2} & P_{B2} & P_{C2} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{5.29}$$

where P_{A1}, P_{B1}, P_{C1}, P_{E1}, P_{A2}, P_{B2}, P_{C2}, P_{E2} are the state transition probabilities derived in (5.6), (5.23), (5.25), (5.27), (5.7), (5.24), (5.26) and (5.28).

We know that the sum of all the probabilities must be 1. To find the steady state we multiply the matrix **M** by the column vector matrix **V** which is given as

$$\mathbf{V} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T \tag{5.30}$$

Throughput Analysis With Markov Chain

The normalized system throughput, defined as the fraction of time used to successfully transmit payload bits is given as

$$\mathbf{S} = \frac{\text{Successfull Data Transmission}}{\text{Total time taken to transmit the Data}} \tag{5.31}$$

Combining the expressions for the state probabilities derived, the generalized throughput equation we have

$$\mathbf{S} = \frac{NP_A d_v + NP_B d_m + NP_C d_d}{NP_A t_v + NP_B t_m + NP_C t_d} \tag{5.32}$$

where P_A, P_B, P_C, and P_E are the steady state probabilities, N is the total number of users, d_v is the data rate for voice, d_m is the data rate for video/multimedia, d_d is the data rate for D2D communication. The value t_v is the time spent to transmit voice traffic for the users and can vary with the probability of the voice traffic. The value t_m is the time spent to transmit multimedia traffic for the users when D2D communication is not taking place. The higher the probability of multimedia traffic with no D2D communications, the larger the value of t_m is. The lower the probability of multimedia traffic with no D2D communications, the smaller the value of t_m. Similarly, t_d is the time spend to transmit the multimedia traffic using D2D communication. The higher the probability of D2D communication the larger the value of t_d is. The smaller the probability of D2D communication the smaller the value of t_d is. t_d always takes smaller time to transmit, as compared to t_m and t_v as the distance between the users is small. The data rate is the same for all users.

5.4 Results and Analysis

In this section, the analytical results of Scalable MAC Protocol (SC-MP) model derived semi-Markov process and Markov chain are validated through simulation. The SC-MP is implemented in Matlab simulation tool and is compared with the proposed analytical models.

The performance of effective capacity of SC-MP is investigated by varying the QoS exponent and the number of users. The behaviour of effective capacity is also observed by changing the voice probability, multimedia probability, neighbour traffic probability and channel condition probabilities for time slot allocation and D2D communication. Further the analytical

Parameter	Value
SIFS Length	10 μs
Time slot Length	20 μs
DIFS	50 μs
PIFS	30 μs
Beacon Size	36 byte
CF- End Size	20 byte
CF-Poll	20 byte
CF -ACK	20 byte

Table 5.1: Simulation Settings

throughput behaviour of SC-MP is validated using traditional Markov chain through simulation.

We consider the IEEE 802.11g WLAN PHY [182] characteristics for simulations. The PHY layer modulation mode is set to 64 QAM, 3/4 code rate and 54 Mb/s data rate. In the simulations, the traffic model is constant with packets arrival using the Poisson distribution with a mean rate λ = 20 frames/sec. The users are randomly and uniformly distributed in a 250 m $\times$ 250 m area with the Wi-Fi BS at the centre. The ITU-UMI model is used to generate the channels between the users and the BS, where σ^2 = -174 dbm/Hz at each user and P_t = 23 dbm. The transmit data rate of a user is calculated using R = (B/N)(1+$G_n P_n/\sigma^2$), where n = $1, \ldots, N$. The simulation parameter settings are indicated in Table 5.1

In Figure. 5.3, the analytical effective capacity is compared with the simulation results, where P_V=0.3, P_M=0.7, P_D=0.3 and P_N=0.7. Each curve is plotted by increasing the persistent incoming rate of the Wi-Fi users and evaluating the achieved QoS exponent θ. The QoS exponent is specified at the x-axis, while the incoming rate is specified at the y-axis. The analytical results coincide with the simulation results for voice and video traffic, which proves that our analysis is correct. As the $\theta > 10^{-6}$, the QoS is too stringent and the effective capacity is very small. In the case of $\theta < 10^{-7}$ the effective capacity is large and QoS is not strict. Further, it can also be observed that the effective capacity of voice is higher than the

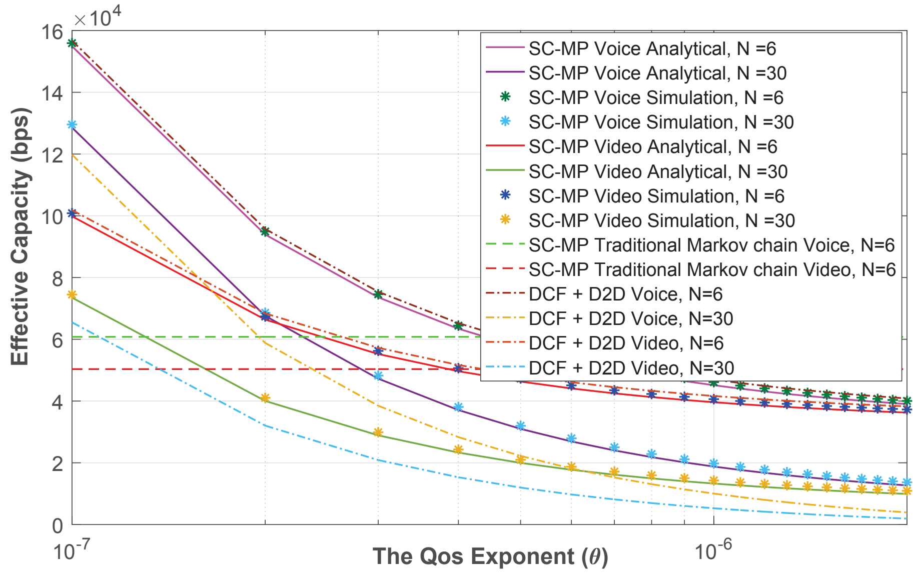


Figure 5.3: Effective Capacity vs The QoS Exponent: P_V=0.3, P_M=0.7, P_N=0.7, and P_D=0.3

video as video requires higher bit rate than voice in a single standard link. We also analytically plot the effective capacity using the existing traditional Markov chain with taking the average time of the traffic of the semi-Markovian model. It is observed that the semi-Markov process is able to capture the stochastic characteristics of voice and video/multimedia packets, while the traditional Markov model cannot. Using Markov model, the effective capacity performance can be biased and cannot capture QoS. In case θ is large the Markov model is over-estimated and in case θ is small the Markov model is under-estimated.

Further, we compare SC-MP to IEEE 802.11 Distributed Coordination Function (DCF) with D2D communication supported. In Fig. 5.3, the effective capacity of DCF with D2D communication is slightly better than SC-MP when the number of users is smaller. As the number of users increases the effective capacity of DCF+D2D decreases, as compared to SC-MP. In a large network, the DCF+D2D MAC protocol can cause high collisions between the users.

In Figure. 5.4, the values of probabilities are changed to further observe the behaviour

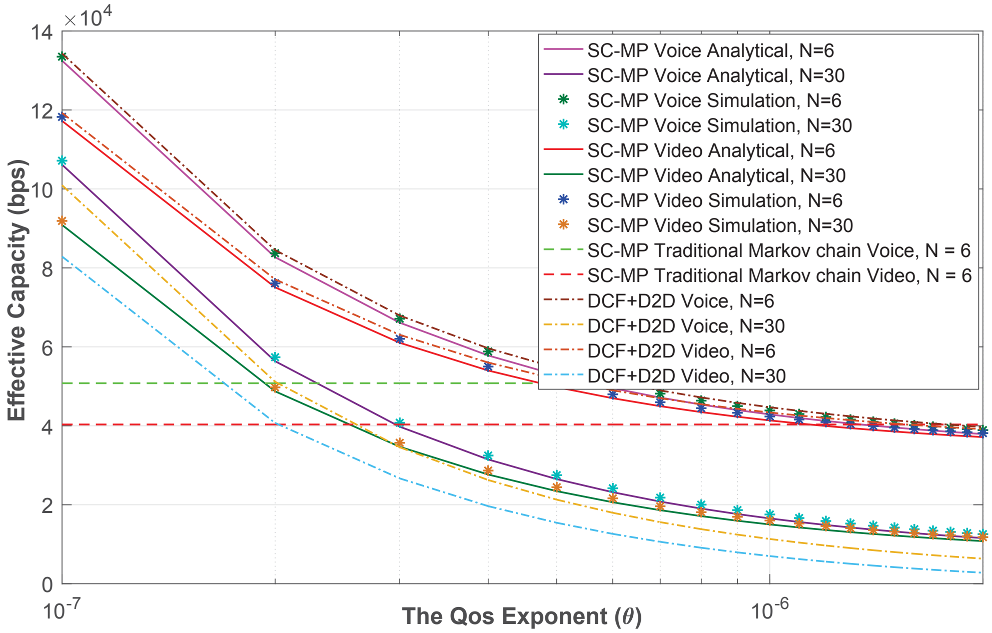


Figure 5.4: Effective Capacity vs The QoS Exponent: P_V=0.45, P_M=0.55, P_N=0.4, and P_D=0.4

of effective capacity with respect to QoS exponent. The effective capacity for P_V=0.45, P_M=0.55, P_D=0.4 and P_N=0.4 is plotted. The traffic probability for voice increases and the probability for video traffic decreases in Fig. 5.4, as compared to Fig. 5.3. By increasing the voice traffic probability, the effective capacity for voice decreases, as compared to Fig. 5.3. Similarly, the video effective increases in Fig. 5.4, as the video traffic probability decreases, as compared to Fig. 5.3. The effective capacity increases with low traffic probability and decreases with high traffic probability. It can also be observed that with the increase of the number of Wi-Fi users also the effective capacity decreases. As the $\theta < 10^{-6}$ there is very little change in effective capacity. Similarly, we plotted the traditional Markovian model as well and there is a slight change in the values and similar result is observed that for large θ the Markov model is over-estimated and for small θ the Markov model is under-estimated. Similarly in Fig. 5.4, the performance of DCF+D2D degrades as the number of users increases as compared to SC-MP.

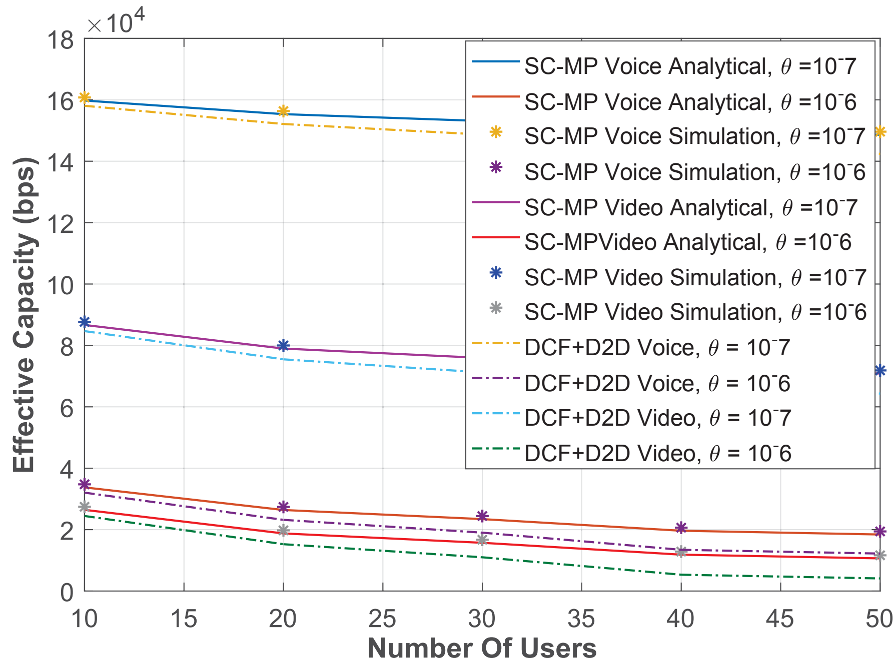


Figure 5.5: Effective Capacity vs Number of Wi-Fi users: P_V=0.3, P_M=0.7, P_N=0.7, and P_D=0.3

In Figures. 5.5 and 5.6, the effective capacity is plotted with the growing number of Wi-Fi users taking $\theta = 10^{-7}$ and 10^{-6} respectively. The effective capacity higher for θ^{-7} than θ^{-6}. The effective capacity decreases with the increase in the number of Wi-Fi users. It is also observed the decrease in effective capacity slows down with increasing number of Wi-Fi users. This is because the effective capacity is less susceptible to the increasing number of Wi-Fi users. The effective capacity decreases as the probability of the traffic increases.

In Figures. 5.5 and 5.6, the effective capacity for voice and video of DCF+D2D decreases rapidly as the number of users increases. SC-MP out performs DCF+D2D in a scalable network, as there is no collision under the PCF access mechanism.

In Figure.5.5, the probability of voice users is less than the video users and probability for D2D communication is very low. The effective capacity for voice is higher than the video as video requires higher bit rate. Whereas, in Figure. 5.6, the effective capacity of voice

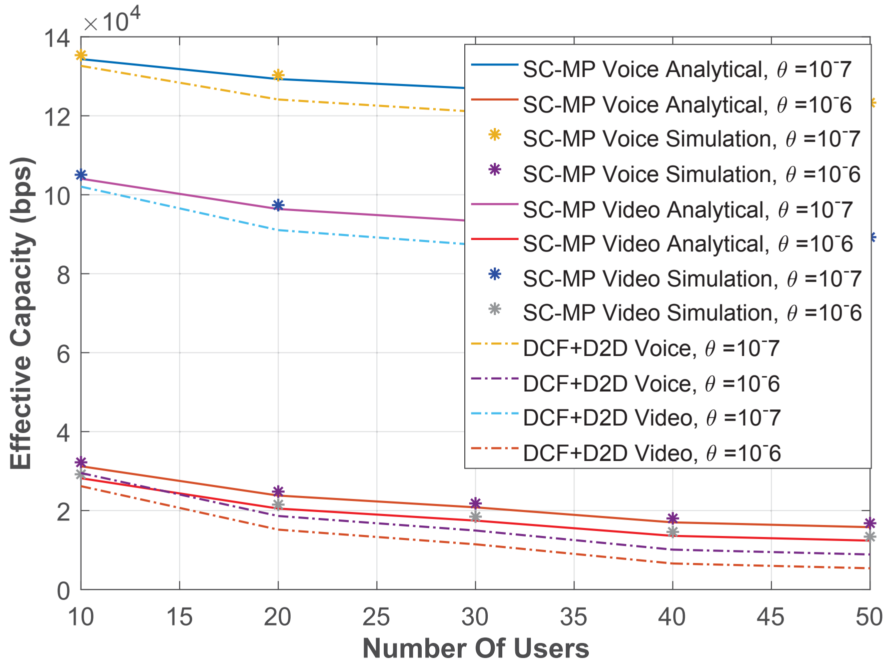


Figure 5.6: Effective Capacity vs Number of Wi-Fi users: P_V=0.45, P_M=0.55, P_N=0.4, and P_D=0.4

decreases with the increase of the probability of voice traffic. Similarly, the video effective capacity increases with the decrease of the video traffic probability and due to D2D communication more number of Wi-Fi users can be accommodated at a particular effective capacity as compared to Figure. 5.5.

