Resource Allocation in Cellular Machine-to-Machine Networks

by

Nedaa Alhussien
B.Sc., Birzeit University, 2010
M.Sc., United Arab Emirates University, 2014

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Electrical and Computer Engineering

Resource Allocation in Cellular Machine-to-Machine Networks

by

Nedaa Alhussien
B.Sc., Birzeit University, 2010
M.Sc., United Arab Emirates University, 2014

Supervisory Committee

Dr. T. Aaron Gulliver, Supervisor
(Department of Electrical and Computer Engineering)

Dr. Wu-Sheng Lu, Departmental Member
(Department of Electrical and Computer Engineering)

Dr. Yvonne Coady, Outside Member
(Department of Computer Science)

ABSTRACT

With the emergence of the Internet-of-Things (IoT), communication networks have evolved toward autonomous networks of intelligent devices capable of communicating without direct human intervention. This is known as Machine-to-Machine (M2M) communications. Cellular networks are considered one of the main technologies to support the deployment of M2M communications as they provide extended wireless connectivity and reliable communication links. However, the characteristics and Quality-of-Service (QoS) requirements of M2M communications are distinct from those of conventional cellular communications, also known as Human-to-Human (H2H) communications, that cellular networks were originally designed for. Thus, enabling M2M communications poses many challenges in terms of interference, congestion, spectrum scarcity and energy efficiency. The primary focus is on the problem of resource allocation that has been the interest of extensive research effort due to the fact that both M2M and H2H communications coexist in the cellular network. This requires that radio resources be allocated such that the QoS requirements of both groups are satisfied. In this work, we propose three models to address this problem.

In the first model, a two-phase resource allocation algorithm for H2H/M2M coexistence in cellular networks is proposed. The goal is to meet the QoS requirements of H2H traffic and delay-sensitive M2M traffic while ensuring fairness for the delay-tolerant M2M traffic. Simulation results are presented which show that the proposed algorithm is able to balance the demands of M2M and H2H traffic, meet their diverse QoS requirements, and ensure fairness for delay-tolerant M2M traffic.

With the growing number of Machine-Type Communication Devices (MTCDs) the problem of spectrum scarcity arises. Hence, Cognitive Radio (CR) is the focus of the second model where clustered Cognitive M2M (CM2M) communications underlaying cellular networks is proposed. In this model, MTCDs are grouped in clusters based on their spatial location and communicate with the Base Station (BS) via Machine-Type Communication Gateways (MTCGs). An underlay CR scheme is implemented where the MTCDs within a cluster share the spectrum of the neighbouring Cellular User Equipment (CUE). A joint resource-power allocation problem is formulated to maximize the sum-rate of the CUE and clustered MTCDs while adhering to MTCD minimum data rate requirements, MTCD transmit power limits, and CUE interference constraints. Simulation results are presented which show that the proposed scheme significantly improves the sum-rate of the network compared to

other schemes while satisfying the constraints.

Due to the limited battery capacity of MTCDs and diverse QoS requirements of both MTCDs and CUE, Energy Efficiency (EE) is critical to prolonging network lifetime to ensure uninterrupted and reliable data transmission. The third model investigates the power allocation problem for energy-efficient CM2M communications underlaying cellular networks. Underlay CR is employed to manage the coexistence of MTCDs and CUE and exploit spatial spectrum opportunities. Two power allocation problems are proposed where the first targets MTCD power consumption minimization while the second considers MTCD EE maximization subject to MTCD transmit power constraints, MTCD minimum data rate requirements, and CUE interference limits. Simulation results are presented which indicate that the proposed algorithms provide MTCD power allocation with lower power consumption and higher EE than the (Equal Power Allocation) EPA scheme while satisfying the constraints.

Contents

List of Figures

List of Acronyms

3GPP	3rd Generation Partnership Project
ADC	Analog-to-Digital Conversion
ARBA	Resource Block Allocation
AWGN	Additive White Gaussian Noise
BB	Branch and Bound
BLP	Binary Linear Programming
BPP	Binomial Point Process
BS	Base Station
CM2M	Cognitive Machine-to-Machine
CR	Cognitive Radio
CR-VANET	Cognitive Radio Vehicular Ad Hoc Networks
CUE	Cellular User Equipment
CVX	Convex Optimization
D2D	Device-to-Device
DAC	Digital-to-Analog Conversion
D.C.	Difference of Convex
DSP	Digital Signal Processing
EE	Energy Efficiency
EESM	Exponential Effective SNR Mapping
EH	Energy Harvesting
EPA	Equal Power Allocation
ETSI	European Telecommunications Standards Institute
GP	Geometric Programming
H2H	Human-to-Human
H2H UE	H2H User Equipment
ICT	Information and Communication Technology
IEEE	Institute of Electrical and Electronics Engineers
IoT	Internet-of-Things
IPA	Iterative Power Allocation
KPIs	Key Performance Indicators
LTE	Long Term Evolution
M2M	Machine-to-Machine
MAB	Multi-Armed Bandit

MIESM	Mutual Information Effective SINR Mapping
MINLP	Mixed Integer Non-Linear Programming
MOS	Mean Opinion Score
MTCD	Machine-Type Communication Device
MTCG	Machine-Type Communication Gateway
MWC	Mobile and Wireless Communications
NOMA	Non-Orthogonal Multiple Access
NORA-DT	Non-Orthogonal Random Access and Data Transmission
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PDBV	Probability of Delay-Bound Violation
PUs	Primary Users
QoE	Quality-of-Experience
QoS	Quality-of-Service
RACH	Random Access Channel
RBs	Resource Blocks
RF	Radio Frequency
SC-FDMA	Single Carrier Frequency Division Multiple Access
SINR	Signal-to-Interference Plus Noise Ratio
SS	Spectrum Sharing
SUs	Secondary Users
TCA	Threshold Controlled Access
TDMA	Time Division Multiple Access
TTI	Transmission Time Interval
TVWS	TV White Space
UHF	Ultra High Frequency
VHF	Very High Frequency
VoIP	Voice-over-IP

List of Tables

ACKNOWLEDGEMENTS

This wonderful and challenging journey would not have been possible without the support of many individuals. First and foremost, I owe my deepest gratitude to my supervisor, Prof. T. Aaron Gulliver, for his continuous support, deep knowledge, and invaluable guidance. Thank you Professor for your encouragement and patience that gave me the strength and ability to progress in my research. You have been a true inspiration to me.

I would like to express my appreciation to my supervisory committee members, Prof. Wu-Sheng Lu and Prof. Yvonne Coady, for their guidance and professional advice on my research. I would like also to thank my external examiner, Prof. Kemal Tepe, for his time and valuable feedback.

I wish to thank my beloved family. Your endless love and support not only during my study but also throughout my life is what kept me going during the hard times.

Words cannot express how grateful I am for my sweet angels, Lamar and Eyas. You are the joy of my life and the core of my soul. Last, but not least, I wish to thank my beloved husband, Younes. You have always been a constant source of love and support for me during the good and hard times. Words fail me to express my appreciation for your patience and confidence in me. Without your encouragement this journey would not have started in the first place. It is to you that I owe all the success in life.

DEDICATION

To the soul of my father, Yousef, who instilled in me the love for knowledge and raised me to believe in myself and never give up.

To my dear mother, Maryam, for her boundless love, encouragement, and silent prayers.

To my beloved husband, Younes, for his unconditional love and support.

To my precious kids, Lamar and Eyas, who filled my life with hope and joy.

Chapter 1

Introduction

The telecommunications industry has gone through tremendous growth in recent decades. A multitude of technologies exists today with the objective of providing ubiquitous connectivity for everything and everyone. Therefore, a significant rise in the number of connected devices that are not mobile phones and do not require human control is expected over the coming years. In this regard, Machine-to-Machine (M2M) refers to technologies that enable networked devices to exchange information among each other as well as with business and industrial applications.

1.1 Internet-of-Things

Recent advancements in Information and Communications Technology (ICT) and the tremendous growth in the number of smart devices have introduced the Internet-of-Things (IoT) as a new communication paradigm. IoT is envisioned as a network of connected devices with identities and virtual personalities operating in smart spaces using intelligent interfaces to connect and communicate within social, environmental, and user contexts [1]. In the IoT, devices including sensors, smart home appliances, control systems, healthcare devices, and vehicles autonomously connect with each other and exchange data using a variety of communication technologies such as WiFi, Bluetooth, ZigBee, and cellular.

The number of connected devices is expected to reach 29.3 billion by 2023 [86]. Thus, IoT networks are required to be flexible, scalable, secure, and energy-efficient in order to accommodate this massive number of devices. To facilitate this, M2M communications have emerged as the enabling technology for the practical realiza-

tion of IoT. M2M communications play a prominent role in the areas of cloud computing, smart grid (demand/response management, grid control, smart metering), vehicular communications (vehicle-to-vehicle and vehicle-to-infrastructure communications), and industrial and e-health monitoring [3].

1.2 Machine-to-Machine Communications

With the emergence of the IoT, communication networks have evolved toward autonomous networks of intelligent devices capable of communicating without direct human intervention. These are known as Machine-Type Communication Devices (MTCDs) and include sensors, smart home appliances, control systems, and healthcare devices. Communication between MTCDs is known as M2M communication and is different from conventional cellular communications, also known as Human-to-Human (H2H) communications. An overview of M2M communications is given below.

1.2.1 Features of M2M Communications

M2M communications have unique characteristics and service requirements [4]. Some of the key features and service requirements are as follows.

- Massive Number of Devices – It is expected that the number of connected MTCDs in an M2M network will exceed that of all devices that directly interact with humans, e.g. mobile phones and tablets. This exponentially growing number of MTCDs will generate a massive amount of traffic resulting in scalability issues for existing networks.

- Small Data Transmissions – Most MTCDs have small data transmissions at regular time intervals. Moreover, M2M transmissions generate higher demand on the uplink channel as most MTCDs are involved in sensing and measurement (data collection) tasks. This can lead to congestion in the physical channel, resulting in low access rates [5]. Hence, existing networks must be able to support these transmissions with minimal impact in terms of signalling overhead and resource utilization.

- Energy Constrained – The majority of MTCDs are battery-operated and may be deployed in remote areas where frequent access for battery replacement and

maintenance is infeasible [90]. Therefore, energy efficiency for M2M communications is critical to prolong network lifetime [92]. Energy-aware approaches can reduce energy consumption such as putting MTCDs into sleep mode when they are not in operation, and using lightweight energy-efficient communication protocols [8].

- Interoperability and Devices Heterogeneity – In M2M networks, a tremendous number of MTCDs run different types of applications with diverse service requirements. This creates interoperability challenges where internetworking and seamless communication between these devices with different networking protocols is required.

- Quality-of-Service (QoS) – In M2M networks, QoS requirements vary based on the MTCD application. M2M traffic can be classified into two categories: event-based and time-based [9]. Event-based traffic includes emergency alarms and critical messages from applications such as e-health monitoring and vehicle tracking where messages are sent only when events occur. MTCDs in this category are called delay-sensitive and have QoS requirements in terms of strict delay limits. On the other hand, time-based traffic comprises infrequent small data transmissions at regular time intervals that span from a few milliseconds to minutes or hours [10]. MTCDs in this category are known as delay-tolerant and require reliable communications to satisfy their delay bounds.

- Security – Security challenges exist in M2M networks due to the large number of MTCDs that can be targets for attacks such as denial-of-service attacks, physical attacks, and identity privacy attacks. Hence, security is critical in M2M communications and the need for suitable security mechanisms at both the MTCD and network levels is essential. For example, MTCDs must have the intelligence to recognize and counteract potential security threats.