In Figure. 5.7, we compare the analytical and simulation results for voice throughput and video throughput. The value of probabilities are P_V=0.3, P_M=0.7, P_D=0.3 and P_N=0.7. The traffic model is constant arrival in Markov modelling. The throughput of the video is higher than than the voice because the probability for video users is more than the voice users. The D2D traffic is low as the probability for D2D channel condition and neighbouring downloaded traffic is low and the video throughput saturates at around 120 users, whereas the total throughput saturates before than 120 users as it will combine the throughput of voice and video.

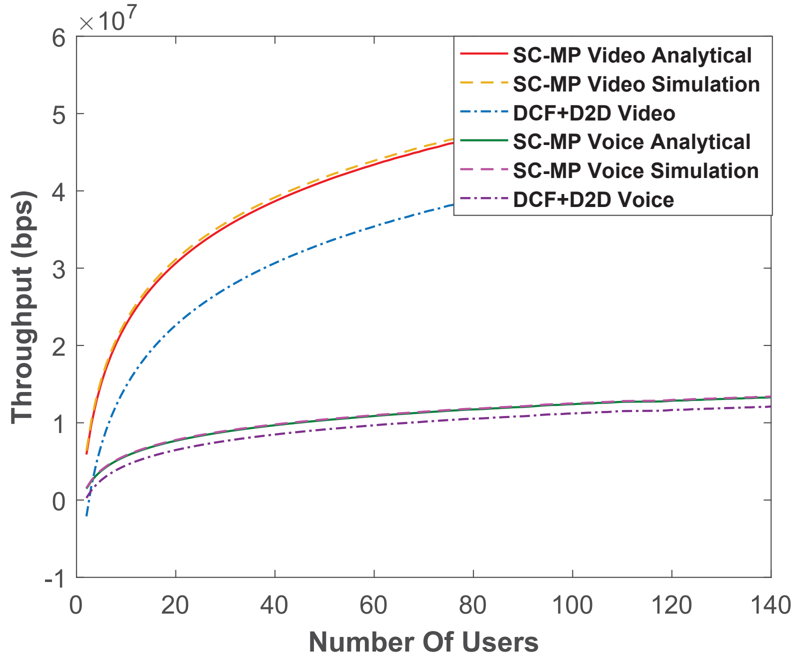


Figure 5.7: Throughput Graph: P_V=0.3, P_M=0.7, P_N=0.7, and P_D=0.3

The throughput of SC-MP for voice and video is higher than DCF+D2D. The throughput of DCF+D2D is lower than the proposed scheme because collisions increasingly occur with the increase of users, where as in the proposed scheme there is no collision as it works under the PCF access mechanism.

5.5 Conclusion

In this chapter, an analytical model is developed to investigate the performance of SC-MP. A three-state semi-Markovian model is developed to analyze the effective capacity and a closed-form expressions is derived from establishing the connections between the effective capacity, channel conditions and transmission durations in a polling based network. Further the three-state semi-Markovian model is modified to four-state traditional Markovian model using discrete Markov chain to analyze the throughput and the behaviour of D2D communication in the network regardless of QoS. The analytical results were validated using simulations. Our investigation gives a good understanding of PCF with D2D communication for a scalable Wi-Fi network.

The work in this chapter also identifies several interesting areas or future work, including implementing of load balancing between LTE and Wi-Fi network on the basis of QoS and a detailed examination of effective capacity in scenarios using more complex mobility models. In Chapter 6, the derived effective capacity for the unlicensed band in Chapter 5 is used to obtain a solution for optimal queue scheduling and resource allocation problem under various statical delay constraints of a three-tier network.

Chapter 6

PCF-Based LTE Wi-Fi Aggregation for Coordinating and Offloading the Cellular Traffic to D2D Network

6.1 Introduction

In this chapter, optimal queuing scheduling and resource allocation problem for three-tier heterogeneous network based on LTE Wi-Fi aggregation (LWA), to offload diverse traffic; voice and multimedia, from licensed band to unlicensed band using Scalable MAC Protocol (SC-MP) under various static delay constraints. The access mechanism used for Wi-Fi in SC-MP is Point Coordination Function (PCF), which further offloads the multimedia traffic using D2D communication in unlicensed band. The proposed scheme builds on and further extends the approach proposed in [183].

The key contributions can be summarized as:

- A three-tier network based on LWA technology is proposed to offload diverse traffic from licensed to unlicensed band with modification in resource allocation scheme based on IEEE 802.11 PCF to access Wi-Fi channels and further offload multimedia files through D2D communication. The optimal joint queue scheduling and resource allocation problem of three-tier network is formulated to minimize the bandwidth of licensed band by guaranteeing the QoS.

- An iterative algorithm is proposed to convexify the problem using generic block coordinated descent (BCD) and difference of convex functions (D.C) program.

- Simulation is performed using two scenarios of the proposed schemes; in the first scenario the voice traffic uses licensed band whereas multimedia traffic uses unlicensed band and in the second scenario half of the voice traffic goes through licensed band and other half of the traffic goes to unlicensed band along with the multimedia traffic. The results are compared with an existing scheme SMS and scheme proposed in [183].

The chapter is organized as follows. In Section 6.2, foundation and related works are discussed and surveyed. In Section 6.3, the system model is presented. In Section 6.4, the closed form expression for the unlicensed band and the licensed band is derived. In Section 6.5, the optimal joint queue scheduling and resource allocation problems with the QoS guarantee between licensed and unlicensed band are formulated to minimize the usage of licensed bandwidth of three-tier heterogeneous network. A new iterative algorithm is proposed based on BCD and DC programs to solve the problem. In Section 6.6, the simulation results are shown, and finally, conclusion is presented in Section 6.7.

6.2 Related Work

In this section, related works are discussed. In [184], LWA analytical model was proposed using Markovian model. In [185], convex optimization technique was used to optimize the power of licensed band and time duration of the unlicensed band to maximize the total algorithmic utility of users. In [186], a linear program technique was used to optimize the licensed bandwidth allocation and rate allocation in the unlicensed band to maximize the overall throughput. In [187], a cross-system learning approach is proposed to optimize the power, cell range expansion bias, sub-band selection and traffic scheduling and the delay-tolerant traffic is steered to unlicensed band. The access mechanism used in Wi-Fi in [185–187] is DCF which is contrary to our work as we will be using PCF as an access mechanism and further offloading the traffic using D2D communication, which is very useful for the dense environment as assigning QoS.

Some related work have been focused on the QoS of the queues transmitted through a single interface. In [170], a semi-Markovian model was proposed for Wi-Fi based on distributed coordinated function (DCF) access channel mechanism. It provides for an asymptotically tight linkage between source characteristics, system resources, and QoS. In [188], a unified framework for LWA was proposed based on QoS class indicator, but there was no analytical model. A non-trivial effort would be required to extend these works to multiple heterogeneous interfaces, as proposed in this paper. In [189], Karush-Kuhn-Tucker technique was used to maximize the effective capacity of the mobile video traffic. In [174,175], the effective capacity is focused on the modelling of rate fluctuations at the physical layer. In [183], author proposes to offload traffic from licensed band to unlicensed band using LWA technology under delay constraints and also derive effective capacity using semi-Markovian model. Our work differs from [183] as we will be introducing a three tier network to offload voice/multimedia traffic from licensed to unlicensed band with modification in resource allocation scheme based on IEEE 802.11 PCF to access Wi-Fi channels and further offload multimedia files through D2D communication. We can conclude that our proposed scheme is scalable and perform better with no contention problem in it.

This chapter addresses an analytical model of three-tier heterogeneous network using Scalable MAC protocol [32] for Wi-Fi network, to evaluate an optimal joint queue scheduling and resource allocation problem under various static delay constraints. The simulation results prove that the proposed scheme has significant performance gain as compared to the state of art.

6.3 System Model

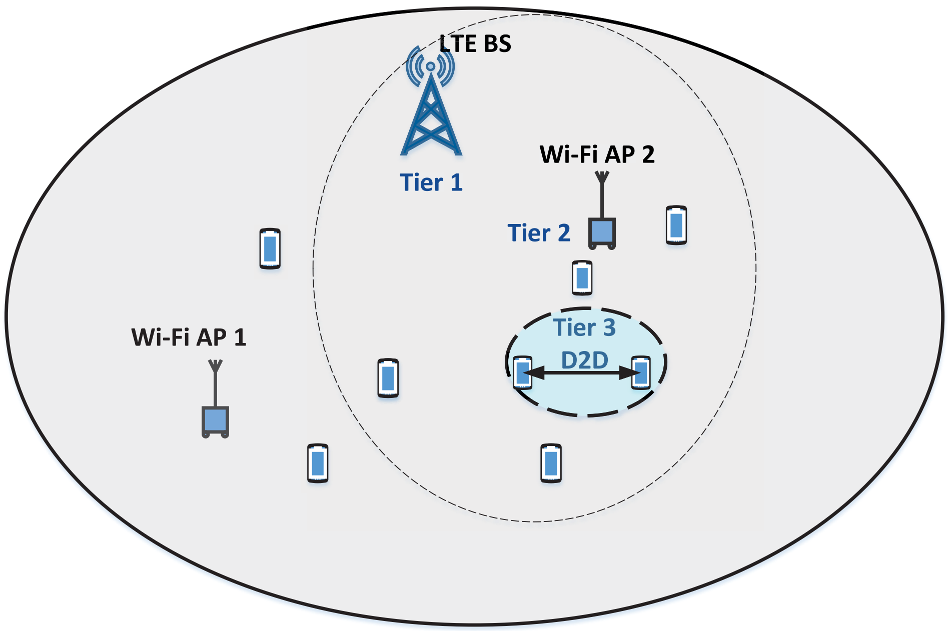


Figure 6.1: Proposed Three-Tier Heterogeneous Network

The proposed architecture builds on the LTE-Wi-Fi aggregation (LWA) technology with modification in resource allocation scheme based on IEEE 802.11 PCF to access WLAN channel and D2D communication itegration. The Figure. 6.1, shows the architecture of proposed three-tier network.

A macro LTE base station (BS) is considered with an LTE air interface in the licensed band overlayed by a Wi-Fi base station (BS) with Wi-Fi air interface in the unlicensed band. We have also considered another Wi-Fi node operating in an unlicensed band within the coverage of three-tier network. Wi-Fi AP 1 and Wi-Fi AP 2 are static, causing a static interference between them. Two frequency bands are considered; band a is the licensed bandwidth and band b is the unlicensed bandwidth. The total bandwidth is given as B_n where n = a, b. Both voice and multimedia traffic is taken into account and licensed and unlicensed bandwidth are further sub-divided. Let B_{av} and B_{am} be the licensed bandwidth for voice and multimedia and B_{bv} and B_{bm} be the unlicensed bandwidth for voice and multimedia. There are total K number of users, where $\gamma_{k,n}$ represent the signal-to-interference-noise ratio (SINR) of a user k in band n and $k = 1, \ldots, K$. The packets are delivered to the users either using licensed LTE air interface or using unlicensed Wi-Fi air interface.

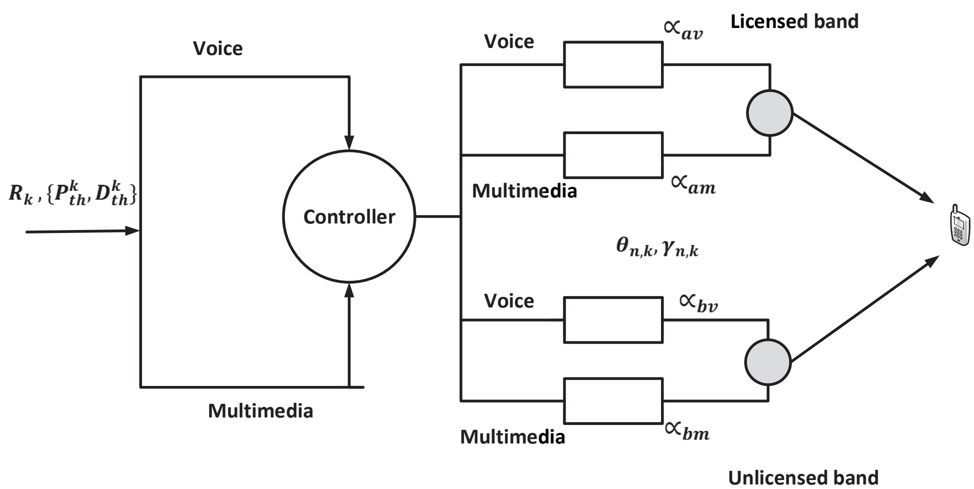


Figure 6.2: Queuing Model Of Three-Tier Network

The Wi-Fi users operating in the unlicensed band use Scalable MAC Protocol (SC-MP) based on point coordination function (PCF) to access the channel [32]. D2D communication is further applied for caching video/multimedia files using CP in PCF [32].

Rayleigh block flat-fading channels is assumed in both licensed and unlicensed bands. The channel remains unchanged during a time frame T, but can vary independently across different time frames. The packets arriving at the scheduler for the user k can be scheduled to proceed using either licensed LTE air interface or unlicensed Wi-Fi air interface. Two transmit queues are formed for the two air interface as shown in the Figure. 6.2. A binary variable $y_{n,k}$ is defined to select the band for the packets. The bandwidth allocated to a user with binary variable $y_{n,k}$ is $\alpha_{n,k}$.

The QoS exponent θ_k is the QoS packets intended for the user k. The larger the value of θ_k, the more stringent is the QoS and the smaller the value of θ_k the more loose is the QoS. If both air interfaces are selected, we propose to decouple θ_k between the two interfaces. QoS exponent of user k in band n is denoted as $\theta_{n,k}$. We propose to precisely design the binary variable $y_{n,k}$, bandwidth $\alpha_{n,k}$ and QoS $\theta_{n,k}$ such that the maximum number of packets can be delivered without compromising θ_k for all the users.