1.2.2 Architecture of M2M Communications

According to the European Telecommunications Standards Institute (ETSI) standards for Information and Communications Technology (ICT), a M2M network consists of five elements [11].

1. Device – Sends and/or receives data or replies to requests. M2M devices usually exchange information with a remote server or another device. Examples of

M2M devices are body sensors in e-healthcare system or sensors in surveillance systems.

2. Gateway – Manages communications between devices and other networks.
3. M2M Area Network – Provides connections between devices and gateways.
4. Communication Network – Provides connections between gateways and applications.
5. Application – Forwards data through application services to processing engines that analyze data, take action and report data.

Figure 1.1 shows the M2M communications architecture that is composed of three interconnected domains: M2M, network, and application [12]. The M2M domain consists of the M2M area network and gateway, while the network domain provides the infrastructure for communications between the M2M domain and the application domain. The application domain contains a M2M server as a data integration point for storing sensor data from the M2M domain and provides real-time data to M2M applications for data processing and remote monitoring management.

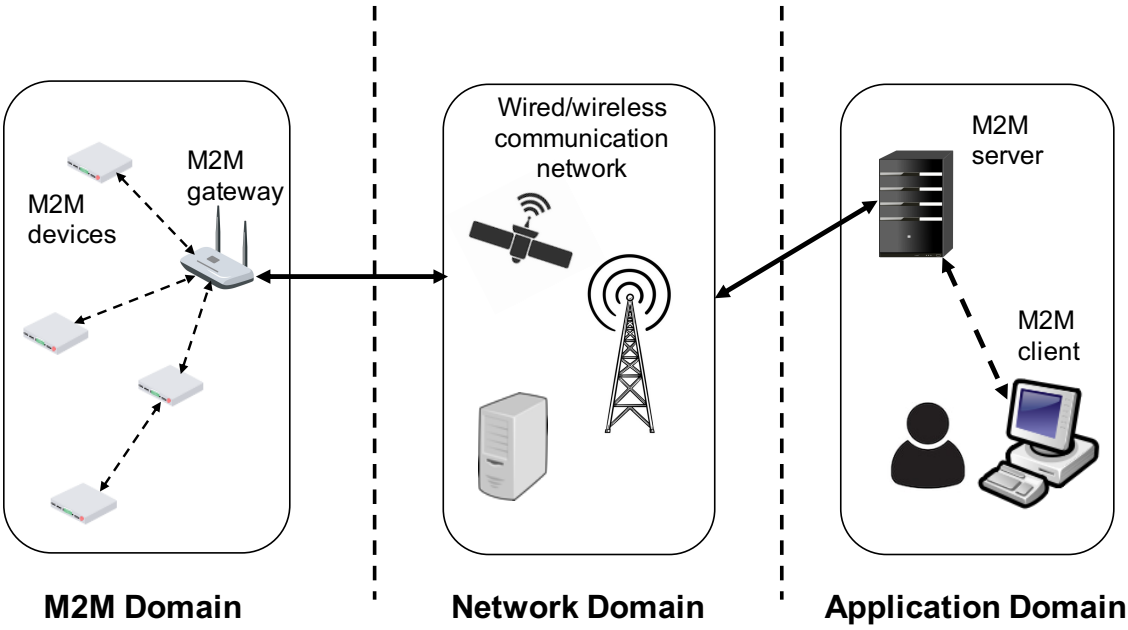

Figure 1.1: M2M communications architecture.

1.2.3 Standards for M2M Communications

Standardization for M2M communications is proceeding in standard developing organizations such as 3rd Generation Partnership Project (3GPP), the Institute of Electrical and Electronics Engineers (IEEE), and ETSI [13]. The main focus of 3GPP standardization efforts is improving the radio access network for enabling M2M communications. On the other hand, the IEEE focuses on enhancements required to support M2M applications such as for 802.16 networks. ETSI is considering internet connectivity solutions for M2M communications through the development of functional architectures and the service middleware layer.

1.3 Cognitive Radio for M2M Communications

The characteristics of M2M communications are distinct from those of traditional cellular communications, also known as Human-to-Human (H2H) communications, that cellular networks were originally designed for [43]. M2M communications are characterized by infrequent transmissions of small amounts of data, delay tolerance, and higher traffic demand on the uplink channel while the majority of the cellular traffic is on the downlink channel. These characteristics in addition to the exponentially growing number of MTCDs connected to the network infrastructure pose significant challenges to cellular networks in terms of power consumption, interference, congestion in the Random Access Channel (RACH), spectrum scarcity, and energy efficiency [5] [44].

Cognitive Radio (CR) [14] is a promising technology for M2M communications [42]. CR can be utilized in a wide range of M2M applications such as public safety, disaster management, healthcare and medical applications, smart grids, and smart metering [45], [46]. Moreover, applications in the areas of military, emergency networks, and Cognitive Radio Vehicular Ad Hoc Networks (CR-VANET) can benefit from CR technology [47]. Integrating CR with M2M communications results in Cognitive M2M (CM2M) communications [92]. CM2M can be used to overcome some of the main challenges of M2M communications as discussed below.

- Spectrum Scarcity: The continuously growing number of MTCDs creates challenges for existing communication networks such as spectrum congestion. With CR opportunistic spectrum access, existing spectrum can be efficiently utilized to overcome spectrum shortages and support reliable data transmission.

- Connectivity and Coverage: In some M2M applications such as smart grids, wireless connectivity can be an issue due to the massive number of MTCDs that can be deployed. CM2M networks can effectively overcome this issue through dynamic spectrum access of communication channels. For example, CR can be used to explore alternate spectrum opportunities such as TV White Space (TVWS) [15], which refers to the large portions of the UHF/VHF spectrum that are available on a geographical basis as a result of switchover from analog to digital TV. TVWS is attractive because of the significant bandwidth availability and superior propagation characteristics.

- Interference: The massive number of MTCDs connected to the M2M network in addition to the coexistence of MTCDs and cellular users creates significant interference issues. Interference will not only degrade the performance of the M2M network, but also affect cellular communications. CR can be utilized to overcome these issues through Spectrum Sharing (SS) of the licensed spectrum occupied by cellular users or dynamic spectrum access in under-utilized spectrum such as TVWS.

- Green Requirements: Energy efficiency is a fundamental requirement for battery-operated MTCDs to enhance network lifetime. The use of CR has been demonstrated to be energy-efficient as MTCDs can adaptively adjust their transmission power levels based on operational environments [16].

- Device Heterogeneity: M2M networks have diverse QoS requirements based on the types of application. This creates interoperability issues between MTCDs with different networking protocols. CR can be used to handle the MTCDs and protocol heterogeneity as M2M networks will be more efficient and flexible if the MTCDs can communicate using CR.

In CM2M communications, MTCDs utilize the spectrum resources of Cellular User Equipment (CUE). This is achieved by sharing the spectrum resources of the CUE with the MTCDs where the CUE are considered to be Primary Users (PUs) while the MTCDs are Secondary Users (SUs). Spectrum Sharing (SS) models in CM2M communications include interweave, overlay, and underlay. In interweave SS, also known as opportunistic SS, SUs sense the spectrum of PUs and opportunistically access the unused spectrum resources, also called spectrum holes, over time and/or space [39]. In overlay SS, PUs and SUs simultaneously communicate as long as PU

transmissions are not affected. To facilitate this, SUs can identify PU signals as they have prior knowledge regarding transmissions. Hence, a SU can transmit at any power level and employ coding to protect against PU interference [40]. In underlay SS, SUs share PU spectrum while keeping their interference below an interference threshold. In M2M environments, underlay SS is promising due to its high spectral and power efficiencies with low transmission delay [41].

1.4 Energy Efficiency in M2M Communications

M2M communications rely on MTCDs in the M2M domain to collect data and send it to the BS in the network domain via a wired or wireless network. Due to the massive number of MTCDs involved in M2M communications, Energy Efficiency (EE) is a challenging issue especially in the M2M domain [12]. Moreover, the majority of MTCDs are battery-operated and may be deployed in remote areas where frequent access for battery replacement and maintenance is infeasible [90]. The success of M2M communications relies on the availability of the MTCDs. Therefore, EE for M2M communications is crucial to prolong the network lifetime [92]. In M2M networks, communications dominate energy consumption, hence EE can be improved by employing efficient communication protocols [63]. Activity scheduling is another approach that can be used to achieve energy-efficient M2M networks. In activity scheduling, some MTCDs are switched into sleep-mode and only a subset of MTCDs remain active to preserve the functionality of the network [64].

Energy-efficient M2M communications can also be attained by adjusting the MTCD transmit power such that their power consumption is minimized while achieving the network performance requirements. Another approach to achieve energy-efficient M2M communication is clustering. In clustered M2M communication, MTCDs are grouped into clusters based on their spatial locations or Quality-of-Service (QoS) requirements, and communicate with the Base Station (BS) via a Machine-Type Communication Gateway (MTCG). An MTCG collects data from the MTCDs within its cluster and forwards it to the BS. This reduces the number of direct accesses to the BS and the MTCD power consumption. Moreover, CR can improve the EE in M2M networks where the MTCDs can adaptively adjust their transmit power levels based on the operating environment as in underlay CR.

1.5 Scope of the Work

M2M communications create many challenges for existing communication networks such as managing the increasing number of devices, meeting resource constraints, and providing the required QoS. Therefore, solutions need to be developed to handle M2M communications so their QoS requirements are met with minimal impact on coexisting H2H communications.

Several non-cellular wireless technologies including WiFi, Bluetooth, and Zigbee have been considered for enabling M2M communications, but they are inefficient and insufficient to support the growing demands of M2M applications [87]. On the other hand, cellular networks provide extended wireless connectivity, more flexibility for radio resource management, and higher capacity to support M2M deployment. However, the rapidly growing number of MTCDs in addition to the distinct QoS requirements of M2M communications that are different from those of H2H communications make the problem of resource allocation challenging. Since cellular networks were originally designed for H2H communications, resource allocation for coexisting H2H/M2M communications has to consider the QoS requirements of M2M communications while controlling the impact on H2H communications.

This dissertation investigates the challenges of resource allocation for M2M communications in cellular networks in terms of meeting the diverse QoS requirements, and achieving spectral and energy efficiencies. Efficient resource allocation algorithms to solve these challenges are proposed and evaluated. The focus of Chapter 2 is resource allocation for coexisting H2H/M2M communications in cellular networks. A two-phase resource allocation algorithm is proposed to meet the QoS requirements of H2H traffic and delay-sensitive M2M traffic while ensuring fairness for delay-tolerant M2M traffic.

CM2M communications is expected to be indispensable in IoT and M2M networks. With the large number of connected devices, spectrum scarcity is a major challenge. CM2M is envisioned to improve spectrum utilization through opportunistic spectrum access capabilities. Moreover, clustering can be used to coordinate communications between the Base Station (BS) and MTCDs to overcome congestion and manage the coexistence of the M2M and CUE. Chapter 3 addresses clustered CM2M communications underlaying cellular networks. A joint resource-power allocation problem is formulated and solved using a two-phase resource and power allocation scheme. The goal is to maximize the sum-rate of the CUE and clustered MTCDs while satisfying

MTCD minimum data rate requirements, MTCD transmit power limits, and CUE interference thresholds.

Energy Efficiency (EE) in M2M communications is crucial to prolonging network lifetime where the majority of MTCDs are battery-operated. This is the focus of Chapter 4 that investigates power allocation for energy-efficient CM2M communications underlaying cellular networks. Two power allocation problems are proposed where the first targets MTCD power consumption minimization while the second considers MTCD EE maximization subject to MTCD transmit power constraints, MTCD minimum data rate requirements, and CUE interference limits.

1.6 Dissertation Contributions and Organization

The contributions of this dissertation are as follows.

1. The problem of resource allocation for coexisting H2H/M2M communications in cellular networks is addressed in Chapter 2. A two-phase resource allocation algorithm is proposed. The proposed algorithm targets maximizing the total achievable data rate of H2H UE while considering the delay constraints of delay-sensitive MTCDs in order to control the impact of M2M traffic on H2H traffic. Moreover, dynamic resource allocation is performed in order to adapt to current traffic loads and channel conditions. Hence, meet the diverse QoS requirements of H2H and M2M traffic. The proposed algorithm also guarantees that a minimum amount of resources is allocated to delay-tolerant MTCDs to ensures fairness.

2. The problem of resource allocation for clustered CM2M communications underlaying cellular networks is addressed in Chapter 3. A clustered CM2M model is proposed based on the spatial locations of the MTCDs to overcome the problem of congestion. Underlay CR is employed so that MTCDs within each cluster share the spectrum of neighbouring CUE to overcome spectrum scarcity and improve spectrum utilization. A two-phase resource and power allocation scheme is proposed to maximize the sum-rate of CUE and clustered MTCDs while adhering to MTCD minimum data rate requirements, MTCD transmit power limits, and CUE interference constraints.

3. The problem of power allocation for energy-efficient CM2M communications

underlaying cellular networks is addressed in Chapter 4. Two power allocation problems are proposed where the first targets MTCD power consumption minimization while the second considers MTCD EE maximization subject to MTCD transmit power constraints, MTCD minimum data rate requirements, and CUE interference limits. Underlay CR is employed so that MTCDs can exploit spatial spectrum opportunities and share the spectrum of surrounding CUE to overcome spectrum scarcity. The proposed power consumption minimization problem is transformed into a Geometric Programming (GP) problem and solved iteratively. The proposed EE maximization problem is a nonconvex fractional programming problem. Hence, parametric transformation is used to convert it into an equivalent convex form and this is solved using an iterative approach.

The detailed contributions of each chapter are presented therein.

Chapter 2

Optimal Resource Allocation in Cellular Networks with H2H/M2M Coexistence

M2M communications via cellular networks can provide MTCDs with significant benefits including extended wireless connectivity and reliable communication links. However, cellular networks were originally designed to support only H2H communications [43]. Thus, enabling M2M communications poses a significant challenge to cellular networks. For example, the majority of H2H traffic is on the downlink channel. This differs from M2M traffic which generates higher demands on the uplink channel as most MTCDs are involved in sensing and measurement (data collection) tasks. Therefore, implementing M2M communications in cellular networks can lead to congestion in the physical channel, resulting in low access rates [5].

Another challenge is the diverse Quality-of-Service (QoS) requirements of H2H and M2M communications [26]. The QoS metric for H2H traffic is the achievable data rate, while for M2M traffic it varies based on the MTCD application. M2M traffic can be classified into two categories: event-based and time-based [9].

Event-based traffic includes emergency alarms and critical messages from applications such as e-health monitoring and vehicle tracking where messages are sent only when events occur. These applications require reliable real-time resources and impose strict delay requirements. MTCDs in this category are called delay-sensitive and have QoS requirements in terms of strict delay limits and achievable data rates. Conversely, time-based traffic comprises infrequent small data transmissions at regu-

lar time intervals that span from a few milliseconds to minutes or hours [10]. MTCDs in this category are known as delay-tolerant and require reliable communications to satisfy their delay bounds.

2.1 Related Work

Several studies have considered the resource allocation problem for M2M communications. Some examined H2H/M2M coexistence and the effect of M2M communications on the QoS of H2H communications. Others investigated M2M communications only and focused on meeting the QoS requirements of delay-sensitive MTCDs in terms of delay limits while providing an acceptable level of fairness for delay-tolerant MTCDs.

Resource allocation for coexistent H2H/M2M users was considered in [10] where radio resources were partitioned between H2H User Equipment (H2H UE) and MTCDs before scheduling. Uplink scheduling is performed separately while considering throughput, fairness and the maximum tolerable delay. A balanced alternating technique to achieve the throughput and delay requirements of M2M traffic was considered in [17]. However, channel conditions for H2H UE and MTCDs were not jointly considered.

In [38], resource allocation for delay-sensitive MTCDs coexisting with H2H UE was considered. The aggregate throughput of H2H UE was maximized while taking into account the channel conditions and statistical QoS requirements of delay-sensitive MTCDs. However, delay-tolerant MTCDs were not considered. Energy-efficient resource allocation was considered in [26]. The proposed algorithm maximized the bits-per-joule capacity under statistical delay constraints for M2M/H2H coexistence. An energy-efficient power control and time scheduling scheme with Non-Orthogonal Multiple Access (NOMA) and energy harvesting was proposed in [111]. The total energy consumption of both MTCDs and H2H UE was minimized under throughput and harvested energy constraints. Resource allocation for relay-aided MTCDs coexisting with H2H UE was examined in [19]. The end-to-end data rate of H2H UE and delay-sensitive MTCDs was maximized under minimum data rate and statistical QoS constraints.

Some approaches to resource allocation focus on M2M communications only. M2M scheduling from a queueing theory perspective was examined in [31] considering statistical delay requirements but ignoring channel conditions. In [20], group-based M2M resource allocation was proposed which classifies MTCDs based on their transmission protocol and QoS requirements. Then, joint power and resource allocation was

performed to maximize the sum-throughput. In [21], a predictive resource allocation algorithm was presented for event-based M2M applications in the Long Term Evolution (LTE) uplink. A Threshold Controlled Access (TCA) algorithm for M2M communications in 5G networks was presented in [22]. In this algorithm, MTCDs choose Resource Blocks (RBs) based on their battery status and associated application power profile.

A two-layer architecture for cluster-grouped MTCDs was considered in [23] where MTCDs transmit their data to a cluster gateway which in turn relays the data to the Base Station (BS). Then, a resource allocation problem to minimize packet loss probability and maximize spectral efficiency was proposed. Hybrid Non-Orthogonal Random Access and Data Transmission (NORA-DT) was proposed in [24] to overcome the signalling overhead and resource allocation problem for M2M communications. In [25], a hybrid network model considering M2M cellular communications coexisting with Device-to-Device (D2D) communications was examined. A stochastic geometry approach was employed to obtain rate expressions and evaluate network performance in terms of effective throughput.

In this chapter, resource allocation for M2M communications coexisting with H2H communications is investigated. A two-phase resource allocation algorithm for H2H/M2M coexistence is proposed. The first phase performs joint power-resource allocation in order to satisfy the QoS requirements of H2H traffic while considering the delay constraints of delay-sensitive M2M traffic. Then, the second phase focuses on meeting the QoS requirements of M2M traffic. Different from the work discussed above, the proposed algorithm addresses the diverse QoS requirements of H2H traffic and M2M traffic (both delay-sensitive and delay-tolerant), and ensures fairness while considering channel conditions. This is summarized in Table 2.1 which indicates the Key Performance Indicators (KPIs) addressed in the related work.

2.2 Contributions of the Chapter

The contributions of this chapter are as follows.

- To meet the diverse QoS requirements of H2H and M2M traffic, resource allocation is dynamically performed over two phases in order to adapt to current traffic loads and channel conditions.

- To control the impact of M2M traffic on H2H traffic, the first phase of resource

Table 2.1: Key Performance Indicators (KPIs) Addressed in the Related Work.

Reference	**Impact of M2M Traffic on H2H Traffic**	**Delay-Sensitive M2M Traffic**	**Delay-Tolerant M2M Traffic**	**Delay Requirements**	**Fairness**	**Channel Conditions**
Proposed algorithm	✓	✓	✓	✓	✓	✓
[31]	✓	✓	✓	✓	✓	×
[88]	✓	✓	✓	✓	✓	×
[17]	✓	✓	✓	✓	×	✓
[38]	✓	✓	×	✓	×	✓
[20]	×	N/A *	N/A*	✓	✓	✓
[19]	✓	✓	×	✓	×	✓
[21]	×	✓	×	✓	×	×
[26]	✓	N/A*	N/A*	✓	×	✓
[18]	✓	N/A*	N/A*	✓	×	✓
[22]	×	N/A*	N/A*	×	×	✓
[24]	×	N/A*	N/A*	×	×	✓
[23]	×	N/A*	N/A*	×	×	✓
[25]	✓	N/A*	N/A*	×	×	✓

* M2M communications is considered without specifying the type of MTCDs

allocation targets maximizing the total achievable data rate of H2H UE while considering the delay constraints of delay-sensitive MTCDs.

- To meet the QoS requirements of delay-sensitive MTCDs, the second phase of resource allocation focuses on maximizing their total achievable data rate within the required delay limits.
- To overcome the problem of resource starvation and ensure fairness for delay-tolerant MTCDs, a minimum amount of resources is guaranteed for delay-tolerant MTCDs in the second phase.
- Simulation results are presented which show that the proposed algorithm is able to balance the demands of M2M and H2H traffic, meet their diverse QoS requirements, and ensure fairness for delay-tolerant M2M traffic.

2.3 System Model

Consider uplink traffic in a single cell of an LTE network with H2H and M2M co-existence. In LTE networks, radio resources are divided in the time and frequency domains, as shown in Figure 2.1. In the time domain, radio resources are distributed every Transmission Time Interval (TTI) which has a duration of 1 ms and consists of two slots. One LTE frame consists of 10 TTIs or 20 slots. In the frequency domain, the available bandwidth is divided into a number of sub-channels each has 12 subcarriers. Each sub-channel has 180 KHz bandwidth and along with 7 symbols in the time domain constitutes a Resource Block (RB). In the proposed model, H2H UE and MTCDs directly communicate with the Base Station (BS). The channel is assumed to experience path loss and shadowing in addition to AWGN. As shown in Figure 2.2, once a user has data to transmit, it sends an access request to the BS. The QoS classifier then assigns the user to the corresponding QoS class based on its QoS requirements. There are three QoS classes: H2H, delay-sensitive M2M, and delay-tolerant M2M. The network scheduler then assigns available RBs to the users based on the proposed resource allocation algorithm.

Available RBs are dynamically allocated to H2H UE and MTCDs over two phases as shown in Figure 2.3. The motivation for this approach is as follows.