We also propose that each user n requires a minimum data rate of R_k, a delay bound of D_{th}^k and the maximum probability threshold of the delay bound being violated is P_{th}^k. The effective capacity $C_{n,k}(\theta_{n,k})$, can be defined to the maximum consistent arrival rate at the input of the transmit queue for user k in band n, as given by [174, 178];

$$C_{n,k}(\theta_{n,k}) = -\lim_{t\to\infty} \frac{1}{\theta_{n,k}t} \log(\mathbb{E}\{e^{-\theta_{n,k}S_{n,k}(t)}\}) \tag{6.1}$$

where $S_k(t)$ is the number of bits successfully delivered to Wi-Fi user k during (0,t] and $\mathbb{E}\{.\}$ denotes expectation.

According to the effective capacity theory [189], the delay-bound violation probability of the transmit queue in each individual band can be approximated as;

$$Pr\{D > D_{th}^k\} \approx e^{-\theta_{n,k}C_{n,k}(\theta_{n,k})D_{th}^n}, \forall k \in K, \forall n \in N \tag{6.2}$$

Overall, the delay-bound violation probability needs to satisfy;

$$\frac{\sum_{n\in N} y_{n,k} e^{-\theta_{n,k}C_{n,k}(\theta_{n,k})D_{th}^k} C_{n,k}(\theta_{n,k})}{\sum_{n\in N} x_{n,k}C_{n,k}(\theta_{n,k})} \leq P_{th}^k, \forall k \in K, \tag{6.3}$$

where $\sum_{n\in N} y_{n,k} e^{-\theta_{n,k} C_{n,k}(\theta_{n,k}) D_{th}^{k}} C_{n,k}(\theta_{m,n})$ accounts for the total number of packets delivered to user k before the delay bound through the two bands, and $\sum_{n\in N} y_{n,k} C_{n,k}(\theta_{n,k})$ is the total number of packets delivered to user.

6.4 Effective Capacity of LTE and Scalable MAC Protocol (SC-MP)

In this section we will discuss the effective capacity of SC-MP operating in unlicensed band and LTE operating in licensed band, while preserving the QoS.

6.4.1 Effective Capacity Of SC-MP in unlicensed band

The effective capacity of SC-MP in unlicensed band is evaluated in Chapter 5 and is given as;

$|\mathbf{H}(-\theta,-\theta C)-\phi(-\theta,-\theta C)\mathbf{I}| =$

$$\begin{vmatrix} -\phi(-\theta,-\theta C) & M_1(-u)P_{A1} & M_1(-u)P_{F1} & M_1(-u)P_{E1} & 0 & 0 \\ 0 & -\phi(-\theta,-\theta C) & 0 & M_2(Rs-u) & 0 & 0 \\ 0 & 0 & -\phi(-\theta,-\theta C) & M_3(Rs-u) & 0 & 0 \\ M_4(-u)P_{E2} & 0 & 0 & -\phi(-\theta,-\theta C) & M_4(-u)P_{A2} & M_4(-u)P_{F2} \\ M_5(Rs-u) & 0 & 0 & 0 & -\phi(-\theta,-\theta C) & 0 \\ M_6(Rs-u) & 0 & 0 & 0 & 0 & -\phi(-\theta,-\theta C) \end{vmatrix} \tag{6.4}$$

$$= \phi(-\theta,-\theta C)^6 - \phi(-\theta,-\theta C)^4 M_1(\theta C)M_4(\theta C) - \phi(-\theta,-\theta C)^3 M_1(\theta C)M_4(\theta C)X - \\ \phi(-\theta,-\theta C)^2 M_1(\theta C)M_4(\theta C)Y \tag{6.5}$$

where,

$$X = M_2(-R\theta+\theta C)P_{A1}P_{E2} + M_3(-R\theta+\theta C)P_{F1}P_{E2} + \\ M_5(-R\theta+\theta C)P_{E1}P_{A2} + M_6(-R\theta+\theta C)P_{E1}P_{F2}$$

$$Y = P_{A1}P_{A2}M_2(-R\theta+\theta C)M_5(-R\theta+\theta C) + P_{A1}P_{F2} \\ M_2(-R\theta+\theta C)M_6(-R\theta+\theta C) + P_{A2}P_{F1} \\ M_3(-R\theta+\theta C)M_5(-R\theta+\theta C) + P_{F2}P_{F1} \\ M_3(-R\theta+\theta C)M_6(-R\theta+\theta C)$$

We substitute the eigen value $\lambda = \phi(-\theta,-\theta C) = 1$ in (6.5).

The closed form expression of effective capacity for voice and multimedia of 2 users of SC-MP in unlicensed band is given as;

$$C_{bv}(\alpha_{bv},\theta_b) = \frac{z}{\theta_b} + \alpha_{bv} log(1+\overline{\gamma_b}) \tag{6.6}$$

$$C_{bm}(\alpha_{bm},\theta_b) = \frac{z}{\theta_b} + \alpha_{bm} log(1+\overline{\gamma_b}) \tag{6.7}$$

where $\overline{\gamma_b}$ is signal-to-interference-noise-ratio (SINR) of user in unlicensed band and z z is a constant value and varies with the value of probabilities defined, that is; P_V, P_M, P_D and P_N for every user. We can be calculate the value of z by inserting the values of P_V, P_M, P_D and P_N in equation 6.5.

6.4.2 Effective Capacity of LTE in Licensed Band

The effective capacity of LTE in licensed band of user k is given by [183, 190];

$$C_{a,k}(\alpha_{a,k},\theta_{a,k}) = -\frac{1}{\theta_{a,k}T}\log(\mathbb{E}_{\gamma}\{e^{-\theta_{a,k}\alpha_{a,k}T\log_2(1+\overline{\gamma}_{a,k})}\}) \tag{6.8}$$

where $\overline{\gamma}_{a,k}$ is the signal-to-interference-noise ratio of user (SINR) k in licensed band.

The closed form expression of effective capacity for voice and multimedia of LTE in licensed band is given as:

$$C_{av,k}(\alpha_{av,k},\theta_{a,k}) = -\frac{1}{\theta_{a,k}T}\log(\mathbb{E}_{\gamma}\{e^{-\theta_{a,k}\alpha_{av,k}T\log_2(1+\overline{\gamma}_{a,k})}\}) \tag{6.9}$$

$$C_{am,k}(\alpha_{am,k},\theta_{a,k}) = -\frac{1}{\theta_{a,k}T}\log(\mathbb{E}_{\gamma}\{e^{-\theta_{a,k}\alpha_{am,k}T\log_2(1+\overline{\gamma}_{a,k})}\}) \tag{6.10}$$

6.5 Minimizing the Bandwidth Of Licensed Band

The important objective of this section is to minimize the requirement of licensed bandwidth in three-tier heterogeneous network while guaranteeing the QoS for all users. For current problem formulation, we have assumed that voice traffic uses both licensed and unlicensed band, whereas multimedia traffic goes through the unlicensed band to minimize the usage of licensed band. In next Section 6.6, we compare this scenario with another scenario where voice traffic goes to licensed band and multimedia traffic goes through unlicensed band and observe their performance.

According to [183], we can also formulate the problem in our scheme as

$$\mathbf{P1} : \underset{\alpha_{n,k},\theta_{n,k},y_{n,k}}{minimize} \sum_{k\in K} y_{n,k}\alpha_{n,k} \tag{6.11a}$$

$$s.t, \frac{\sum\limits_{n\in N} y_{n,k} e^{-\theta_{n,k} C_{n,k}(\theta_{n,k}) D^k_{th}} C_{n,k}(\theta_{n,k})}{\sum\limits_{n\in N} y_{n,k} C_{n,k}(\theta_{n,k})} \leq P^k_{th}, \forall k \in K \tag{6.11b}$$

$$\sum_{k\in K} y_{b,k}\alpha_{b,k} \leq B_b \tag{6.11c}$$

$$\sum_{n\in N} y_{n,k} C_{n,k}(\alpha_{n,k}, \theta_{n,k}) \geq R_k, \forall k \in K \tag{6.11d}$$

$$y_{n,k} = 0, 1, \forall k \in K, \forall n \in N \tag{6.11e}$$

$$\alpha_{n,k} \geq 0, \theta_{n,k} \geq 0, \forall k \in N, \forall n \in N \tag{6.11f}$$

where (6.11b) describes to meet the requirement of QoS of user k, (6.11c) states the total unlicensed bandwidth that should not exceed B_b, (6.11d) represents the minimum data rate of the user k, , (6.11e), and (6.11f) are the generic restraints to specify the problem.

It can be observed that from **P1**, $e^{-\theta_{n,k} C_{n,k} D^k_{th}}$ and $y_{n,k} C_{n,k}$ are coupled and the effective capacity in (6.11d) is not joint convex on $\alpha_{n,k}$ and $\theta_{n,k}$.

As $y_{n,k}$ being a binary variable and according to [183], it is said that

$$y_{n,k} C_{n,k}(\alpha_{n,k}, \theta_{n,k}) = C_{n,k}(y_{n,k}\alpha_{n,k}, \theta_{n,k}) \tag{6.12}$$

However, **P1** is a combinational mixed integer program. The effective capacity in (6.11d) is not joint convex on $\alpha_{n,k}$ and $\theta_{n,k}$ as $e^{-\theta_{n,k} C_{n,k} D^k_{th}}$ and $y_{n,k} C_{n,k}$ are coupled in multiplicative way.

Using Chebyshev's sum inequality [191], (6.11d) can be written as

$$\frac{1}{|N|}\sum_{n\in N} e^{-\theta_{n,k}C_{n,k}(\theta_{n,k})D^k_{th}} \sum_{n\in N} C_{n,k} \leq P^k_{th} \sum_{n\in N} C_{n,k} \tag{6.13}$$

where $|N|$ stands for cardinality. If a packet is transmitted to user k in band n the (6.11b) can be rewritten as ;

$$e^{-\theta_{n,k}C_{n,k}(\theta_{n,k})D^k_{th}} \leq P^k_{th} \tag{6.14}$$

Combining (6.13) and (6.14), (6.11b) can be rewritten as

$$\sum_{n\in N} e^{-\theta_{n,k}C_{n,k}(\theta_{n,k})D^k_{th}} - 1 + y_{n,k} \leq P^k_{th} \sum_{n\in N} y_{n,k} \tag{6.15}$$

The (6.11e), can be relaxed as the intersection of the following region [192];

$$0 \leq y_{n,k} \leq 1, \forall k \in K, \forall n \in N \tag{6.16}$$

$$\sum_{n\in N}\sum_{k\in K}(y_{n,k} - (y_{n,k})^2) \tag{6.17}$$

We state two variables for simplification; $\omega_{n,k} = \alpha_{n,k}\theta_{n,k}$ and $\beta_{n,k} = \frac{1}{\theta_{n,k}}$, where $\alpha_{n,k} = \omega_{n,k}\beta_{n,k}$. The effective capacity of SC-MP and licensed band in (6.6), (6.7), and (6.9) can be rewritten in terms of auxiliary variables as

$$C_{av,k}(\omega_{av,k}, \beta_{a,k}) = -\frac{\beta_{a,k}}{T} log(\mathbb{E}_\gamma\{e^{-\omega_{av,k}Tlog_2(1+\gamma_a)}\}) \tag{6.18}$$

$$C_{bv}(\alpha_{bv}, \beta_b) = z\beta_b + \omega_{bv}\beta_b log_2(1+\gamma_b) \tag{6.19}$$

$$C_{bm}(\alpha_{bm}, \beta_b) = z\beta_b + \omega_{bm}\beta_b log_2(1 + \gamma_b) \quad (6.20)$$

Using (6.18), (6.19) and (6.20), (6.15) can be rewritten as

$$e^{\frac{log(\mathbb{E}_\gamma\{e^{-\omega_{av,k} T log_2(1+\gamma_a)}\})}{T} D^k_{th}} + e^{-(2z+\omega_{bv}\beta_b log_2(1+\gamma_b)+\omega_{bm}\beta_b log_2(1+\gamma_b) D^k_{th}} - 2 + \sum_{n\in N} y_{n,k} \leq P^k_{th} \sum_{n\in N} y_{n,k}, \forall k \in K \quad (6.21)$$

We will also relax the binary constraint (6.11f) as the intersection of the following region [192];

$$0 \leq y_{n,k} \leq 1, \forall k \in K, \forall n \in N \quad (6.22)$$

$$\sum_{n\in N}\sum_{k\in K}(y_{n,k} - (y_{n,k})^2) \quad (6.23)$$

As an outcome and according to [183], **P1** can be simplified to be a continuous problem and can be written as

$$\mathbf{P2}: \underset{\{\omega_{n,k}\},\{\beta_{n,k}\},\{y_{n,k}\}}{minimize} \sum_{k\in K} \omega_{a,k}\beta_{a,k} \quad (6.24a)$$

$$e^{\frac{log(\mathbb{E}_\gamma\{e^{-\omega_{av,k} T log_2(1+\gamma_a)}\})}{T} D^k_{th}} + e^{-(2z+\omega_{bv}\beta_b log_2(1+\gamma_b)+\omega_{bm}\beta_b log_2(1+\gamma_b) D^k_{th}} - 2 + \sum_{n\in N} y_{n,k} \leq P^k_{th} \sum_{n\in N} y_{n,k}, \forall k \in K \quad (6.24b)$$

$$\sum_{k\in K}(\omega_{bv,k}+\omega_{bm,k})\beta_{b,k}\leq B_b \tag{6.24c}$$

$$s.t, \sum_{n\in N} C_{n,k}(\omega_{n,k},\beta_{n,k})\geq R_k, \forall k\in K \tag{6.24d}$$

$$\sum_{n\in N} y_{n,k}\geq 1, \forall k\in K \tag{6.24e}$$

$$\sum_{n\in N}\sum_{k\in K}(y_{n,k}-(y_{n,k})^2) \tag{6.24f}$$

$$\omega_{n,k}\geq 0, \beta_{n,k}\geq 0, 0\leq y_{n,k}\leq 1, \forall n\in N, \forall k\in K \tag{6.24g}$$

$$0\leq \omega_{a,n}\beta_{a,n}\leq y_{n,k}\chi, \forall k\in K, \forall n\in N \tag{6.24h}$$

As (6.24a), (6.24c), and (6.24d) are affine, therefore **P2** is linear on $\{\beta_{n,k}\}$.