- H2H traffic has priority over M2M traffic so sufficient resources should be assigned to H2H UE in order to satisfy their QoS requirements and control the

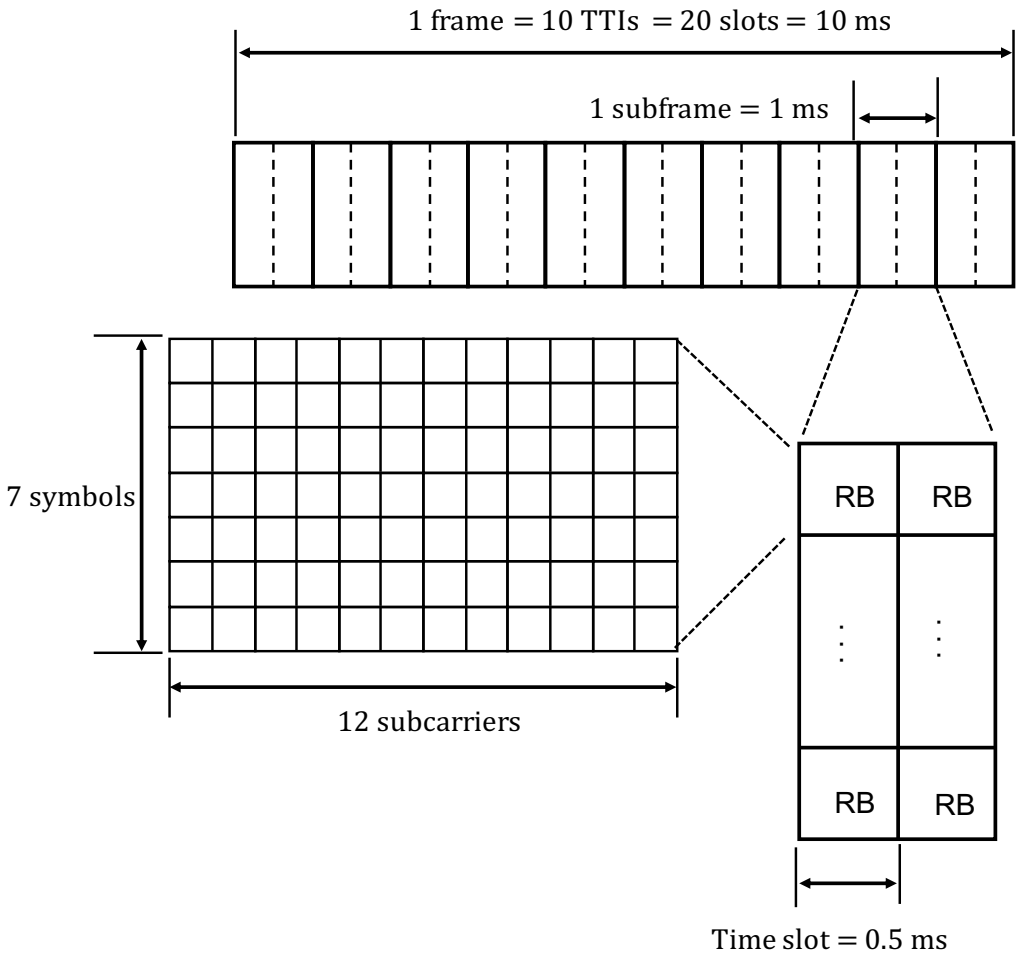

Figure 2.1: RB in LTE frame.

impact of M2M traffic on H2H traffic.

- Resource allocation for coexistent H2H and M2M traffic is challenging due to their diverse QoS requirements. The achievable data rate is the key QoS measure for the former while delay is important for the latter.

- The tolerable delay limit of M2M traffic, also known as the maximum delay-bound, ranges from tens of milliseconds to minutes or hours depending on the M2M application [27]. Therefore, higher priority in assigning resources should be given to delay-sensitive MTCDs while ensuring a certain level of fairness for delay-tolerant MTCDs.

In the first phase of the allocation process, available RBs are divided between H2H and M2M traffic such that the total achievable data rate of H2H UE is maximized

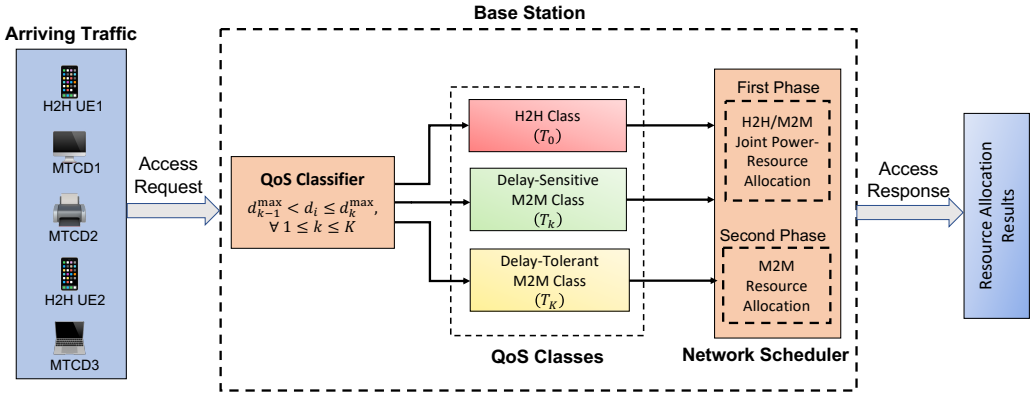

Figure 2.2: System model.

while considering the delay constraints of delay-sensitive MTCDs. In the second phase, the resources allocated to M2M traffic during the first phase are divided between the two M2M classes. The focus of this phase is on the QoS of delay-sensitive MTCDs so that their total achievable data rate is maximized considering their delay limits while ensuring fairness for delay-tolerant MTCDs.

The advantages of the proposed approach are as follows. First, it guarantees the QoS of H2H UE in terms of the total achievable data rate while considering the QoS of delay-sensitive MTCDs. Second, it adapts to the current M2M traffic load, i.e., when the network is heavily loaded with delay-sensitive M2M traffic, delay-sensitive MTCDs are allocated more RBs to accommodate their traffic and fulfill their delay requirements. On the other hand, when the delay-sensitive M2M traffic load is light, more RBs are allocated to delay-tolerant MTCDs.

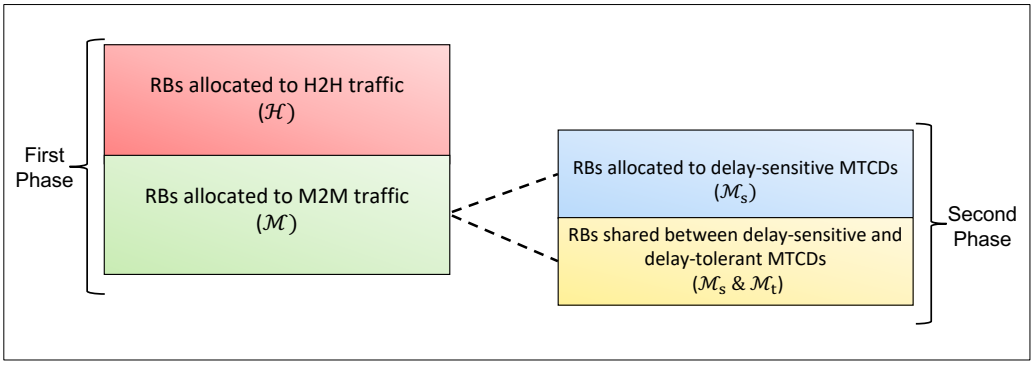

Figure 2.3: The proposed resource allocation approach.

2.4 Problem Formulation

Assume the set of traffic types within the network is indexed by $\mathcal{K} = \{T_0, \cdots, T_k, \cdots, T_K\}$, where T_k is the kth traffic type with maximum delay-bound d_k. High-priority H2H traffic corresponds to T_0 while delay-tolerant M2M traffic belongs to T_K. Moreover, delay-sensitive M2M traffic with delay limit $d_i^{\max}$ belongs to T_k if $d_{k-1} < d_i^{\max} \leq d_k,\ \forall\ 1 \leq k \leq K-1$. Let $\mathcal{N} = \{1, \cdots, n, \cdots, N\}$ be the set of available RBs. The set of users, including H2H UE and MTCDs, is indexed by $\mathcal{I} = \{1, \cdots, i, \cdots, I\}$, and is divided into three subsets. Subset $\mathcal{H}$ is for H2H UE with minimum data rate requirements, $R_i^{\mathrm{h,min}}\ \forall,\ i \in \mathcal{H}$. Subset $\mathcal{M}_{\mathrm{s}}$ includes delay-sensitive MTCDs with minimum data rate requirements, $R_i^{\mathrm{ms,min}}(d_i^{\max}, \varepsilon_i^{\max}),\ \forall\ i \in \mathcal{M}_s$. Finally, subset $\mathcal{M}_{\mathrm{t}}$ contains delay-tolerant MTCDs with minimum data rate requirements, $R_i^{\mathrm{mt,min}},\ \forall\ i \in \mathcal{M}_t$.

The delay constraints of delay-sensitive MTCDs are expressed in terms of the Probability of Delay-Bound Violation (PDBV). This is defined as the probability that the delay d_i exceeds the maximum delay-bound $d_i^{\max}$ which must be less than the maximum delay-violation threshold $\varepsilon_i^{\max}$.

The resource allocation process shown in Figure 2.2 can be divided into two sub-problems. The first sub-problem solves the first phase of the allocation process and is formulated as

$$\max_{x_{i,n}, p_{i,n}} \quad \sum_{i \in \mathcal{H}} \sum_{n \in \mathcal{N}} x_{i,n} R_{i,n}(p_{i,n}) \tag{2.1}$$

$$\text{s.t.} \quad x_{i,n} \in \{0,1\},\ \ \forall\ i \in \mathcal{I}, \forall n \in \mathcal{N} \tag{2.1a}$$

$$\sum_{i \in \mathcal{I}} x_{i,n} \leq 1,\ \ \forall\ n \in \mathcal{N} \tag{2.1b}$$

$$\sum_{n \in \mathcal{N}} x_{i,n} R_{i,n}(p_{i,n}) \geq R_i^{\mathrm{h,min}},\ \ \forall\ i \in \mathcal{H} \tag{2.1c}$$

$$\Pr(d_i > d_i^{\max}) \leq \varepsilon_i^{\max},\ \ \forall\ i \in \mathcal{M}_{\mathrm{s}}. \tag{2.1d}$$

The objective function to be maximized in (2.1) is the total achievable data rate of H2H UE in subset $\mathcal{H}$, with $R_{i,n}$ defined as the achievable data rate of the ith H2H UE over the nth RB. Decision variables $p_{i,n}$ and $x_{i,n}$ are defined as the transmit power of the ith H2H UE/MTCD over the nth RB and the binary resource allocation indicator of the ith H2H UE/MTCD over the nth RB, respectively. If the nth RB is assigned to the ith H2H UE/MTCD, then $x_{i,n}$ is equal to 1, otherwise it is equal to 0. Constraint (2.1b) indicates that an RB can be assigned to one H2H UE/MTCD only.

The minimum guaranteed data rate requirement of the ith H2H UE in subset $\mathcal{H}$ is considered in constraint (2.1c). Constraint (2.1d) considers the delay constraints of delay-sensitive MTCDs in subset $\mathcal{M}_{\rm s}$ by ensuring that the PDBV of the ith delay-sensitive MTCD is less than a maximum delay-violation threshold, $\varepsilon_i^{\max}$.

The second sub-problem solves the second phase of the resource allocation process. The resources allocated to M2M traffic during the first phase are partitioned between the two MTCD groups based on their QoS requirements. The corresponding optimization problem can be formulated as

$$\max_{x_{i,n}^{\rm m2m}} \sum_{i\in\mathcal{M}_{\rm s}} w_i \sum_{n\in\mathcal{N}_{\rm m2m}} x_{i,n}^{\rm m2m} R_{i,n} \tag{2.2}$$

$$\text{s.t. } x_{i,n}^{\rm m2m} \in \{0,1\}, \ \forall\, i \in \mathcal{M}_{\rm s}\cup\mathcal{M}_{\rm t}, \ \forall n \in \mathcal{N}_{\rm m2m} \tag{2.2a}$$

$$\sum_{i\in\mathcal{M}_{\rm s}\cup\mathcal{M}_{\rm t}} x_{i,n}^{\rm m2m} \le 1, \ \ \forall\, n \in \mathcal{N}_{\rm m2m} \tag{2.2b}$$

$$\Pr(d_i > d_i^{\max}) \le \varepsilon_i^{\max}, \ \ \forall\, i \in \mathcal{M}_{\rm s} \tag{2.2c}$$

$$\sum_{i\in\mathcal{M}_{\rm t}} \sum_{n\in\mathcal{N}_{\rm m2m}} x_{i,n}^{\rm m2m} R_{i,n} \ge \rho R^{\rm mt,min}. \tag{2.2d}$$

The objective function in (2.2) is the weighted total achievable data rate of the MTCDs in subset $\mathcal{M}_{\rm s}$. This objective function considers the delay requirements of the ith MTCD through weight w_i where $w_i \in [0,1]$. The value of w_i is based on the corresponding maximum delay-violation threshold $\varepsilon_i^{\max}$ of the ith MTCD, i.e.

$$w_i = 1 - \varepsilon_i^{\max}, \quad \forall\, i \in \mathcal{M}_{\rm s}. \tag{2.3}$$

Large values of w_i correspond to strict delay limits while a small w_i indicates a loose delay requirement. For example, the ith MTCD can tolerate minimal delay when $w_i \approx 1$, while $w_i \approx 0$ implies that the MTCD can tolerate a long delay.