Using difference of convex (DC) programming, **P2** can be reformulated using **P3**

$$\begin{aligned}\textbf{P3:} \underset{\{\omega_{n,k}\},\{y_{n,k}\}}{minimize} \sum_{k\in K}\omega_{a,k}\beta_{a,k}+\lambda\sum_{n\in N}\sum_{k\in K}y_{n,k}-\\ \lambda\sum_{n\in N}\sum_{k\in K}(y_{n,k})^2,\end{aligned} \tag{6.25}$$

$$s.t(6.24b)-(6.24d),(6.24f)-(6.24h), \tag{6.26}$$

where λ is a large penalty factor.

Let;

$$f_1(\omega_{n,k}, y_{n,k}) = \sum_{k\in K} \omega_{a,k}\beta_{a,k} + \lambda \sum_{n\in N}\sum_{k\in K} y_{n,k} \tag{6.27}$$

$$f_2(y_{n,k}) = \lambda \sum_{n\in N}\sum_{k\in K}(y_{n,k})^2 \tag{6.28}$$

$$f_1(\omega_{n,k}, y_{n,k}) - f_2(y_{n,k}) \tag{6.29}$$

where (6.29) is the difference of two convex functions.

According to [193], **P3** is equivalent to **P2** for large value of λ.

An algorithm is summarized based on a block coordinated descent (BCD) framework [194]. In the algorithm we initial number of users K, minimum data rate for each user R_k, signal-to-noise ratio of user $\gamma_{n,k}$, requirement for QoS, and total unlicensed bandwidth B_b. The algorithm consist of 2 loops. In the inner loop, given $\{\omega_{n,k}\}$ and $\{y_{n,k}\}$, **P2** is linear programming on $\{\beta_{n,k}\}$, and can be solved efficiently using interior-point method [195] and has a complexity of $O((2XY)^3(3Y+X+XY))$. In the outer loop, given $\{\beta_{n,k}\}$, **P2** is a DC program in $\{\omega_{n,k}, y_{n,k}\}$ and has a complexity of $O((XY)^3)$. The total complexity of algorithm 1 is $O((XY)^6(3Y+X+XY))$ based on [192].

6.6 Simulation

In this section, we evaluate the performance of the proposed algorithm through simulation. There is one LTE Base station (LTE-BS) over-layed by WLAN BS (WBS). In our proposed scheme the WLAN BS contend the channel with the existing Wi-Fi systems before it can poll the users. We have also considered another Wi-Fi node operating in an unlicensed band within the coverage of three-tier network. The channel model used for licensed and unlicensed bands are the ITU-UMi Models. We have assumed in total 10 users uniformly distributed and all have the same minimum data rate requirement and QoS requirement. We assumed the time frame length of LTE T = 1 ms.

We have used both voice and multimedia traffic, for voice we used persistent scheduling and for video we use adaptive scheduling. We have taken two different scenarios for our proposed scheme. In scenario 1, the voice traffic is assigned to the licensed band and the multimedia traffic is assigned to the unlicensed band. In scenario 2, half of the voice traffic goes on the licensed band and other half will go through the unlicensed band along with the multimedia traffic. The proposed scenarios are also compared with the existing algorithm, Static mapping scheme (SMS) and scheme proposed in [183], to evaluate the performance.

Static Mapping Scheme (SMS): In this scheme, a static mapping table is established where a design parameter γ ($0 \leq \gamma \geq 1$) is decided according to the QoS Class Indicator (QCI) or the types of traffics. In our simulations, we set $\gamma = 0.6$.

Cellular users are first ordered in descending order in terms of SINR of the unlicensed band. The LWA BS sequentially assigns the unlicensed bandwidth and the QoS exponent to the ordered users by solving the following equations

$$e^{-\theta_{b,k} C_{b,k} D_{th}^{k}} = P_{th}^{k} \tag{6.30}$$

$$C_{b,k}(\alpha_{b,k}, \theta_{b,k}) = \frac{z}{\theta_{b,k}} + \alpha_{b,k} log(1 + \overline{\gamma_{b,k}}) \tag{6.31}$$

LWA BS will keep on assigning the unlicensed bandwidth until the total allocated bandwidth reaches the total unlicensed bandwidth or all the QoS users are satisfied. If the unlicensed band is insufficient to satisfy the QoS of the users, the rest of the traffic goes to the licensed band. The remaining users are ordered in descending SINR of the licensed band. The LWA BS sequentially allocate the licensed bandwidth and the QoS exponent to the ordered users by solving the following equations,

$$e^{-\theta_{a,k} C_{a,k} D_{th}^{k}} = P_{th}^{k} \tag{6.32}$$

6.6.1 Performance evaluation for minimizing the bandwidth of licensed band

In Figure. 6.3, we investigate the bandwidth of licensed band with the delay requirements, where P_V=0.3, P_M=0.7, P_D=0.6 and P_N=0.4. There are in total 10 users within the coverage area of LTE and Wi-Fi. The figure shows that as the delay bound increases the required bandwidth of the licensed band decreases. It can also be observed that the our proposed scenarios performs better in terms of delay as compared to the existing scheme in [183] and SMS scheme. Also, proposed scenario 2 can perform much better than scenario 1, when the voice traffic is split between the licensed and unlicensed band.

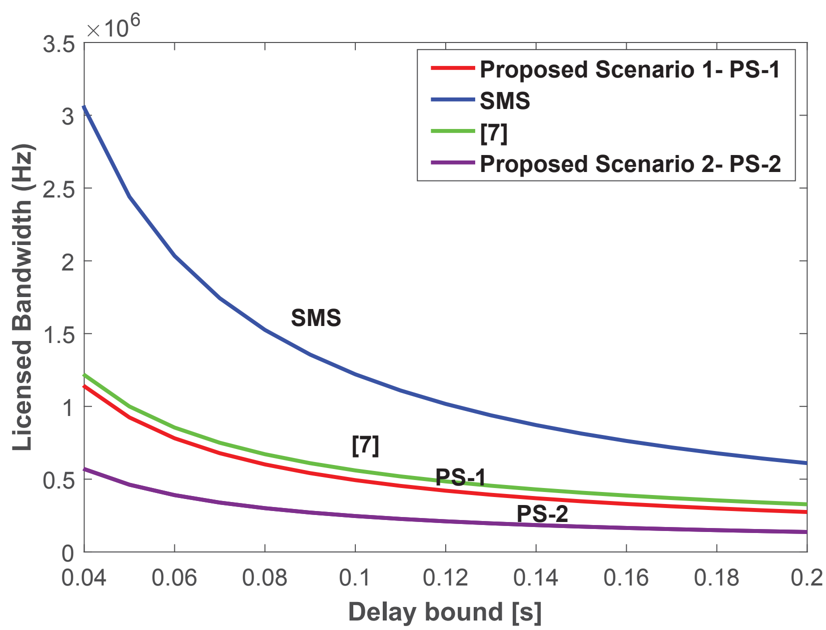


Figure 6.3: Licensed Bandwidth vs Delay bound: P_V=0.3, P_M=0.7, P_N=0.4, and P_D=0.6

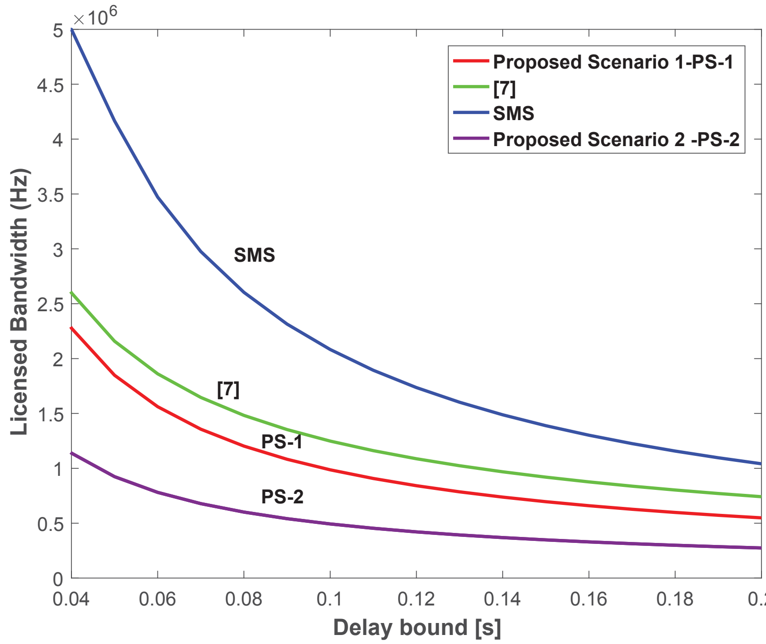


Figure 6.4: Licensed Bandwidth vs Delay Bound: P_V=0.6, P_M=0.4, P_N=0.4, and P_D=0.4

In Figure. 6.4, the values of probabilities are changed to further observe the behaviour of licensed bandwidth with the delay requirements. The Fig. 5 is plotted for P_V=0.4, P_M=0.6, P_D=0.4 and P_{ND}=0.4. It can be observed that by increasing the probability of voice users will increase the requirement of licensed bandwidth as well. Also, our proposed scheme performs better than the existing scheme in [183] and the SMS scheme.

In Figure. 6.5 and 6.6, the required licensed bandwidth is investigated with the number of transmitters by changing the value of probabilities. The required licensed bandwidth increases with the increase of transmitters. In Fig. 6.6, the voice probability is increased and the requirement for licensed bandwidth also increases. Also, it can be observed that our proposed scheme can reduce the licensed bandwidth as compared to the existing scheme in [183] and SMS. Scenario 2 can provide 50% less requirement of licensed band as than scenario 1.

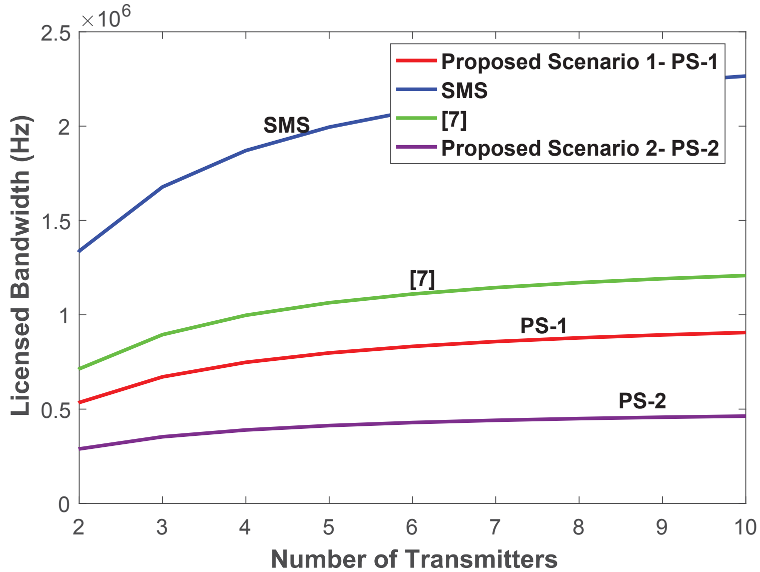


Figure 6.5: Licensed Bandwidth vs Number of Transmitters: P_V=0.3, P_M=0.7, P_N=0.4, and P_D=0.6

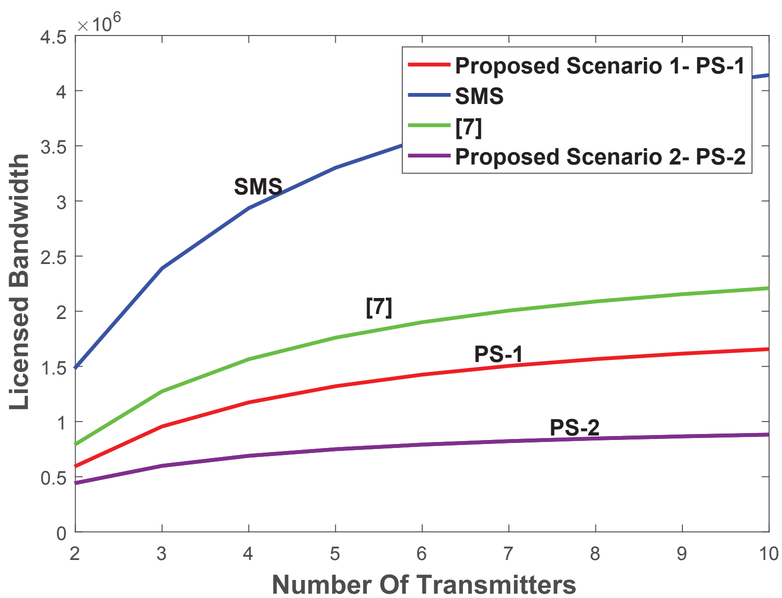


Figure 6.6: Licensed Bandwidth vs Number of Transmitters: P_M=0.6, P_M=0.4, P_N=0.4, and P_D=0.4

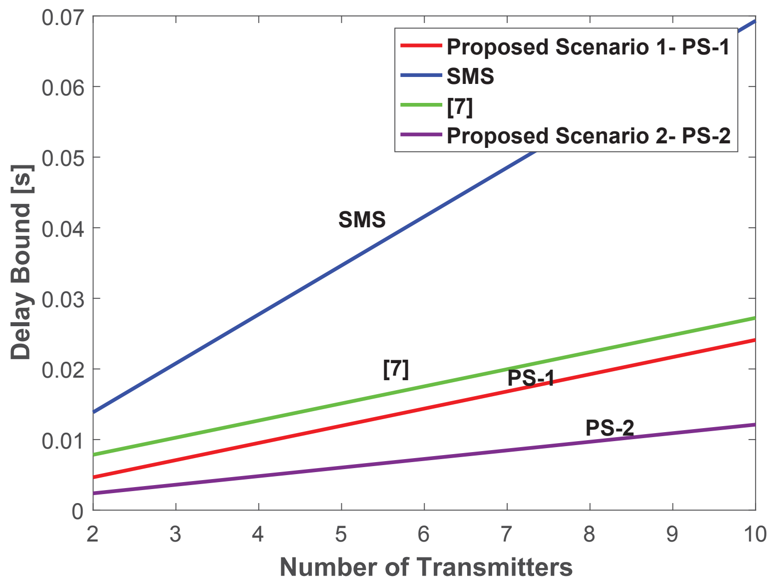


Figure 6.7: Delay Bound vs Number of Transmitters: P_V=0.3, P_M=0.7, P_N=0.4, and P_D=0.6

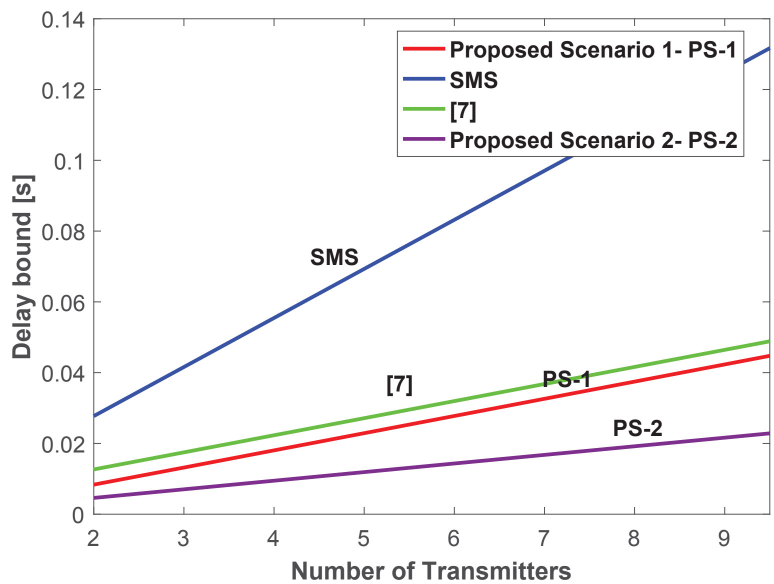


Figure 6.8: Delay Bound vs Number of Transmitters: P_V=0.6, P_M=0.4, P_N=0.4, and P_D=0.4

In Figure. 6.7 and 6.8, we investigate the delay bound with the number of transmitters by changing the probabilities. All users have the same delay threshold $D_{th}^{n} = 0.05$ s. The delay increases as the transmitter increases but our proposed schemes can reduce the delay as compared to the existing scheme in [183] and the SMS scheme. Scenario 2 can further reduce the delay as compared with scenario 1.

6.7 Conclusion

In this chapter, an optimal queue scheduling and resource allocation problem is investigated under various statistical delay constraints of three-tier network based on LWA technology with modification in resource allocation scheme based on IEEE 802.11 PCF to access WLAN channels and further offload multimedia files through D2D communication. a closed-form expression of effective capacity for the unlicensed band is derived using semi-Markovian process. Secondly, we formulate the optimal joint queue scheduling and resource allocation problem with the QoS guarantee between licensed and unlicensed band to minimize the bandwidth of licensed band. A new iterative algorithm is proposed to convexify the problem as a series of block coordinated descent (BCD) and difference of convex functions (D.C) program.

Further, an extensive simulation was carried out with the proposed scheme with two scenarios, the first scenario had all the voice traffic on the licensed band and multimedia traffic traffic on the unlicensed band and the second scenario splitted the voice traffic to licensed and unlicensed band whereas all the multimedia traffic goes to unlicensed band. The simulation results showed that the proposed scheme can provide improved performance than SMS and existing scheme in [183]. Also scenario 2 can perform 50% better than scenario 1.

Chapter 7

Conclusion and Future Works

7.1 Conclusion

In the next decade, the tremendous growth of data traffic volume poses a significant challenge for current cellular networks and is one of the main reason for the next generation of mobile networks, referred to as the fifth generation (5G). There are several solutions proposed for 5G that can increase the efficiency of the available resources or aim at providing new radio resources or infrastructures. The primary factors that will help to enhance the capacity and coverage in 5G include; device-to-device communication, heterogeneous networks, and software defined networks. Device-to-Device communication in heterogeneous networks is a good example to increase the capacity and coverage in a dense environment, by which users can communicate directly without the involvement of base station. The merger of different networks and D2D communication will be very beneficial to further enhance the performance of the network and hence allow us to look towards 5G.

In this thesis, a new Scalable MAC Protocol based on IEEE 802.11 PCF access mechanism is proposed for a three-tier network to increase the coverage and capacity in a dense environment while decreasing the delay and giving fairness to the users. In this chapter, we give the summary of the results and the contributions.

7.1.1 Literature Review

A comprehensive literature review of heterogeneous network and MAC protocols of D2D communication is presented. Both infrastructures assisted cellular D2D communication and ad-hoc techniques are studied and their scalability is compared. Following conclusion is concluded from the literature review:

- Heterogeneous Networks are a center of attraction during LTE and will be an essential part of 5G networks. CRRM will play an important role in HetNets to offload traffic from a larger cell to a smaller cell. It can bring significant improvement in load balancing, interference distribution, reduction of unnecessary handovers (HOs) and reduction of dropping/blocking probability [26]. It may increase the signaling but a trade-off is required between signaling and performance to get the best results.

- Device-to-Device (D2D) communication will play a vital role to further enhance the capacity and coverage. Although D2D communication was already introduced in LTE-A (LTE advanced) with further improvement, it will be an integral part of 5G networks. In the literature review, different techniques are discussed for resource allocation in D2D communication. Most of the papers discussed in the literature review require tight interference control. It has been assumed that BS is aware of instantaneous CSI of cellular/D2D links e.g [59], [64]. The assumption is required as BS must participate in making scheduling decisions for both cellular and D2D users.

- D2D communication in infrastructure assisted networks are operating under centralized control and a majority of the scheduling is handled by the BS. Although ad-hoc networks are distributed, it is important for D2D operating systems should be combined with ad-hoc networks in order to further enhance the coverage and capacity problems because this can enhance overall performance.

- The MAC protocols in ad-hoc networks based on IEEE 802.11 DCF, PCF and HCF are discussed in detail using different techniques and strategies. Clustering strategy is proposed in [115] using DCF, hence can play a significant role in improving network performance and solving capacity problems. For future generation clustering strategies can also play an important role to fulfill the requirements of capacity, coverage, and QoS.

- To increase capacity in certain areas, D2D communication in HetNets will be the best possible solution especially for a special event like exhibitions, stadium or fairs where a large crowd gathers for a certain period of time. The access mechanism used for WLAN should be PCF as it does not have contention problems and is centralized control which is suitable for dense environment

7.1.2 Performance Investigation of D2D Communication over LTE Solutions Operating in Unlicensed Band

A detailed overview of LTE technologies operating in an unlicensed 5GHz band that includes; carrier Wi-Fi, LTE-U, LTE-LAA, LWA, and MuLTEfire is discussed. These technologies are compared and their advantages and disadvantages are concluded.
Simulation is done for LTE-LAA, LWA, and MuLTEfire in the presence of Wi-Fi hot-spot. From the simulation and results, it is analyzed that MuLTEfire can provide higher throughput but limit the overall scalability of the network coexistence. Consequently, Wi-Fi users have a higher impact in the presence of MuLTEfire and can degrade the performance of Wi-Fi throughput. LTE-LAA has higher throughput for smaller load but the throughput decreases as the load increases and also Wi-Fi throughput degrades. Whereas LWA can provide better throughput than LTE-LAA for a large network and has minimal effect on the

Wi-Fi users as there is no contention between the users. Similarly, when D2D communication is investigated in these techniques MuLTEfire can increase the throughput but cannot increase the users whereas LWA and LTE-LAA can accommodate more users, where LWA showing better throughput. LWA gives better results when a Wi-Fi AP is deployed near it and is more cost-effective and easy to implement as compared to LTE-LAA. For future challenges, it will be very important that all the technologies operating in the 5GHz band should have equal control to access the medium to satisfy fairness.

7.1.3 Scalable MAC Protocol For D2D Communication For Future 5G Networks

An innovative three-tier 5G architecture comprising of LTE BS, WLAN BS and D2D communication is proposed to offload traffic from cellular to WLAN. The proposed architecture builds on the LWA technology with modification in resource allocation scheme based on IEEE 802.11 PCF to access WLAN channel with an additional increase in throughput by D2D communication. A scalable MAC protocol (SC-MP) for D2D communication is proposed based on PCF access mechanism. The importance of PCF access mechanism is that it operates in a centralized manner and highly suitable for the dense environment, hence, can create a centralized control in a distributive manner. The best SNR polling scheme is used for PCF which can take into consideration the channel condition of the user instead of on simple round robin scheme. D2D communication further increases the network throughput by sharing content files between the WLAN users and decreases the delay of the normal users.

The simulation results show that the proposed MAC scheme can increase the capacity of the network and perform better relative to the legacy Distributed Coordination Function (DCF) defined in IEEE 802.11. It is also observed that greater number of users can be accommodated without wasting the resources by applying D2D communication in WLAN.

7.1.4 Analysis of Effective Capacity and Throughput of Polling Based Heterogeneous Networks

An analytical model is developed to investigate the performance of SC-MP based on IEEE 802.11 PCF access mechanism with D2D communication. A three-state semi-Markovian model is developed to analyze the effective capacity and a closed-form expression is derived from establishing the connections between the effective capacity, channel conditions and transmission durations in a polling based network. Further, the three-state semi-Markovian model is modified to four-state traditional Markovian model using discrete Markov chain to analyze the throughput and the behavior of D2D communication in the network regardless of QoS. The analytical results were validated using simulations, hence proving the appropriateness of the model.

7.1.5 Optimal Scheduling and Resource Allocation in LWA-Driven D2D-Enabled Networks

An optimal queue scheduling and resource allocation problem is investigated under various statistical delay constraints of three-tier network based on LWA technology with modification in resource allocation scheme based on IEEE 802.11 PCF to access WLAN channels and further offload multimedia files through D2D communication. The optimal joint queue scheduling and resource allocation problem is formulated with the QoS guarantee between licensed and unlicensed band to minimize the bandwidth of licensed band. A new iterative algorithm is proposed to convexify the problem as a series of block coordinated descent (BCD) and difference of convex functions (D.C) program. Further, the simulation was carried out with the proposed scheme with two scenarios, the first scenario had all the voice traffic on the licensed band and multimedia traffic on the unlicensed band and the second scenario split the voice traffic to licensed and unlicensed band whereas all the multimedia traffic goes to the unlicensed band. The simulation results showed that our proposed scheme can provide improved performance than SMS. Also, scenario 2 can perform 50% better than scenario 1. The results proved that the optimized algorithm has a significant gain compared to the scheme SMS.

7.2 Future Works

In this section, based on the assumptions, results, and observations discussed in this thesis, a number of interesting opportunities for future work are listed below;

- In Chapter 3, the LTE technologies operating in the unlicensed band; LTE-U, LTE-LAA, LWA, and MuLTEfire, are compared through simulation with further applying D2D communication. It will be very interesting to investigate the behaviour of D2D communication in these technologies through testbed and comparing the result with the simulation result. The testbeds can give more realistic and accurate results towards the real life.

- The proposed scheme SC-MP can also be further extended by using multiple WLAN in a multiple cells. The main challenge will be to manage the offloading of traffic and the interference problem. Further, the power levels of the user equipment can be optimized for the D2D communications. The performance gain can be quantified through extensive mobile simulations, including realistic vehicular mobility model with taking energy consumption is account as well.

- The proposed scheme SC-MP can also be further extended by using multiple WLAN in a single cell and finding the best possible position for WLAN to increase the coverage and the capacity. The main challenge will be the interference between the cells.

- The analytical model can further be extended to multi-objective function in which the aggregated effective capacity can be maximized and aggregated delay can be minimized with the QoS guarantee between the licensed and unlicensed band.

Bibliography

[1] mepits. 5G,Mobile Technology. `https://www.mepits.com/tutorial/175/Communication/5G,-Mobile-Technology`, 2014. [Online; accessed 16-Aug-2014].

[2] Mugen Peng, Yong Li, Zhongyuan Zhao, and Chonggang Wang. System architecture and key technologies for 5g heterogeneous cloud radio access networks. *Network, IEEE*, 29(2):6–14, 2015.

[3] Patrick Agyapong, Mikio Iwamura, Dirk Staehle, Wolfgang Kiess, and Anass Benjebbour. Design considerations for a 5g network architecture. *Communications Magazine, IEEE*, 52(11):65–75, 2014.

[4] Cheng-Xiang Wang, Fourat Haider, Xiqi Gao, Xiao-Hu You, Yang Yang, Dongfeng Yuan, Hadi Aggoune, Harald Haas, Sam Fletcher, and Erol Hepsaydir. Cellular architecture and key technologies for 5g wireless communication networks. *Communications Magazine, IEEE*, 52(2):122–130, 2014.

[5] Kishor Chandra, Zizheng Cao, TM Bruintjes, R Venkatesha Prasad, G Karagiannis, Eduward Tangdiongga, HPA van den Boom, and ABJ Kokkeler. mcran: A radio access network architecture for 5g indoor communications. In *Communication Workshop (ICCW), 2015 IEEE International Conference on*, pages 300–305. IEEE, 2015.

[6] Jeffrey G Andrews, Stefano Buzzi, Wan Choi, Stephen V Hanly, Aurelie Lozano, Anthony CK Soong, and Jianzhong Charlie Zhang. What will 5g be? *Selected Areas in Communications, IEEE Journal on*, 32(6):1065–1082, 2014.

[7] John Thompson, Xing-Lai Ge, Hsiao-Chun Wu, Ralf Irmer, Hongbo Jiang, Gerhard Fettweis, and Siavash Alamouti. 5g wireless communication systems: prospects and challenges [guest editorial]. *Communications Magazine, IEEE*, 52(2):62–64, 2014.

[8] Boyd Bangerter, Shilpa Talwar, Reza Arefi, and Kyle Stewart. Networks and devices for the 5g era. *Communications Magazine, IEEE*, 52(2):90–96, 2014.

[9] Woon Hau Chin, Zhong Fan, and Robert Haines. Emerging technologies and research challenges for 5g wireless networks. *Wireless Communications, IEEE*, 21(2):106–112, 2014.

[10] Panagiotis Demestichas, Andreas Georgakopoulos, Dimitrios Karvounas, Kostas Tsagkaris, Vera Stavroulaki, Jianmin Lu, Chunshan Xiong, and Jing Yao. 5g on the horizon: key challenges for the radio-access network. *Vehicular Technology Magazine, IEEE*, 8(3):47–53, 2013.

[11] Shanzhi Chen and Jian Zhao. The requirements, challenges, and technologies for 5g of terrestrial mobile telecommunication. *Communications Magazine, IEEE*, 52(5):36–43, 2014.

[12] wikipedia. MIMO. `https://en.wikipedia.org/wiki/MIMO`, 2015. [Online; accessed Jan-2015].

[13] Frederick W Vook, Amitava Ghosh, and Timothy A Thomas. Mimo and beamforming solutions for 5g technology. In *Microwave Symposium (IMS), 2014 IEEE MTT-S International*, pages 1–4. IEEE, 2014.

[14] Cheng-Xiang Wang, Xuemin Hong, Xiaohu Ge, Xiang Cheng, Gong Zhang, and John Thompson. Cooperative mimo channel models: A survey. *Communications Magazine, IEEE*, 48(2):80–87, 2010.

[15] Beibei Wang and KJ Liu. Advances in cognitive radio networks: A survey. *Selected Topics in Signal Processing, IEEE Journal of*, 5(1):5–23, 2011.