The reason for considering the achievable data rate of delay-sensitive M2M traffic is to ensure that critical messages involved in this type of M2M traffic, such as emergency alarms and messages from applications like e-health monitoring and vehicle tracking, are delivered within the specified delay limits. The binary resource allocation indicator $x_{i,n}^{\rm m2m} \in \mathbf{X}_{\rm m2m}$ where $\mathbf{X}_{\rm m2m}$ is an $I_{\rm m2m} \times N_{\rm m2m}$ matrix with $I_{\rm m2m}$ as the total number of MTCDs (both delay-sensitive and delay-tolerant) and $N_{\rm m2m}$ is the number of RBs assigned to M2M traffic during the first phase of the resource allocation process. Constrant (2.2b) indicates that an RB can be assigned to one

MTCD only.

The PDBV of each MTCD in subset $\mathcal{M}_{\rm s}$ should be below the corresponding maximum delay-violation threshold $\varepsilon_i^{\max}$ as shown in constraint (2.2c). Fairness is provided to delay-tolerant MTCDs in subset $\mathcal{M}_{\rm t}$ through constraint (2.2d), where $\rho \in [0,1]$ is the fairness factor. This ensures that these MTCDs get a minimal amount of resources even when the network is heavily loaded with H2H and delay-sensitive M2M traffic. Since the number of delay-sensitive MTCDs in real systems is far less than the number of delay-tolerant MTCDs [28], the resource allocation in (2.2) guarantees that every MTCD gets sufficient resources to satisfy its QoS requirements.

The maximum achievable data rate of the ith H2H UE/MTCD over the nth RB can be expressed as

$$R_{i,n} = \ B_{\rm RB} \log_2(1+\gamma_{i,n}), \tag{2.4}$$

where $\gamma_{i,n}$ is the instantaneous SNR of the ith H2H UE/MTCD over the nth RB which is given by

$$\gamma_{i,n} = \frac{p_{i,n}|h_{i,n}|^2}{\sigma^2}, \tag{2.5}$$

where $|h_{i,n}|^2$ is the channel power gain, σ^2 is the Additive White Gaussian Noise (AWGN) power, and $p_{i,n}$ is the transmit power of the ith H2H UE/MTCD over the nth RB.

According to [30], the arrival process of delay-sensitive M2M traffic can be modeled as a Poisson process. Hence, a delay-sensitive MTCD can be modeled as an M/D/1 queue with arrival rate λ_i and service rate R_i, $\forall\, i \in \mathcal{M}_s$ [31]. The MTCDs queues are considered simultaneously during the resource allocation process. With an M/D/1 queue, the wait time of the ith MTCD (W_i) is given by [32]

$$\Pr(W_i \le t_i) = (1-\lambda_i\tau_i)\sum_{v=0}^{z}\frac{[-\lambda_i(t_i-v\tau_i)]^v}{v!}e^{\lambda_i(t_i-v\tau_i)}, \ \ \forall i \in \mathcal{M}_{\rm s}, \tag{2.6}$$

where z is an integer satisfying $z\tau_i \le t_i \le (z+1)\tau_i$ and $\tau_i = 1/R_i$ is the deterministic service time. The total delay of the ith MTCD can be expressed as

$$d_i = W_i + \tau_i, \quad \forall\, i \in \mathcal{M}_{\rm s}. \tag{2.7}$$

Replacing t_i by ($d_i^{\max}$ - τ_i) in (2.6), the PDBV can be expressed as

$$\Pr(d_i > d_i^{\max}) = 1 - (1 - \lambda_i \tau_i) \times \sum_{v=0}^{z} \frac{[-\lambda_i (d_i^{\max} - \tau_i - v\tau_i)]^v}{v!} e^{\lambda_i (d_i^{\max} - \tau_i - v\tau_i)}, \quad \forall i \in \mathcal{M}_{\mathrm{s}}. \tag{2.8}$$

Consequently, optimization sub-problem (2.1) can be reformulated as

$$\max_{x_{i,n}, p_{i,n}} \sum_{i \in \mathcal{H}} \sum_{n \in \mathcal{N}} x_{i,n} \, B_{\mathrm{RB}} \log_2(1 + \frac{p_{i,n} |h_{i,n}|^2}{\sigma^2}) \tag{2.9}$$

$$\text{s.t.} \quad x_{i,n} \in \{0, 1\}, \quad \forall \, i \in \mathcal{I}, \forall \, n \in \mathcal{N} \tag{2.9a}$$

$$\sum_{i \in \mathcal{I}} x_{i,n} \leq 1, \quad \forall \, n \in \mathcal{N} \tag{2.9b}$$

$$\sum_{n \in \mathcal{N}} x_{i,n} \, B_{\mathrm{RB}} \log_2(1 + \frac{p_{i,n} |h_{i,n}|^2}{\sigma^2}) \geq R_i^{\mathrm{h,min}}, \quad \forall i \in \mathcal{H} \tag{2.9c}$$

$$\Pr(d_i > d_i^{\max}) \leq \varepsilon_i^{\max}, \quad \forall \, i \in \mathcal{M}_{\mathrm{s}}. \tag{2.9d}$$

Constraint (2.9d) is satisfied when the achievable data rate of the ith MTCD in subset $\mathcal{M}_{\mathrm{s}}$ is greater than the minimum data rate, $R_i^{\min}(d_i^{\max}, \varepsilon_i^{\max})$, i.e.,

$$R_i \geq R_i^{\mathrm{ms,min}}(d_i^{\max}, \varepsilon_i^{\max}), \quad \forall \, i \in \mathcal{M}_{\mathrm{s}}. \tag{2.10}$$

where $R_i^{\mathrm{ms,min}}(d_i^{\max}, \varepsilon_i^{\max})$ is the minimum achievable data rate of the ith delay-sensitive MTCD that achieves a PDBV equal to the maximum delay-violation threshold $\varepsilon_i^{\max}$ when the maximum delay-bound is $d_i^{\max}$. The value of $R_i^{\mathrm{ms,min}}(d_i^{\max}, \varepsilon_i^{\max})$ is obtained from (2.8). Hence, constraint (2.9d) can be replaced by

$$\sum_{n \in \mathcal{N}} x_{i,n} \, B_{\mathrm{RB}} \log_2(1 + \frac{p_{i,n} |h_{i,n}|^2}{\sigma^2}) \geq R_i^{\mathrm{ms,min}}(d_i^{\max}, \varepsilon_i^{\max}), \quad \forall i \in \mathcal{M}_{\mathrm{s}}. \tag{2.11}$$

The optimization problem in (2.9) is a joint power-resource allocation problem with two decision variables $x_{i,n}$ and $p_{i,n}$ as the binary resource allocation indicator and transmit power of the ith H2H UE/MTCD over the nth RB, respectively.

The transmit power of the ith delay-sensitive MTCD is assigned using (2.9).

Hence, the optimization problem in (2.2) can be reformulated as

$$\max_{x_{i,n}^{\mathrm{m2m}}} \sum_{i\in\mathcal{M}_{\mathrm{s}}} w_i \sum_{n\in\mathcal{N}_{\mathrm{m2m}}} x_{i,n}^{\mathrm{m2m}} R_{i,n} \tag{2.12}$$

$$\text{s.t. } x_{i,n}^{\mathrm{m2m}} \in \{0,1\}, \ \ \forall\, i \in \mathcal{M}_{\mathrm{s}} \cup \mathcal{M}_{\mathrm{t}}, \ \forall\, n \in \mathcal{N}_{\mathrm{m2m}} \tag{2.12a}$$

$$\sum_{i\in\mathcal{M}_{\mathrm{s}}\cup\mathcal{M}_{\mathrm{t}}} x_{i,n}^{\mathrm{m2m}} \leq 1, \ \ \forall\, n \in \mathcal{N}_{\mathrm{m2m}} \tag{2.12b}$$

$$\sum_{n\in\mathcal{N}_{\mathrm{m2m}}} x_{i,n}^{\mathrm{m2m}} R_{i,n} \geq R_i^{\mathrm{ms,min}}(d_i^{\max}, \varepsilon_i^{\max}), \ \ \forall\, i \in \mathcal{M}_{\mathrm{s}} \tag{2.12c}$$

$$\sum_{i\in\mathcal{M}_{\mathrm{t}}} \sum_{n\in\mathcal{N}_{\mathrm{m2m}}} x_{i,n}^{\mathrm{m2m}} R_{i,n} \geq \rho R^{\mathrm{mt,min}}, \tag{2.12d}$$

2.5 Proposed Resource Allocation Algorithm

2.5.1 First Phase: H2H/M2M Joint Power-Resource Allocation

Optimization problem (2.9) is a Mixed Integer Non-Linear Programming (MINLP) problem with decision variables $x_{i,n}$ and $p_{i,n}$. To solve this problem, alternating optimization is employed [33]. In alternating optimization, the original optimization problem is divided into two sub-problems. Each sub-problem is with respect to one decision variable while the other decision variable is considered constant. For (2.9), the first sub-problem is a resource allocation problem with respect to $x_{i,n}$ while $p_{i,n}$ is fixed, and the second sub-problem is a power allocation problem with decision variable $p_{i,n}$ while $x_{i,n}$ is fixed.

For given $p_{i,n}$, the objective function in (2.9) is a function of $x_{i,n}$ only. Therefore, finding the optimal resource allocation is to solve the following optimization problem

$$\max_{x_{i,n}} \quad \sum_{i\in\mathcal{H}}\sum_{n\in\mathcal{N}} x_{i,n}\, B_{\mathrm{RB}} \log_2(1+\frac{p_{i,n}|h_{i,n}|^2}{\sigma^2}) \tag{2.13}$$

$$\text{s.t.} \quad x_{i,n}\in\{0,1\}, \quad \forall\, i\in\mathcal{I}, \forall\, n\in\mathcal{N} \tag{2.13a}$$

$$\sum_{i\in\mathcal{I}} x_{i,n} \leq 1, \quad \forall\, n\in\mathcal{N} \tag{2.13b}$$

$$\sum_{n\in\mathcal{N}} x_{i,n}\, B_{\mathrm{RB}} \log_2(1+\frac{p_{i,n}|h_{i,n}|^2}{\sigma^2}) \geq R_i^{\mathrm{h,min}}, \quad \forall i\in\mathcal{H} \tag{2.13c}$$

$$\sum_{n\in\mathcal{N}} x_{i,n}\, B_{\mathrm{RB}} \log_2(1+\frac{p_{i,n}|h_{i,n}|^2}{\sigma^2}) \geq R_i^{\mathrm{ms,min}}(d_i^{\max}, \varepsilon_i^{\max}), \quad \forall i\in\mathcal{M}_{\mathrm{s}}. \tag{2.13d}$$

The solution to (2.13) is

$$x_{i,n} = \begin{cases} 1 & n = \arg\max_{n'\in\mathcal{N}} \log_2(1+\frac{p_{i,n'}|h_{i,n'}|^2}{\sigma^2}),\ \forall\, i\in\mathcal{H}, \\ 0 & \text{otherwise}, \end{cases} \tag{2.14}$$

which suggests that the ith H2H UE should be assigned the RB n that maximizes its achievable data rate $R_{i,n}$ for the given power allocation $p_{i,n}$.

For given $x_{i,n}$, the optimal power allocation is obtained by solving the following problem

$$\max_{p_{i,n}} \quad \sum_{i\in\mathcal{H}}\sum_{n\in\mathcal{N}} x_{i,n}\, B_{\mathrm{RB}} \log_2(1+\frac{p_{i,n}|h_{i,n}|^2}{\sigma^2}) \tag{2.15}$$

$$\text{s.t.} \quad \sum_{n\in\mathcal{N}} x_{i,n}\, B_{\mathrm{RB}} \log_2(1+\frac{p_{i,n}|h_{i,n}|^2}{\sigma^2}) \geq R_i^{\mathrm{h,min}}, \quad \forall i\in\mathcal{H} \tag{2.15a}$$

$$\sum_{n\in\mathcal{N}} x_{i,n}\, B_{\mathrm{RB}} \log_2(1+\frac{p_{i,n}|h_{i,n}|^2}{\sigma^2}) \geq R_i^{\mathrm{ms,min}}(d_i^{\max}, \varepsilon_i^{\max}), \quad \forall i\in\mathcal{M}_{\mathrm{s}}. \tag{2.15b}$$

For notational simplicity, we replace $\frac{p_{i,n}|h_{i,n}|^2}{\sigma^2}$ with $\gamma_{i,n}$ and drop the constant B_{RB}. When the SNR is greater than 0 dB, which is the case in our model, $(1+\gamma_{i,n})$ can be replaced by $\gamma_{i,n}$ [34]. Then, (2.15) can be reformulated as

$$\max_{p_{i,n}} \sum_{i\in\mathcal{H}}\sum_{n\in\mathcal{N}} \log_2 \gamma_{i,n}^{(x_{i,n})} \tag{2.16}$$

$$\text{s.t.} \sum_{n\in\mathcal{N}} \log_2 \gamma_{i,n}^{(x_{i,n})} \geq R_i^{\text{h,min}}, \ \ \forall\, i \in \mathcal{H} \tag{2.16a}$$

$$\sum_{n\in\mathcal{N}} \log_2 \gamma_{i,n}^{(x_{i,n})} \geq R_i^{\text{ms,min}}(d_i^{\max}, \varepsilon_i^{\max}), \ \ \forall i \in \mathcal{M}_{\text{s}}. \tag{2.16b}$$

The objective function in (2.16) can be expressed as

$$\sum_{i\in\mathcal{H}}\sum_{n\in\mathcal{N}} \log_2 \gamma_{i,n}^{(x_{i,n})} = \log_2 \prod_{i\in\mathcal{H}}\prod_{n\in\mathcal{N}} \gamma_{i,n}^{(x_{i,n})}, \tag{2.17}$$

so the objective function in (2.16) can be replaced by the product of the $\gamma_{i,n}$, and this is equivalent to minimizing the product of the $\frac{1}{\gamma_{i,n}}$. Hence, (2.16) is equivalent to

$$\min_{p_{i,n}} \ \frac{1}{\prod_{i\in\mathcal{H}}\prod_{n\in\mathcal{N}} \gamma_{i,n}^{(x_{i,n})}} \tag{2.18}$$

$$\text{s.t.} \ \log_2 \prod_{n\in\mathcal{N}} \gamma_{i,n}^{(x_{i,n})} \geq R_i^{\text{h,min}}, \ \ \forall\, i \in \mathcal{H} \tag{2.18a}$$

$$\log_2 \prod_{n\in\mathcal{N}} \gamma_{i,n}^{(x_{i,n})} \geq R_i^{\text{ms,min}}(d_i^{\max}, \varepsilon_i^{\max}), \ \ \forall i \in \mathcal{M}_{\text{s}}\ , \tag{2.18b}$$

and rearranging yields

$$\min_{p_{i,n}} \ \frac{1}{\prod_{i\in\mathcal{H}}\prod_{n\in\mathcal{N}} \gamma_{i,n}^{(x_{i,n})}} \tag{2.19}$$

$$\text{s.t.} \ \frac{1}{a_i \prod_{n\in\mathcal{N}} \gamma_{i,n}^{(x_{i,n})}} \leq 1, \ \forall\, i \in \mathcal{H} \tag{2.19a}$$

$$\frac{1}{b_i \prod_{n\in\mathcal{N}} \gamma_{i,n}^{(x_{i,n})}} \leq 1, \ \forall\, i \in \mathcal{M}_{\text{s}}\ , \tag{2.19b}$$

where $a_i = 2^{-R_i^{\text{h,min}}}$ and $b_i = 2^{-R_i^{\text{ms,min}}(d_i^{\max}, \varepsilon_i^{\max})}$. The optimization problem in (2.19) is a standard form Geometric Programming (GP) problem [35] (see the Appendix A for details).

GP is a type of nonlinear optimization with an objective function and constraints that have a specific form. A standard form GP targets minimizing a posynomial objective function with upper-bound posynomial inequality constraints and monomial

equality constraints. This form often occurs in network resource allocation problems and efficient algorithms such as interior-point methods can be used to solve large scale GP problems with polynomial-time complexity [34].

Now, alternating optimization is used to solve (2.9). First, starting with an initial power allocation $p_{i,n}^{(0)}, \forall\ i \in \mathcal{H} \cup \mathcal{M}_s$, where $p_{i,n}^{(0)}$ is the transmit power that achieves the minimum data rate requirements $R_i^{\mathrm{h,min}}$ and $R_i^{\mathrm{ms,min}}(d_i^{\max}, \varepsilon_i^{\max})$ of the ith H2H UE and delay-sensitive MTCD, respectively. Then, (2.13) is solved to obtain the corresponding resource allocation $x_{i,n}^{(0)}$. Next, the power allocation $p_{i,n}^{(1)}$ is updated by solving (2.19) using $x_{i,n}^{(0)}$. Based on the updated power allocation, the resource allocation is obtained by solving (2.13) again. This procedure is repeated until $p_{i,n}^{(l)} - p_{i,n}^{(l-1)}$ is less than a predefined threshold, ϵ, or the maximum number of iterations, L, is reached. The proposed algorithm is summarized in Algorithm 2.1.

2.5.2 Second Phase: M2M Resource Allocation

The set of resources allocated to M2M traffic using (2.9) is $\mathcal{N}_{\mathrm{m2m}} = \{1, \cdots, n_{\mathrm{m2m}}, \cdots, N_{\mathrm{m2m}}\}$. These resources are optimized to satisfy the QoS requirements of the two MTCD subsets $\mathcal{M}_s$ and $\mathcal{M}_t$. This is formulated in (2.12) as a Binary Linear Programming (BLP) problem with decision variable $x_{i,n}^{\mathrm{m2m}},\ \forall\ i \in \mathcal{M}_s \cup \mathcal{M}_t, n \in N_{\mathrm{m2m}}$.

In BLP problems, the objective function and constraints are linear and the decision variables can only have values 0 or 1. BLP problems are non-convex due to the binary decision variable and algorithms such as Branch and Bound (BB) can be used to obtain an optimal solution. Hence, (2.12) is a non-convex BLP problem. CVX 2.0 is used to solve this problem as it supports solvers that implement a combination of continuous optimization algorithms (such as interior-point methods), and exhaustive search (such as BB) [36]. The algorithm used to solve (2.12) is summarized in Algorithm 2.2, and a summary of the proposed two-phase resource allocation algorithm is given in Algorithm 2.3. A flowchart of the proposed algorithm is given in Figure 2.4.

2.6 Theoretical Performance Evaluation

In this section, the theoretical performance of the proposed algorithm is derived. Assuming the channel experiences shadowing and path loss, the instantaneous channel power gain between the BS and ith H2H UE/MTCD over the nth RB can be modeled

Algorithm 2.1 : Alternating Optimization Algorithm to Solve (2.9)

1 : *Initialization* : $l = 0,\ p_{i,n}^{(0)}, \forall\ i \in \mathcal{H} \cup \mathcal{M}_s$

2 : Solve (2.13) to obtain $x_{i,n}^{(0)}$ using $p_{i,n}^{(0)}$

3 : **for** $l = 1$ to L **do**

4 : Solve (2.19) to obtain $p_{i,n}^{(l)}$ using $x_{i,n}^{(l-1)}$

5 : **if** $|p_{i,n}^{(l)} - p_{i,n}^{(l-1)}| \geq \epsilon$ **or** $l \leq L$ **then**

6 : Update $x_{i,n}^{(l)}$ with $p_{i,n}^{(l)}$ using (2.13)

7 : **else**

8 : $p_{i,n}^{(l)} = p_{i,n}^{(l-1)}$
$x_{i,n}^{(l)} = x_{i,n}^{(l-1)}$

9 : **end if**

10 : $l = l + 1$

11 : **end for**

12 : **return** $p_{i,n}^{(l)}, x_{i,n}^{(l)}$

Algorithm 2.2 : Optimization Algorithm to Solve (2.12)

1: Use (2.3) to calculate $w_i,\ \forall\ i \in \mathcal{M}_s$

2: Obtain resources allocated to M2M traffic $\mathcal{N}_{\text{m2m}}$ from Algorithm 2.1

3: Obtain $p_{i,n}$ from Algorithm 2.1 such that
$p_{i,n} = p_{i,n}^{(l)},\ \forall\ i \in \mathcal{M}_s, n \in N_{\text{m2m}}$

4: Calculate $R_{i,n}$ where
$R_{i,n} = \log_2(1 + \frac{p_{i,n}|h_{i,n}^2|}{\sigma^2})\ \forall\ i \in \mathcal{M}_s, n \in N_{\text{m2m}}$

5: Solve (2.12) using CVX

6: Return $x_{i,n}^{\text{m2m}}$

Algorithm 2.3 : Proposed Two-Phase Resource Allocation Algorithm

1: *Initialize*: $\mathcal{N} = \{1, \cdots, n, \cdots, N\}, \mathcal{I} = \{1, \cdots, i, \cdots, I\}$,
$\mathcal{K} = \{T_0, \cdots, T_k, \cdots, T_K\}$, and $\mathcal{H}, \mathcal{M}_s, \mathcal{M}_t \subset \mathcal{I}$

First Phase: H2H/M2M Joint Power-Resource Allocation

2: Find the optimal power allocation $p_{i,n}$ and optimal resource allocation $x_{i,n}$, $\forall\, i \in \mathcal{H} \cup \mathcal{M}_s$ using Algorithm 2.1

Second Phase: M2M Resource Allocation

3: Obtain amount of resources available for M2M traffic N_{m2m} from Algorithm 2.1

4: Find the optimal resource allocation $x_{i,n}^{\mathrm{m2m}}$, $\forall\, i \in \mathcal{M}_s \cup \mathcal{M}_t$ using Algorithm 2.2