[16] Yong Niu, Yong Li, Depeng Jin, Li Su, and Athanasios V Vasilakos. A survey of millimeter wave communications (mmwave) for 5g: opportunities and challenges. *Wireless Networks*, 21(8):2657–2676, 2015.

[17] Zheng Ma, ZhengQuan Zhang, ZhiGuo Ding, PingZhi Fan, and HengChao Li. Key techniques for 5g wireless communications: network architecture, physical layer, and mac layer perspectives. *Science China Information Sciences*, 58(4):1–20, 2015.

[18] wikipedia. Radio Access and Spectrum FP7 Future Networks Cluster. `http://fp7-semafour.eu/media/cms_page_media/9/SEMAFOUR_2014_RAScluster%20White%20paper.pdf`. [Online; accessed Jan-2015].

[19] Karolina Ratajczak, Krzysztof Bakowski, and Krzysztof Wesolowski. Two-way relaying for 5g systems: Comparison of network coding and mimo techniques. In *Wireless Communications and Networking Conference (WCNC), 2014 IEEE*, pages 376–381. IEEE, 2014.

[20] Riccardo Trivisonno, Riccardo Guerzoni, Ishan Vaishnavi, and David Soldani. Sdn-based 5g mobile networks: architecture, functions, procedures and backward compatibility. *Transactions on Emerging Telecommunications Technologies*, 26(1):82–92, 2015.

[21] Hsin-Hung Cho, Chin-Feng Lai, Timothy K Shih, and Han-Chieh Chao. Integration of sdr and sdn for 5g. *Access, IEEE*, 2:1196–1204, 2014.

[22] Volkan Yazıcı, Ulas C Kozat, and M Oguz Sunay. A new control plane for 5g network architecture with a case study on unified handoff, mobility, and routing management. *Communications Magazine, IEEE*, 52(11):76–85, 2014.

[23] Sushil Singh and Jeffrey G Andrews. Joint resource partitioning and offloading in heterogeneous cellular networks. *Wireless Communications, IEEE Transactions on*, 13(2):888–901, 2014.

[24] Mahamod Ismail, Atef Abdrabou, and Weihua Zhuang. Cooperative decentralized resource allocation in heterogeneous wireless access medium. *Wireless Communications, IEEE Transactions on*, 12(2):714–724, 2013.

[25] Xiaoyu Duan, Xianbin Wang, and Auon Muhammad Akhtar. Partial mobile data offloading with load balancing in heterogeneous cellular networks using software-defined networking. In *Personal, Indoor, and Mobile Radio Communication (PIMRC), 2014 IEEE 25th Annual International Symposium on*, pages 1348–1353. IEEE, 2014.

[26] Leijia Wu and Kumbesan Sandrasegaran. A survey on common radio resource management. In *Wireless Broadband and Ultra Wideband Communications, 2007. AusWireless 2007. The 2nd International Conference on*, pages 66–66. IEEE, 2007.

[27] Lorenza Giupponi, Ramon Agusti, Jordi Perez-Romero, and Oriol Sallent Roig. A novel approach for joint radio resource management based on fuzzy neural methodology. *Vehicular Technology, IEEE Transactions on*, 57(3):1789–1805, 2008.

[28] Antti Tölli and Petteri Hakalin. Adaptive load balancing between multiple cell layers. In *Vehicular technology conference, 2002. Proceedings. VTC 2002-Fall. 2002 IEEE 56th*, volume 3, pages 1691–1695. IEEE, 2002.

[29] Abdul Hasib and Abraham O Fapojuwo. Performance analysis of common radio resource management scheme in multi-service heterogeneous wireless networks. In *Wire-*

less Communications and Networking Conference, 2007. WCNC 2007. IEEE, pages 3296–3300. IEEE, 2007.

[30] Aleksandar Damnjanovic, Juan Montojo, Yongbin Wei, Tingfang Ji, Tao Luo, Madhavan Vajapeyam, Taesang Yoo, Osok Song, and Durga Malladi. A survey on 3gpp heterogeneous networks. *Wireless Communications, IEEE*, 18(3):10–21, 2011.

[31] Bushra Ismaiel, Mehran Abolhasan, David Smith, Wei Ni, and Daniel Franklin. A survey and comparison of device-to-device architecture using lte unlicensed band. In *Vehicular Technology Conference (VTC Spring), 2017 IEEE 85th*, pages 1–5. IEEE, 2017.

[32] Bushra Ismaiel, Mehran Abolhasan, David Smith, Wei Ni, and Daniel Franklin. Scalable mac protocol for d2d communication for future 5g networks. In *Consumer Communications & Networking Conference (CCNC), 2017 14th IEEE Annual*, pages 542–547. IEEE, 2017.

[33] Bushra Ismaiel, Mehran Abolhasan, Wei Ni, David Smith, Daniel Franklin, and Abbas Jamalipour. Analysis of effective capacity and throughput of polling based device-to-device networks. *IEEE Transactions on Vehicular Technology*, 67(9):8656 – 8666, May 2018.

[34] Bushra Ismaiel, Mehran Abolhasan, Wei Ni, David Smith, Daniel Franklin, Eryk Dutkiewicz, Marwan Krunz, and Abbas Jamalipour. PCF-Based LTE Wi-Fi aggregation for coordinating and offloading the cellular traffic to D2D network. *(to be appear in) IEEE Transactions on Vehicular Technology*, 2018.

[35] Bushra Ismaiel, Mehran Abolhasan, Wei Ni, David SMith, and Daniel Franklin. Mac protocols for Device-to-Device communication in 5G networks-A Survey. *(submitted to) IEEE Access*, 2018.

[36] MNIT Jaipur. Heterogeneous network: a pardigm shift in cellular networks. `http://www.slideshare.net/dev464898/ppt-het-net/`, 2014. [Online; accessed 19-Nov-2014].

[37] Renchao Xie, F Richard Yu, and Hong Ji. Dynamic resource allocation for heterogeneous services in cognitive radio networks with imperfect channel sensing. *Vehicular Technology, IEEE Transactions on*, 61(2):770–780, 2012.

[38] Jiwoong Lee, Jeonghoon Mo, Tran Minh Trung, Jean Walrand, and Hoi-Sheung Wilson So. Design and analysis of a cooperative multichannel mac protocol for heterogeneous networks. *Vehicular Technology, IEEE Transactions on*, 59(7):3536–3548, 2010.

[39] Shu-ping Yeh, Ali Y Panah, Nageen Himayat, and Shilpa Talwar. Qos aware scheduling and cross-radio coordination in multi-radio heterogeneous networks. In *Vehicular Technology Conference (VTC Fall), 2013 IEEE 78th*, pages 1–6. IEEE, 2013.

[40] Radhika Jain and Upkar Varshney. Supporting quality of service in multiple heterogeneous wireless networks. In *Vehicular Technology Conference, 2002. VTC Spring 2002. IEEE 55th*, volume 2, pages 952–956. IEEE, 2002.

[41] Wei Song, Hai Jiang, and Weihua Zhuang. Performance analysis of the wlan-first scheme in cellular/wlan interworking. *Wireless Communications, IEEE Transactions on*, 6(5):1932–1952, 2007.

[42] Phuong Luong, Tri Minh Nguyen, Long Bao Le, and Ngoc-Duung Eao. Admission control design for integrated wlan and ofdma-based cellular networks. In *Wireless Communications and Networking Conference (WCNC), 2014 IEEE*, pages 1386–1391. IEEE, 2014.

[43] J Pérez-Romero, O Sallent, and R Agustí. Policy-based initial rat selection algorithms in heterogeneous networks. In *Proceedings of the 7th IFIP International Conference*

on Mobile and Wireless Communication Networks (MWCN05), Marrakech, Morocco, 2005.

[44] Khuzairi Mohd Zaini, Azizul Rahman Mohd Shariff, and Zheng Shi. Traffic offloading and its application in crrm. In *Wireless and Mobile, 2014 IEEE Asia Pacific Conference on*, pages 132–137. IEEE, 2014.

[45] ZTE Technologies. Application of D2D in 5G Networks. `http://wwwen.zte.com.cn/endata/magazine/ztetechnologies/2015/no3/articles/201505/t20150506_433771.html`, 2015. [Online; accessed MAY-2015].

[46] Gábor Fodor, Erik Dahlman, Gunnar Mildh, Stefan Parkvall, Norbert Reider, György Miklós, and Zoltán Turányi. Design aspects of network assisted device-to-device communications. *Communications Magazine, IEEE*, 50(3):170–177, 2012.

[47] Rawan Alkurd, Raed M Shubair, and Ibrahim Abualhaol. Survey on device-to-device communications: Challenges and design issues. In *New Circuits and Systems Conference (NEWCAS), 2014 IEEE 12th International*, pages 361–364. IEEE, 2014.

[48] Arash Asadi, Qing Wang, and Vincenzo Mancuso. A survey on device-to-device communication in cellular networks. *Communications Surveys & Tutorials, IEEE*, 16(4):1801–1819, 2014.

[49] Jonathan Loo, Jaime Lloret Mauri, and Jesús Hamilton Ortiz. *Mobile ad hoc networks: current status and future trends*. CRC Press, 2016.

[50] Tie Qiu, Ning Chen, Keqiu Li, Daji Qiao, and Zhangjie Fu. Heterogeneous ad hoc networks: Architectures, advances and challenges. *Ad Hoc Networks*, 2016.

[51] Marco Conti and Silvia Giordano. Mobile ad hoc networking: milestones, challenges, and new research directions. *IEEE Communications Magazine*, 52(1):85–96, 2014.

[52] Al-Sakib Khan Pathan. *Security of self-organizing networks: MANET, WSN, WMN, VANET*. CRC press, 2016.

[53] 802-2014 - IEEE Standard for Local and Metropolitan Area Networks: Overview and Architecture, 2014.

[54] Dimitris Tsolkas, Eirini Liotou, Nikos Passas, and Lazaros Merakos. Lte-a access, core, and protocol architecture for d2d communication. In *Smart Device to Smart Device Communication*, pages 23–40. Springer, 2014.

[55] Phil Karn. Maca-a new channel access method for packet radio. In *ARRL/CRRL Amateur radio 9th computer networking conference*, volume 140, pages 134–140, 1990.

[56] Zygmunt J Haas and Jing Deng. Dual busy tone multiple access (dbtma)-a multiple access control scheme for ad hoc networks. *Communications, IEEE Transactions on*, 50(6):975–985, 2002.

[57] Hangguan Shan, Weihua Zhuang, and Zongxin Wang. Distributed cooperative mac for multihop wireless networks. *IEEE Communications magazine*, 47(2):126–133, 2009.

[58] Ian F Akyildiz, Won-Yeol Lee, and Kaushik R Chowdhury. Crahns: Cognitive radio ad hoc networks. *AD hoc networks*, 7(5):810–836, 2009.

[59] Bin Wang, Li Chen, Xiaohang Chen, Xin Zhang, and Dacheng Yang. Resource allocation optimization for device-to-device communication underlaying cellular networks. In *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, pages 1–6. IEEE, 2011.

[60] Vaduvur Bharghavan, Alan Demers, Scott Shenker, and Lixia Zhang. Macaw: a media access protocol for wireless lan's. In *ACM SIGCOMM Computer Communication Review*, volume 24, pages 212–225. ACM, 1994.

[61] Asis Nasipuri, Shengchun Ye, J You, and Robert E Hiromoto. A mac protocol for mobile ad hoc networks using directional antennas. In *Wireless Communications and Networking Confernce, 2000. WCNC. 2000 IEEE*, volume 3, pages 1214–1219. IEEE, 2000.

[62] Tao Guo and Rolando Carrasco. Crbar: Cooperative relay-based auto rate mac for multirate wireless networks. *IEEE Transactions on Wireless Communications*, 8(12), 2009.

[63] Claudia Cormio and Kaushik R Chowdhury. An adaptive multiple rendezvous control channel for cognitive radio wireless ad hoc networks. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2010 8th IEEE International Conference on*, pages 346–351. IEEE, 2010.

[64] Qiaoyang Ye, Mazin Al-Shalash, Constantine Caramanis, and Jeffrey G Andrews. Distributed resource allocation in device-to-device enhanced cellular networks. *Communications, IEEE Transactions on*, 63(2):441–454, 2015.

[65] Fabrizio Talucci, Mario Gerla, and Luigi Fratta. Maca-bi (maca by invitation)-a receiver oriented access protocol for wireless multihop networks. In *Personal, Indoor and Mobile Radio Communications, 1997. Waves of the Year 2000. PIMRC'97., The 8th IEEE International Symposium on*, volume 2, pages 435–439. IEEE, 1997.

[66] Shih-Lin Wu, Chih-Yu Lin, Yu-Chee Tseng, and Jang-Ping Sheu. A new multi-channel mac protocol with on-demand channel assignment for multi-hop mobile ad hoc networks. In *Parallel Architectures, Algorithms and Networks, 2000. I-SPAN 2000. Proceedings. International Symposium on*, pages 232–237. IEEE, 2000.

[67] Jin Zhang, Qian Zhang, and Weijia Jia. Vc-mac: A cooperative mac protocol in vehicular networks. *IEEE Transactions on Vehicular Technology*, 58(3):1561–1571, 2009.

[68] Sang-Jo Yoo, Hao Nan, and Tae-In Hyon. Dcr-mac: distributed cognitive radio mac protocol for wireless ad hoc networks. *Wireless Communications and Mobile Computing*, 9(5):631–653, 2009.

[69] J. Liu, S. Zhang, H. Nishiyama, N. Kato, and J. Guo. A stochastic geometry analysis of d2d overlaying multi-channel downlink cellular networks. In *2015 IEEE Conference on Computer Communications (INFOCOM)*, pages 46–54, April 2015.

[70] Chenxi Zhu and M Scott Corson. A five-phase reservation protocol (fprp) for mobile ad hoc networks. In *INFOCOM'98. Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 1, pages 322–331. IEEE, 1998.

[71] Haiming Yang and Biplab Sikdar. Performance analysis of polling based tdma mac protocols with sleep and wakeup cycles. In *Communications, 2007. ICC'07. IEEE International Conference on*, pages 241–246. IEEE, 2007.

[72] Helmut Adam, Wilfried Elmenreich, Christian Bettstetter, and Sidi Mohammed Senouci. Core-mac: a mac-protocol for cooperative relaying in wireless networks. In *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, pages 1–6. IEEE, 2009.

[73] C-F Shih, Tsung Ying Wu, and Wanjiun Liao. Dh-mac: A dynamic channel hopping mac protocol for cognitive radio networks. In *Communications (ICC), 2010 IEEE International Conference on*, pages 1–5. IEEE, 2010.