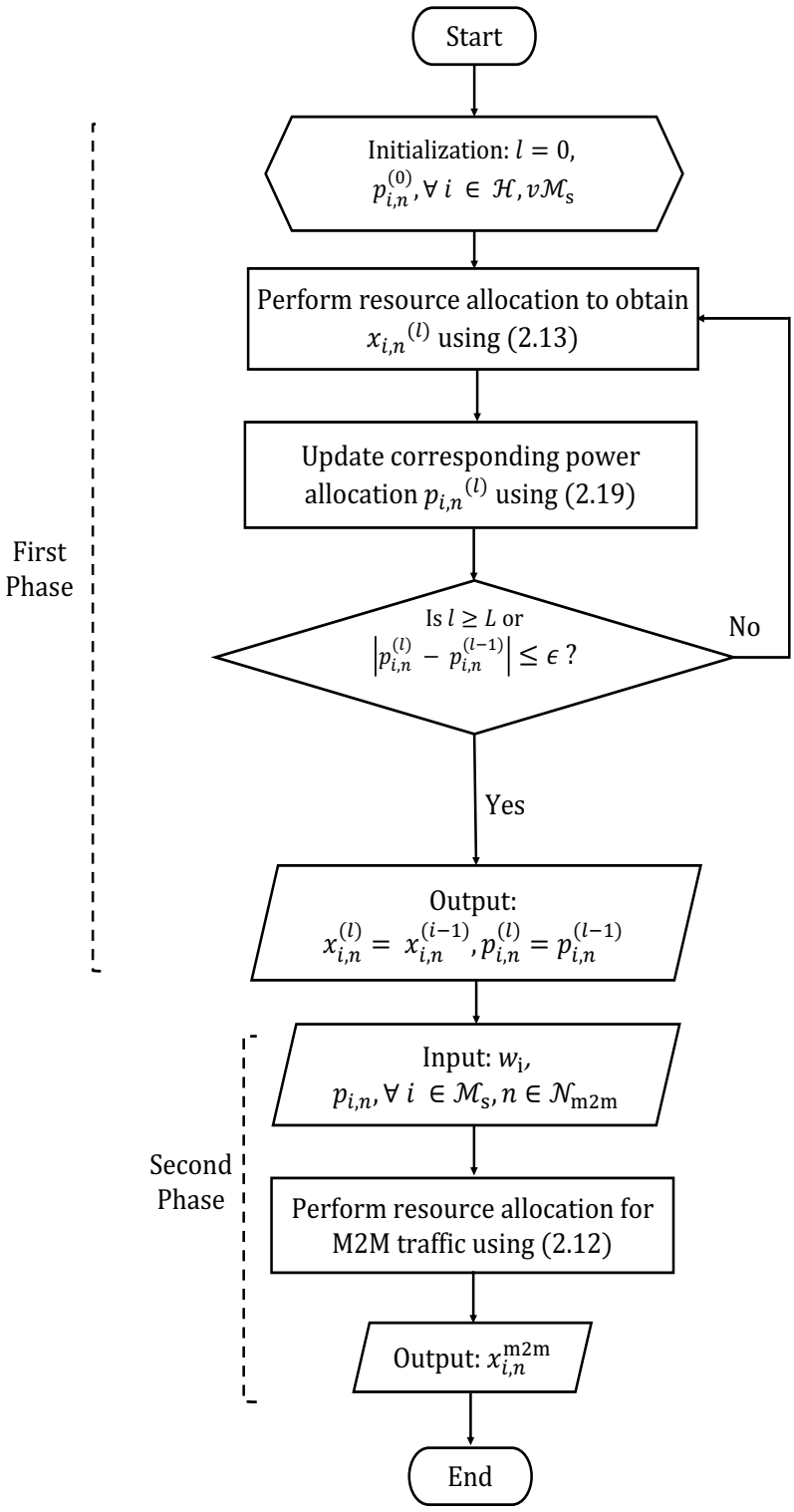

Figure 2.4: Flowchart of the proposed two-phase resource allocation algorithm.

as [120]

$$|h_{i,n}|^2 = G\,\Gamma_{i,n} d_i^{-\alpha}, \tag{2.20}$$

where G is the path loss constant, $\Gamma_{i,n}$ is the shadowing gain which follows a log-normal distribution with zero mean and standard deviation σ_s, d_i is the distance between the BS and ith H2H UE/MTCD, and α is the path loss exponent. Hence, the achievable data rate of the ith H2H UE/MTCD over the nth RB is

$$\begin{aligned} R_{i,n} &= B_{\mathrm{RB}} \log_2(1+\gamma_{i,n}) \\ &= B_{\mathrm{RB}} \log_2(1+\frac{p_{i,n}|h_{i,n}|^2}{\sigma^2}). \end{aligned} \tag{2.21}$$

The total achievable data rate of the users in each subset is

$$\begin{aligned} R^{\mathrm{h,tot}} &= \sum_{i\in\mathcal{H}} R_i \\ &= \sum_{i\in\mathcal{H}} \sum_{n\in\mathcal{N}} x_{i,n}\, B_{\mathrm{RB}} \log_2(1+\frac{p_{i,n}|h_{i,n}|^2}{\sigma^2}), \end{aligned} \tag{2.22}$$

$$\begin{aligned} R^{\mathrm{ms,tot}} &= \sum_{i\in\mathcal{M}_s} R_i \\ &= \sum_{i\in\mathcal{M}_s} \sum_{n\in\mathcal{N}_{\mathrm{m2m}}} x_{i,n}^{\mathrm{m2m}}\, B_{\mathrm{RB}} \log_2(1+\frac{p_{i,n}|h_{i,n}|^2}{\sigma^2}), \end{aligned} \tag{2.23}$$

$$\begin{aligned} R^{\mathrm{mt,tot}} &= \sum_{i\in\mathcal{M}_t} R_i \\ &= \sum_{i\in\mathcal{M}_t} \sum_{n\in\mathcal{N}_{\mathrm{m2m}}} x_{i,n}^{\mathrm{m2m}}\, B_{\mathrm{RB}} \log_2(1+\frac{p_{i,n}|h_{i,n}|^2}{\sigma^2}), \end{aligned} \tag{2.24}$$

After optimal resource allocation, constraint (2.9c) is satisfied, so then

$$R^{\mathrm{h,tot}} \geq R^{\mathrm{h,min}} = \sum_{i\in\mathcal{H}} R_i^{\mathrm{h,min}}. \tag{2.25}$$

For delay-sensitive MTCDs

$$\begin{aligned} R^{\mathrm{ms,tot}} &\geq R^{\mathrm{ms,min}} \\ &= \sum_{i \in \mathcal{M}_s} R_i^{\mathrm{ms,min}}(d_i^{\max}, \varepsilon_i^{\max}), \end{aligned} \tag{2.26}$$

and for delay-tolerant MTCDs, constraint (2.12d) is satisfied after optimal resource allocation with

$$R^{\mathrm{mt,tot}} \geq \rho R^{\mathrm{mt,min}}. \tag{2.27}$$

2.7 Simulation Results

In this section, the performance of the proposed two-phase resource allocation algorithm is evaluated via simulation and compared with the theoretical analysis. The performance of H2H UE is evaluated in terms of the total achievable data rate, while that of delay-sensitive MTCDs is evaluated in terms of the PDBV as well as the total achievable data rate. Finally, the total achievable data rate of delay-tolerant MTCDs is investigated as a measure of fairness in the proposed algorithm.

A single cell in an LTE network with uniformly distributed H2H UE and MTCDs is considered. The available radio resources are divided in the time and frequency domains into RBs where each RB has a time duration of 0.5 ms and a bandwidth of 180 kHz [26]. The channel experiences path loss and shadowing in addition to AWGN. The sources of arriving H2H traffic are video and Voice-over-IP (VoIP) applications [10], while delay-sensitive M2M traffic is event-based with Poisson arrivals [31], and delay-tolerant M2M traffic is uniformly distributed with arrival rates as indicated in Table 2.2. The other simulation parameters are given in this table.

The total achievable data rate of H2H traffic versus the number of H2H UE, I_{h}, is depicted in Figure 2.5 for different numbers of MTCDs (both delay-sensitive and delay-tolerant). This shows that the QoS requirements of H2H UE are satisfied with a total achievable data rate exceeding the minimum data rate threshold, $R^{\mathrm{h,min}}$, in all cases. In addition, the simulation results agree with the theoretical results calculated from (2.22). These results confirm that the proposed algorithm gives priority to H2H traffic where the total achievable data rate is always greater than $R^{\mathrm{h,min}}$ as the number of MTCDs increases from 20 to 65.

Figure 2.6 presents the percentage of RBs allocated to H2H traffic versus the

Table 2.2: Simulation Parameters.

Parameter	**Value**
Cell radius (r)	500 m [26], 2 km
Distribution of H2H UE/MTCDs	Uniform [38]
Bandwidth	10 MHz
No. of RBs (N)	$50, 100$
No. of H2H UE (I_{h})	$2 - 65$
No. of delay-sensitive MTCDs (I_{ms})	$2 - 50$
No. of delay-tolerant MTCDs (I_{mt})	$10 - 60$
Path loss constant (G)	10^{-2} [120]
Path loss exponent (α)	3 [120]
Shadowing standard deviation (σ_s)	8 dB [26]
Maximum transmit power ($P^{\mathrm{RB,\ max}}$)	24 dBm [120]
Noise power (σ^2)	-100 dBm [26]
Maximum delay-bound ($d_i^{\max}$)	0.2 ms [38]
Maximum delay-violation threshold ($\varepsilon_i^{\max}$)	10% [38]
H2H minimum data rate requirements ($R_i^{\mathrm{h,min}}$), $\forall\, i \in \mathcal{H}$	$64, 128$ kbps [38]
M2M minimum data rate requirements ($R_i^{\mathrm{mt,min}}$), $\forall\, i \in \mathcal{M}_t$	$10, 20, 30, 40$ kbps [31]
w_i, $\forall\, i \in \mathcal{M}_s$	0.8
ρ	$0.2, 0.4, 0.6, 0.8$

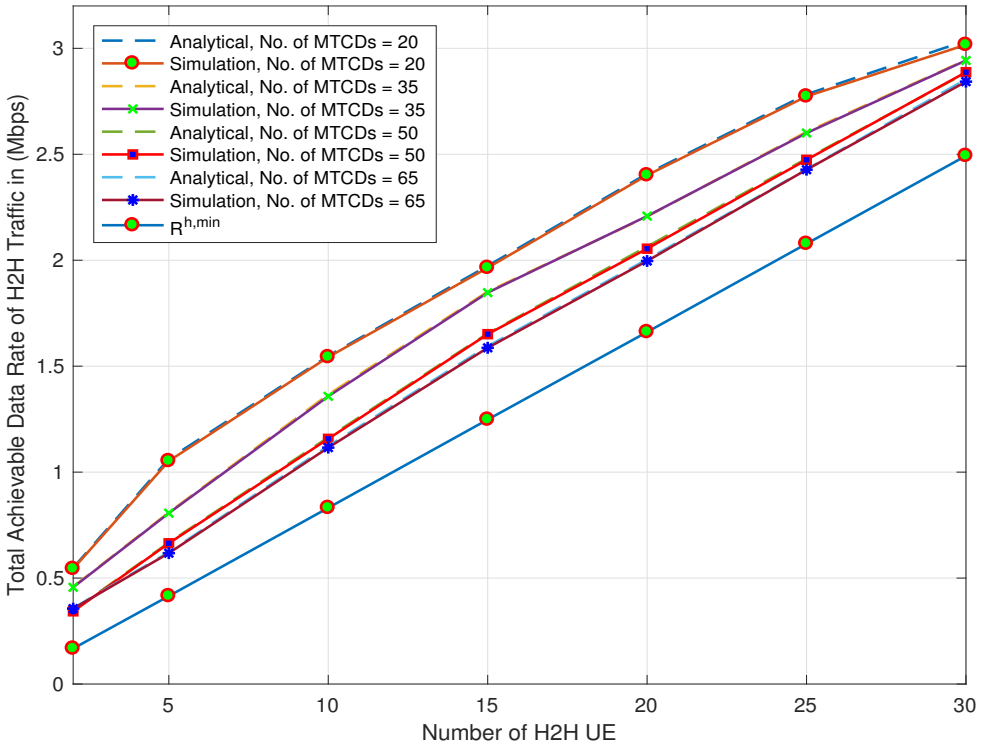

Figure 2.5: Total achievable data rate of H2H traffic versus the number of H2H UE, I_{h}, for different numbers of MTCDs, I_{m2m}.

number of H2H UE, I_{h}. This shows that the percentage of resources allocated to H2H traffic increases with the number of H2H UE. This demonstrates the ability of the proposed algorithm to accommodate increased H2H traffic through dynamic resource allocation and prioritization.