[74] Suresh Singh and Cauligi S Raghavendra. Pamaspower aware multi-access protocol with signalling for ad hoc networks. *ACM SIGCOMM Computer Communication Review*, 28(3):5–26, 1998.

[75] Gavin Holland, Nitin Vaidya, and Paramvir Bahl. A rate-adaptive mac protocol for multi-hop wireless networks. In *Proceedings of the 7th annual international conference on Mobile computing and networking*, pages 236–251. ACM, 2001.

[76] Ruonan Zhang, Rukhsana Ruby, Jianping Pan, Lin Cai, and Xuemin Shen. A hybrid reservation/contention-based mac for video streaming over wireless networks. *IEEE Journal on Selected Areas in Communications*, 28(3), 2010.

[77] Hongqiang Zhai, Jianfeng Wang, Xiang Chen, and Yuguang Fang. Medium access control in mobile ad hoc networks: challenges and solutions. *Wireless Communications and Mobile Computing*, 6(2):151–170, 2006.

[78] Daquan Feng, Lu Lu, Yi Yuan-Wu, Geoffrey Ye Li, Shaoqian Li, and Gang Feng. Device-to-device communications in cellular networks. *IEEE Communications Magazine*, 52(4):49–55, 2014.

[79] Yiyang Pei and Ying-Chang Liang. Resource allocation for device-to-device communications overlaying two-way cellular networks. *IEEE Transactions on Wireless Communications*, 12(7):3611–3621, 2013.

[80] Mei-Ju Shih, Guan-Yu Lin, and Hung-Yu Wei. A distributed multi-channel feedbackless mac protocol for d2d broadcast communications. *Wireless Communications Letters, IEEE*, 4(1):102–105, 2015.

[81] Jing-Wei Kao, Yuan-Yao Shih, Ai-Chun Pang, and Yung-Chun Lin. Radio resource allocation for d2d-assisted full-duplex cellular networks. In *2015 Seventh International Conference on Ubiquitous and Future Networks*, pages 721–726. IEEE, 2015.

[82] Xiaoqiang Ma, Jiangchuan Liu, and Hongbo Jiang. Resource allocation for heterogeneous applications with device-to-device communication underlaying cellular networks. *IEEE Journal on Selected Areas in Communications*, 34(1):15–26, 2016.

[83] Yunan Gu, Walid Saad, Mehdi Bennis, Merouane Debbah, and Zhu Han. Matching theory for future wireless networks: fundamentals and applications. *Communications Magazine, IEEE*, 53(5):52–59, 2015.

[84] Lei Lei, Zhangdui Zhong, Chuang Lin, and Xuemin Shen. Operator controlled device-to-device communications in lte-advanced networks. *IEEE Wireless Communications*, 19(3):96, 2012.

[85] Rui Yin, Caijun Zhong, Guanding Yu, Zhaoyang Zhang, Kai Kit Wong, and Xiaoming Chen. Joint spectrum and power allocation for d2d communications underlaying cellular networks. *IEEE Transactions on Vehicular Technology*, 65(4):2182–2195, 2016.

[86] Min-Hong Han, Byung-Gook Kim, and Jang-Won Lee. Subchannel and transmission mode scheduling for d2d communication in ofdma networks. In *Vehicular Technology Conference (VTC Fall), 2012 IEEE*, pages 1–5. IEEE, 2012.

[87] Monowar Hasan and Ekram Hossain. Distributed resource allocation in d2d-enabled multi-tier cellular networks: An auction approach. In *2015 IEEE International Conference on Communications (ICC)*, pages 2949–2954. IEEE, 2015.

[88] Mohammad Zulhasnine, Changcheng Huang, and Anand Srinivasan. Efficient resource allocation for device-to-device communication underlaying lte network. In *2010 IEEE 6th International Conference on Wireless and Mobile Computing, Networking and Communications*, pages 368–375. IEEE, 2010.

[89] Rui Yin, Guanding Yu, Caijun Zhong, and Zhaoyang Zhang. Distributed resource allocation for d2d communication underlaying cellular networks. In *2013 IEEE In-*

ternational Conference on Communications Workshops (ICC), pages 138–143. IEEE, 2013.

[90] Negin Golrezaei, Parisa Mansourifard, Andreas F Molisch, and Alexandros G Dimakis. Base-station assisted device-to-device communications for high-throughput wireless video networks. *Wireless Communications, IEEE Transactions on*, 13(7):3665–3676, 2014.

[91] Guopeng Zhang, Kun Yang, and Hsiao-Hwa Chen. Socially aware cluster formation and radio resource allocation in d2d networks. *IEEE Wireless Communications*, 23(4):68–73, 2016.

[92] Arash Asadi and Vincenzo Mancuso. On the compound impact of opportunistic scheduling and d2d communications in cellular networks. In *Proceedings of the 16th ACM international conference on Modeling, analysis & simulation of wireless and mobile systems*, pages 279–288. ACM, 2013.

[93] Long Bao Le. Fair resource allocation for device-to-device communications in wireless cellular networks. In *Global Communications Conference (GLOBECOM), 2012 IEEE*, pages 5451–5456. IEEE, 2012.

[94] Karaputugala G Madushan Thilina, Ekram Hossain, and Dong In Kim. Dccc-mac: A dynamic common-control-channel-based mac protocol for cellular cognitive radio networks. *IEEE Transactions on Vehicular Technology*, 65(5):3597–3613, 2016.

[95] Junyu Liu, Min Sheng, Yan Zhang, Xijun Wang, Hongguang Sun, and Yan Shi. A distributed opportunistic scheduling protocol for device-to-device communications. In *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 1715–1719. IEEE, 2013.

[96] Xinzhou Wu, Saurabha Tavildar, Sanjay Shakkottai, Tom Richardson, Junyi Li, Rajiv Laroia, and Aleksandar Jovicic. Flashlinq: A synchronous distributed scheduler for peer-to-peer ad hoc networks. *IEEE/ACM Transactions on Networking (TON)*, 21(4):1215–1228, 2013.

[97] Chai K Toh. *Ad hoc mobile wireless networks: protocols and systems*. Pearson Education, 2001.

[98] Amjad Ali, Wang Huiqiang, Lv Hongwu, and Xiaoming Chen. A survey of mac protocols design strategies and techniques in wireless ad hoc networks. *Journal of Communications*, 9(1):30–38, 2014.

[99] K Joice Olempia, C Pandeeswaran, and Pappa Natarajan. A survey on energy efficient contention based and hybrid mac protocols for wireless sensor networks. *Indian Journal of Science and Technology*, 9(12), 2016.

[100] Rajarshi Mahapatra, Yogesh Nijsure, Georges Kaddoum, Naveed Ul Hassan, and Chau Yuen. Energy efficiency tradeoff mechanism towards wireless green communication: A survey. *IEEE Communications Surveys & Tutorials*, 18(1):686–705, 2015.

[101] Chien-Erh Weng and Hsing-Chung Chen. The performance evaluation of ieee 802.11 dcf using markov chain model for wireless lans. *Computer Standards & Interfaces*, 44:144–149, 2016.

[102] Qamar Jabeen, Fazlullah Khan, Shahzad Khan, and Mian Ahmad Jan. Performance improvement in multihop wireless mobile adhoc networks. *the Journal Applied, Environmental, and Biological Sciences (JAEBS)*, 6:82–92, 2016.

[103] Rajoua Anane, Ridha Bouallegue, and Kosai Raoof. Medium access control (mac) protocols for wireless sensor network: An energy aware survey. In *Proceedings of the*

Mediterranean Conference on Information & Communication Technologies 2015, pages 561–569. Springer, 2016.

[104] Chane L Fullmer and JJ Garcia-Luna-Aceves. *Floor acquisition multiple access (FAMA) for packet-radio networks*, volume 25. ACM, 1995.

[105] CK Toh, Vasos Vassiliou, Guillermo Guichal, and C-H Shih. March: a medium access control protocol for multihop wireless ad hoc networks. In *MILCOM 2000. 21st Century Military Communications Conference Proceedings*, volume 1, pages 512–516. IEEE, 2000.

[106] Cheng-shong Wu and V Li. Receiver-initiated busy-tone multiple access in packet radio networks. In *ACM SIGCOMM Computer Communication Review*, volume 17, pages 336–342. ACM, 1987.

[107] Hongyi Wu, Anant Utgikar, and Nian-Feng Tzeng. Syn-mac: a distributed medium access control protocol for synchronized wireless networks. *Mobile Networks and Applications*, 10(5):627–637, 2005.

[108] Injong Rhee, Ajit Warrier, Mahesh Aia, Jeongki Min, and Mihail L Sichitiu. Z-mac: a hybrid mac for wireless sensor networks. *IEEE/ACM Transactions on Networking (TON)*, 16(3):511–524, 2008.

[109] Fouad Tobagi, Leonard Kleinrock, et al. Packet switching in radio channels: Part ii–the hidden terminal problem in carrier sense multiple-access and the busy-tone solution. *Communications, IEEE Transactions on*, 23(12):1417–1433, 1975.

[110] Wenchi Cheng, Xi Zhang, and Hailin Zhang. Full-duplex spectrum-sensing and mac-protocol for multichannel nontime-slotted cognitive radio networks. *IEEE Journal on Selected Areas in Communications*, 33(5):820–831, 2015.

[111] Khaled H Almotairi and Xuemin Sherman Shen. A distributed multi-channel mac protocol for ad hoc wireless networks. *Mobile Computing, IEEE Transactions on*, 14(1):1–13, 2015.

[112] Charles E Perkins. *Ad hoc networking*. Addison-Wesley Professional, 2008.

[113] Vikram Kanodia, Chengzhi Li, Ashutosh Sabharwal, Bahareh Sadeghi, and Edward Knightly. Distributed multi-hop scheduling and medium access with delay and throughput constraints. In *Proceedings of the 7th annual international conference on Mobile computing and networking*, pages 200–209. ACM, 2001.

[114] Chunhung Richard Lin and Mario Gerla. Asynchronous multimedia multihop wireless networks. In *INFOCOM'97. Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Driving the Information Revolution., Proceedings IEEE*, volume 1, pages 118–125. IEEE, 1997.

[115] Jesús Alonso-Zárate, Elli Kartsakli, Luis Alonso, and Christos Verikoukis. Performance analysis of a cluster-based mac protocol for wireless ad hoc networks. *EURASIP Journal on Wireless Communications and Networking*, 2010:11, 2010.

[116] Andreas Köpsel, Jean-Pierre Ebert, and Adam Wolisz. A performance comparison of point and distributed coordination function of an ieee 802.11 wlan in the presence of real-time requirements. In *7th International Workshop on Mobile Multimedia Communication (MoMuC2000)*, 2000.

[117] Vikrant Gokhale, SK Ghosh, and Arobinda Gupta. Classification of attacks on wireless mobile ad hoc networks and vehicular ad hoc networks. *Security of Self-Organizing Networks*, page 195, 2016.

[118] Bo Li, Qiao Qu, Zhongjiang Yan, and Mao Yang. Survey on ofdma based mac protocols for the next generation wlan. In *Wireless Communications and Networking Conference Workshops (WCNCW), 2015 IEEE*, pages 131–135. IEEE, 2015.

[119] Ravindra S Ranasinghe, Lachlan LH Andrew, and David Everitt. Impact of polling strategy on capacity of 802.11 based wireless multimedia lans. In *Networks, 1999.(ICON'99) Proceedings. IEEE International Conference on*, pages 96–103. IEEE, 1999.

[120] Eustathia Ziouva and Theodore Antonakopoulos. Efficient voice communications over ieee 802.11 wlans using improved pcf procedures. *Proc. INC*, 2, 2002.

[121] Sunghyun Choi and Kang G Shin. A unified wireless lan architecture for real-time and non-real-time communication services. *Networking, IEEE/ACM Transactions on*, 8(1):44–59, 2000.

[122] Oran Sharon and Eitan Altman. An efficient polling mac for wireless lans. *IEEE/ACM Transactions on Networking (TON)*, 9(4):439–451, 2001.

[123] Byung-Seo Kim, Sung Won Kim, Yuguang Fang, and Tan F Wong. Two-step multipolling mac protocol for wireless lans. *Selected Areas in Communications, IEEE Journal on*, 23(6):1276–1286, 2005.

[124] Hao Zhu and Guohong Cao. On improving the performance of ieee 802.11 with relay-enabled pcf. *Mobile Networks and Applications*, 9(4):423–434, 2004.

[125] R. Palacios, G. M. Mekonnen, J. Alonso-Zarate, D. Kliazovich, and F. Granelli. Analysis of an energy-efficient MAC protocol based on polling for IEEE 802.11 WLANs. In *2015 IEEE International Conference on Communications (ICC)*, pages 5941–5947, June 2015.

[126] Ray YW Lam, Victor Leung, and Henry CB Chan. Polling-based protocols for packet voice transport over ieee 802.11 wireless local area networks. *Wireless Communications, IEEE*, 13(1):22–29, 2006.

[127] Fadhil Firyaguna and Marcelo M Carvalho. Performance of polling disciplines for the receiver-initiated binary exponential backoff mac protocol. *Ad Hoc Networks*, 31:1–19, 2015.

[128] David Eckhardt, Peter Steenkiste, et al. Effort-limited fair (elf) scheduling for wireless networks. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 3, pages 1097–1106. IEEE, 2000.

[129] Wen-Tsuen Chen, Sheng-Hsien Chen, and Jen-Chu Liu. An efficient qos guaranteed mac protocol in wireless atm networks. In *Information Networking, 2001. Proceedings. 15th International Conference on*, pages 785–792. IEEE, 2001.

[130] Fang-Yie Leu, Ching-Chien Kuan, and Dr-Jiunn Deng. A qos provision multipolling mechanism for ieee 802.11 e standard. In *Wireless Networks, Communications and Mobile Computing, 2005 International Conference on*, volume 1, pages 392–397. IEEE, 2005.

[131] William Michael Rudnick, Ngan-Cheung Pun, David Clark, and Charles Joseph Datz. Mobile ad hoc network with dynamic tdma slot assignments and related methods, January 27 2015. US Patent 8,942,197.

[132] Ajinkya Rajandekar and Biplab Sikdar. A survey of mac layer issues and protocols for machine-to-machine communications. *IEEE Internet of Things Journal*, 2(2):175–186, 2015.

[133] Petar Djukic and Shahrokh Valaee. Delay aware link scheduling for multi-hop tdma wireless networks. *IEEE/ACM Transactions on Networking (TON)*, 17(3):870–883, 2009.

[134] Marion Souil, Tifenn Rault, and Abdelmadjid Bouabdallah. A new adaptive mac protocol with qos support for heterogeneous wireless sensor networks. In *Computers and Communications (ISCC), 2012 IEEE Symposium on*, pages 000405–000410. IEEE, 2012.

[135] Lichun Bao and JJ Garcia-Luna-Aceves. A new approach to channel access scheduling for ad hoc networks. In *Proceedings of the 7th annual international conference on Mobile computing and networking*, pages 210–221. ACM, 2001.

[136] Venkatesh Rajendran, Katia Obraczka, and Jose Joaquin Garcia-Luna-Aceves. Energy-efficient, collision-free medium access control for wireless sensor networks. *Wireless Networks*, 12(1):63–78, 2006.

[137] Injong Rhee, Ajit Warrier, Jeongki Min, and Lisong Xu. Drand: distributed randomized tdma scheduling for wireless ad-hoc networks. In *Proceedings of the 7th ACM international symposium on Mobile ad hoc networking and computing*, pages 190–201. ACM, 2006.

[138] 802.11e-2005 - IEEE Standard for Information technology–Local and metropolitan area networks–Specific requirements–Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications - Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements, 2005.

[139] Lamia Romdhani, Qiang Ni, and Thierry Turletti. Adaptive edcf: enhanced service differentiation for ieee 802.11 wireless ad-hoc networks. In *Wireless Communications*

and Networking, 2003. WCNC 2003. 2003 IEEE, volume 2, pages 1373–1378. IEEE, 2003.

[140] Mohammad Malli, Qiang Ni, Thierry Turletti, and Chadi Barakat. Adaptive fair channel allocation for qos enhancement in ieee 802.11 wireless lans. In *Communications, 2004 IEEE International Conference on*, volume 6, pages 3470–3475. IEEE, 2004.

[141] Manas Khatua and Sudip Misra. D2d: Delay-aware distributed dynamic adaptation of contention window in wireless networks. *IEEE Transactions on Mobile Computing*, 15(2):322–335, 2016.

[142] Colin Ruby. Provision of Multimedia Services in 802.11-based Networks. `www.csee.umbc.edu/courses/graduate/681/Fall07/Presentations/Colin%2520Roby%2520-%2520present.ppt+&cd=3&hl=en&ct=clnk&gl=au`, 2007. [Online;2007].

[143] Monica Paolini and Senza Fili. Lte unlicensed and wi-fi: Moving beyond coexistence.

[144] Klaus Doppler, Mika Rinne, Carl Wijting, Cássio B Ribeiro, and Klaus Hugl. Device-to-device communication as an underlay to lte-advanced networks. *IEEE Communications Magazine*, 47(12):42–49, 2009.

[145] Muhammad Imadur Rahman, Ali Behravant, Havish Koorapaty, Joachim Sachs, and Kumar Balachandran. License-exempt lte systems for secondary spectrum usage: scenarios and first assessment. In *New Frontiers in Dynamic Spectrum Access Networks (DySPAN), 2011 IEEE Symposium on*, pages 349–358. IEEE, 2011.

[146] Haijun Zhang, Xiaoli Chu, Weisi Guo, and Siyi Wang. Coexistence of wi-fi and heterogeneous small cell networks sharing unlicensed spectrum. *IEEE Communications Magazine*, 53(3):158–164, 2015.

[147] Netmanias. Analysis of LTE WiFi Aggregation Solutions. http://www.netmanias.com/en/post/reports/8532/laa-lte-lte-u-lwa-mptcp/analysis-of-lte-wifi-aggregation-solutions, 2016. [Online;2016].

[148] Ning Zhang, Shan Zhang, Shaohua Wu, Ju Ren, Jon W Mark, and Xuemin Sherman Shen. Beyond coexistence: Traffic steering in lte networks with unlicensed bands.

[149] Bolin Chen, Jiming Chen, Yuan Gao, and Jie Zhang. Coexistence of lte-laa and wi-fi on 5 ghz with corresponding deployment scenarios: A survey.

[150] 3GPP. 3GPP. http://www.3gpp.org/, 2016. [Online;2016].

[151] 3GPP. Collaboration on LTE - WLAN Integration. http://www.3gpp.org/news-events/3gpp-news/1771-wlan_lte, 2016. [Online;2016].

[152] Qualcomm. MulteFire: LTE-like performance with Wi-Fi-like deployment simplicity. https://www.qualcomm.com/invention/technologies/lte/multefire, 2016. [Online;2016].

[153] Ran Zhang, Miao Wang, Lin X Cai, Zhongming Zheng, Xuemin Shen, and Liang-Liang Xie. Lte-unlicensed: the future of spectrum aggregation for cellular networks. *IEEE Wireless Communications*, 22(3):150–159, 2015.

[154] Qualcomm. Extending lte to unlicensed spectrum globally laa.

[155] Roslyn Layton. What The LTE-U Vs. WiFi Debate Is Really About. http://www.forbes.com/sites/roslynlayton/2015/09/02/what-the-lte-u-vs-wifi-debate-is-really-about/#538b71e74a03, 2015. [Online;2015].

[156] Martha DeGrasse. MWC15: Alcatel-Lucent combines Wi-Fi and cellular networks. http://www.rcrwireless.com/20150302/network-infrastructure/wi-fi/alcatel-lucent-blends-wi-fi-and-cellular-tag4, 2015. [Online;2015].

[157] Insoo Hwang, Bongyong Song, and Samy S Soliman. A holistic view on hyper-dense heterogeneous and small cell networks. *Communications Magazine, IEEE*, 51(6):20–27, 2013.

[158] Mehran Abolhasan, Justin Lipman, Wei Ni, and Brett Hagelstein. Software-defined wireless networking: centralized, distributed, or hybrid? *Network, IEEE*, 29(4):32–38, 2015.

[159] Brian P Crow, Indra Widjaja, Jeong Geun Kim, and Prescott Sakai. Investigation of the ieee 802.11 medium access control (mac) sublayer functions. In *INFOCOM'97. Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Driving the Information Revolution., Proceedings IEEE*, volume 1, pages 126–133. IEEE, 1997.

[160] Matthijs A Visser and Magda El Zarki. Voice and data transmission over an 802.11 wireless network. In *Personal, Indoor and Mobile Radio Communications, 1995. PIMRC'95. Wireless: Merging onto the Information Superhighway., Sixth IEEE International Symposium on*, volume 2, pages 648–652. IEEE, 1995.

[161] Jiandong Wu and Guohui Huang. Simulation study based on qos schemes for ieee 802.11. In *Advanced Computer Theory and Engineering (ICACTE), 2010 3rd International Conference on*, volume 6, pages V6–534. IEEE, 2010.

[162] S Sharma, K Kim, C Oh, and A Ahmad. Performance evaluation of ieee 802.11 mac protocol for multimedia services. *Proceedings of APCC/ICCS*, pages 181–185, 1998.

[163] Takahiro Suzuki and Shuji Tasaka. Performance evaluation of integrated video and data transmission with the ieee 802.11 standard mac protocol. In *Global Telecommunications Conference, 1999. GLOBECOM'99*, volume 1, pages 580–586. IEEE, 1999.

[164] Liu Qianlin and Zhao Dongfeng. Analysis of two-level-polling system with mixed access policies. In *Intelligent Computation Technology and Automation, 2009. ICICTA'09. Second International Conference on*, volume 4, pages 357–360. IEEE, 2009.

[165] Ding Hongwei, Zhao Dongfeng, and Zhao Yifan. Analysis of polling system with multiple vacations and using exhaustive service. In *Innovative Computing & Communication, 2010 Intl Conf on and Information Technology & Ocean Engineering, 2010 Asia-Pacific Conf on (CICC-ITOE)*, pages 294–296. IEEE, 2010.

[166] Biplab Sikdar. An analytic model for the delay in ieee 802.11 pcf mac-based wireless networks. *IEEE Transactions on Wireless Communications*, 6(4), 2007.

[167] Md Atiur Rahman Siddique and Joarder Kamruzzaman. Performance analysis of pcf based wlans with imperfect channel and failure retries. In *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, pages 1–6. IEEE, 2010.

[168] Yu Gu, Qimei Cui, Yu Chen, Wei Ni, Xiaofeng Tao, and Ping Zhang. Effective capacity analysis in ultra-dense wireless networks with random interference. *arXiv preprint arXiv:1803.10891*, 2018.

[169] Lingyun Lu, Haifeng Du, and Ren Ping Liu. Choker: A novel aqm algorithm with proportional bandwidth allocation and tcp protection. *IEEE Transactions on Industrial Informatics*, 10(1):637–644, 2014.

[170] Emmanouil Kafetzakis, Kimon Kontovasilis, and Ioannis Stavrakakis. A novel effective capacity-based framework for providing statistical qos guarantees in ieee 802.11 wlans. *Computer Communications*, 35(2):249–262, 2012.

[171] Wei Ni, Iain B Collings, and Ren Ping Liu. Decentralized user-centric scheduling with low rate feedback for mobile small cells. *IEEE Transactions on Wireless Communications*, 12(12):6106–6120, 2013.

[172] Angelos Antonopoulos, Christos Verikoukis, Charalabos Skianis, and Ozgur B Akan. Energy efficient network coding-based mac for cooperative arq wireless networks. *Ad Hoc Networks*, 11(1):190–200, 2013.

[173] Jiping Li, Yaoming Ding, Qiang Ye, Ning Zhang, and Weihua Zhuang. On effective capacity and effective energy efficiency in relay-assisted wireless networks. *IEEE Transactions on Vehicular Technology*, 2017.

[174] Dapeng Wu and Rohit Negi. Effective capacity: a wireless link model for support of quality of service. *IEEE Transactions on wireless communications*, 2(4):630–643, 2003.

[175] Xi Zhang, Jia Tang, Hsiao-Hwa Chen, Song Ci, and Mohsen Guizani. Cross-layer-based modeling for quality of service guarantees in mobile wireless networks. *IEEE Communications Magazine*, 44(1):100–106, 2006.

[176] Qimei Cui, Yu Gu, Wei Ni, and Ren Ping Liu. Effective capacity of licensed-assisted access in unlicensed spectrum: From theory to applications. *IEEE Journal on Selected Areas in Communications*, 2017.

[177] Murali Krishna Kadiyala, Dipti Shikha, Ravi Pendse, and Neeraj Jaggi. Semi-markov process based model for performance analysis of wireless lans. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2011 IEEE International Conference on*, pages 613–618. IEEE, 2011.

[178] Leila Musavian and Qiang Ni. Effective capacity maximization with statistical delay and effective energy efficiency requirements. *IEEE Transactions on Wireless Communications*, 14(7):3824–3835, 2015.

[179] Kimon Kontovasilis and Nikolas Mitrou. Effective bandwidths for a class of non markovian fluid sources. *ACM SIGCOMM Computer Communication Review*, 27(4):263–274, 1997.

[180] Wei Ni and Iain B Collings. A new adaptive small-cell architecture. *IEEE Journal on Selected Areas in Communications*, 31(5):829–839, 2013.

[181] Wei Ni, Iain B Collings, and Ren Ping Liu. Relay handover and link adaptation design for fixed relays in imt-advanced using a new markov chain model. *IEEE Transactions on Vehicular Technology*, 61(4):1839–1853, 2012.

[182] IEEE Std 802.11g. Analysis of LTE WiFi Aggregation Solutions. `http://infocom.uniroma1.it/alef/802.11/standard/802.11g-2003.pdf`, 2003. [Online;2003].

[183] Yu Gu, Qimei Cui, Wei Ni, Ping Zhang, and Weihua Zhuang. Optimal scheduling across heterogeneous air interfaces of lte/wifi aggregation. *arXiv preprint arXiv:1802.01799*, 2018.

[184] Olga Galinina, Alexander Pyattaev, Sergey Andreev, Mischa Dohler, and Yevgeni Koucheryavy. 5g multi-rat lte-wifi ultra-dense small cells: Performance dynamics, architecture, and trends. *IEEE Journal on Selected Areas in Communications*, 33(6):1224–1240, 2015.

[185] Feilu Liu, Erdem Bala, Elza Erkip, Mihaela C Beluri, and Rui Yang. Small-cell traffic balancing over licensed and unlicensed bands. *IEEE transactions on vehicular technology*, 64(12):5850–5865, 2015.

[186] Ahmed R Elsherif, Wei-Peng Chen, Akira Ito, and Zhi Ding. Resource allocation and inter-cell interference management for dual-access small cells. *IEEE Journal on Selected Areas in Communications*, 33(6):1082–1096, 2015.

[187] Mehdi Bennis, Meryem Simsek, Andreas Czylwik, Walid Saad, Stefan Valentin, and Merouane Debbah. When cellular meets wifi in wireless small cell networks. *IEEE communications magazine*, 51(6):44–50, 2013.

[188] Qimei Cui, Yulong Shi, Xiaofeng Tao, Ping Zhang, Ren Ping Liu, Ningyu Chen, Jyri Hamalainen, and Alexis Dowhuszko. A unified protocol stack solution for lte and wlan in future mobile converged networks. *IEEE wireless communications*, 21(6):24–33, 2014.

[189] Amin Abdel Khalek, Constantine Caramanis, and Robert W Heath. Delay-constrained video transmission: Quality-driven resource allocation and scheduling. *IEEE Journal of Selected Topics in Signal Processing*, 9(1):60–75, 2015.

[190] Fan Jin, Rong Zhang, and Lajos Hanzo. Resource allocation under delay-guarantee constraints for heterogeneous visible-light and rf femtocell. *IEEE Transactions on Wireless Communications*, 14(2):1020–1034, 2015.

[191] Godfrey Harold Hardy, John Edensor Littlewood, and George Pólya. *Inequalities. By GH Hardy, JE Littlewood, G. Pólya.*. University Press, 1952.

[192] Enlong Che, Hoang Duong Tuan, and Ha H Nguyen. Joint optimization of cooperative beamforming and relay assignment in multi-user wireless relay networks. *IEEE Transactions on Wireless Communications*, 13(10):5481–5495, 2014.

[193] Behzad Khamidehi, Ali Rahmati, and Maryam Sabbaghian. Joint sub-channel assignment and power allocation in heterogeneous networks: An efficient optimization method. *IEEE Communications Letters*, 20(12):2490–2493, 2016.

[194] Neyre Tekbiyik, Tolga Girici, Elif Uysal-Biyikoglu, and Kemal Leblebicioglu. Proportional fair resource allocation on an energy harvesting downlink. *IEEE Transactions on Wireless Communications*, 12(4):1699–1711, 2013.

[195] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.