The performance of delay-sensitive MTCDs is given in Figure 2.7 in terms of the PDBV for different numbers of H2H UE, I_{h}. These results show that the PDBV increases with the number of delay-sensitive MTCDs, I_{ms}. However, it is always less than the maximum delay-violation threshold, $\varepsilon_i^{\max}$, even when I_{h} is equal to 15. This indicates that the proposed algorithm can satisfy the delay requirements of these MTCDs.

The effect of delay-tolerant MTCDs on the delay requirements of delay-sensitive MTCDs is illustrated in Figure 2.8 which gives the PDBV for different values of the fairness factor, ρ. These results show that the delay requirements are met in all cases as the PDBV is less than the maximum delay-violation threshold, $\varepsilon_i^{\max}$. Thus, priority

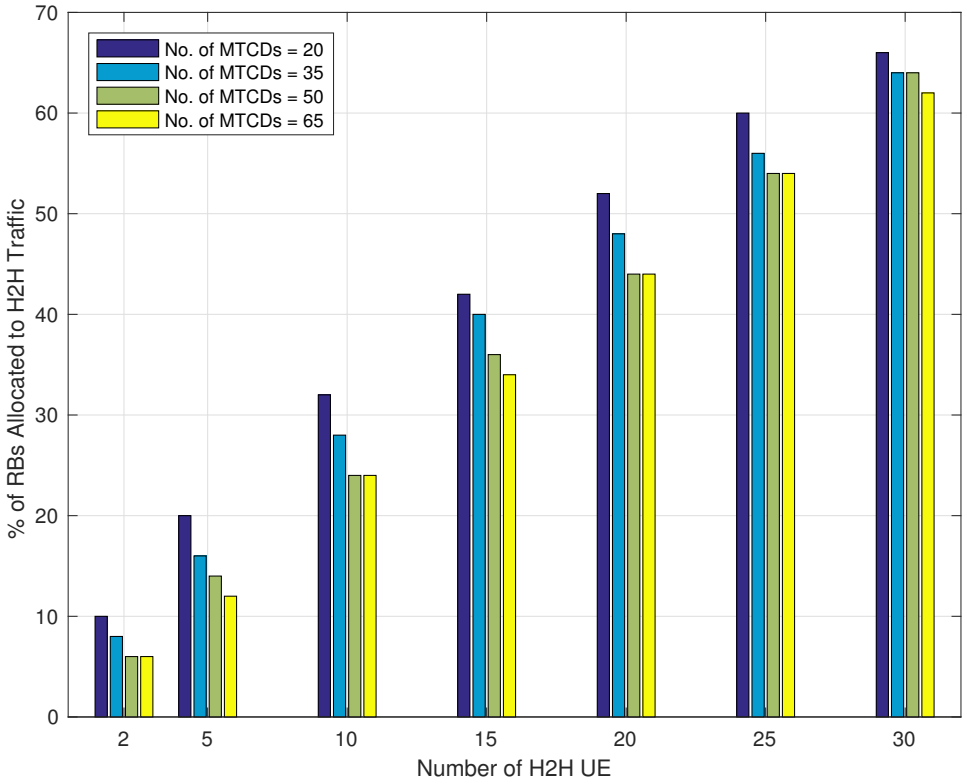

Figure 2.6: Percentage of resources allocated to H2H traffic versus the number of H2H UE, I_{h}, for different numbers of MTCDs, I_{m2m}.

is given to delay-sensitive MTCDs over the delay-tolerant MTCDs.

Figure 2.9 presents the total achievable data rate of delay-sensitive M2M traffic versus the number of delay-sensitive MTCDs, I_{ms}, for different numbers of H2H UE, I_{h}. This shows that for a given I_{h}, the total achievable data rate increases with the number of delay-sensitive MTCDs. On the other hand, as I_{h} increases, more resources are allocated to H2H traffic resulting in a reduction in the total achievable data rate of delay-sensitive MTCDs. However, this is greater than the minimum data rate threshold, $R^{\mathrm{ms,min}}$, which indicates that the QoS requirements of this MTCD group are met. Moreover, there is close agreement between the simulation and theoretical results from (2.23).

Figure 2.10 gives the total achievable data rate versus the number of delay-sensitive MTCDs, I_{ms}, for different values of the fairness factor, ρ. This shows that as ρ increases from 0.2 to 0.8, the total achievable data rate decreases. The larger ρ is, the higher the minimum data rate threshold for delay-tolerant MTCDs, $\rho R^{\mathrm{mt,min}}$.

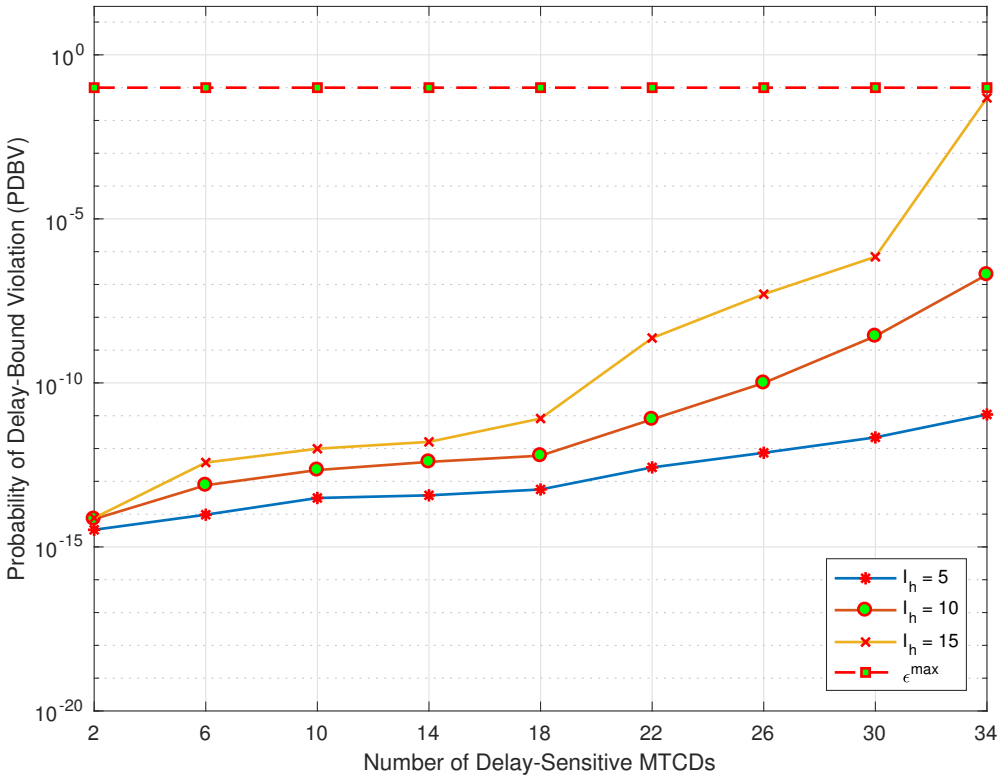

Figure 2.7: PDBV versus the number of delay-sensitive MTCDs, I_{ms}, for different numbers of H2H UE, I_{h}.

Consequently, fewer resources are allocated to delay-sensitive MTCDs in order to satisfy the fairness requirements of delay-tolerant MTCDs. Nevertheless, the total achievable data rate of delay-sensitive MTCDs is greater than $R^{\mathrm{ms,min}}$, so the QoS requirements of this MTCD group are satisfied even at high values of ρ.

Figures 2.9 and 2.10 show that the total achievable data rate of the delay-sensitive MTCDs approaches a constant as the number of these MTCDs increases. This is because there is a maximum amount of available resources to be allocated to delay-sensitive MTCDs after the minimum data rate requirements of H2H UE, $R^{\mathrm{h,min}}$, and the minimum data rate requirements of delay-tolerant MTCDs, $\rho R^{\mathrm{mt,min}}$, are satisfied. Therefore, increasing the number of delay-sensitive MTCDs beyond a certain number will not increase their total achievable data rate.

The performance of delay-tolerant MTCDs is presented in Figure 2.11. This gives the total achievable data rate of delay-tolerant M2M traffic versus the number of delay-tolerant MTCDs, I_{mt}, for different numbers of delay-sensitive MTCDs,

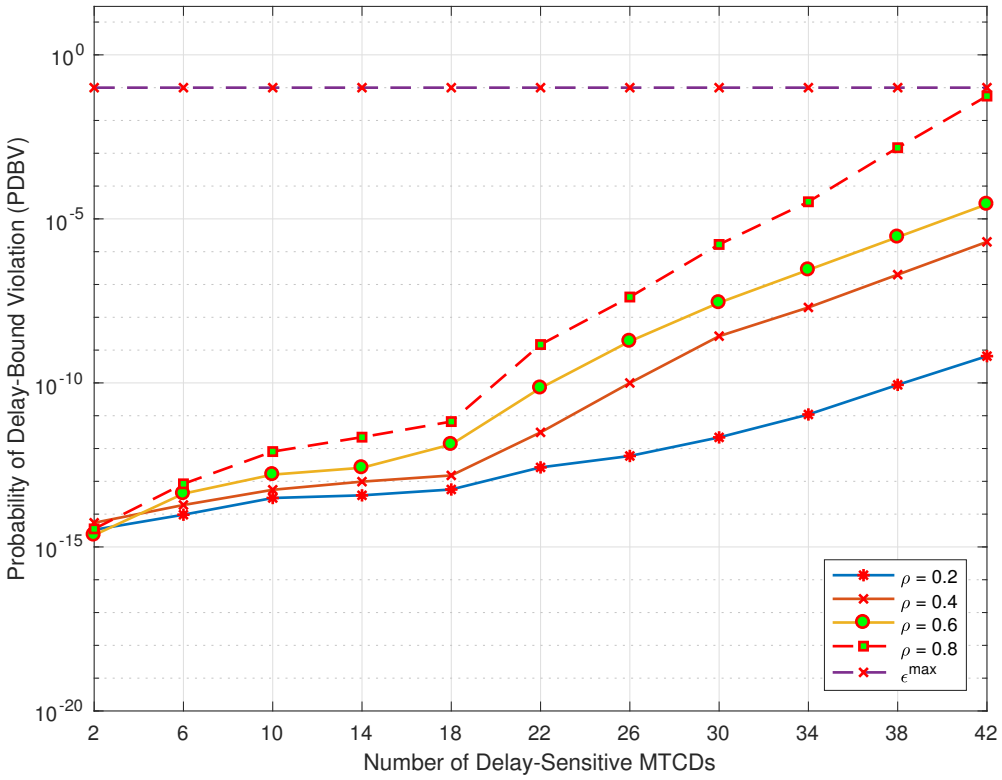

Figure 2.8: PDBV versus the number of delay-sensitive MTCDs, I_{ms}, for different values of the fairness factor, ρ.

I_{ms}. These results show that the total achievable data rate of this MTCD group decreases as the number of delay-sensitive MTCDs increases from 5 to 15. This is due to the higher delay-sensitive traffic load associated with the increased number of delay-sensitive MTCDs. Thus, fewer resources are allocated to delay-tolerant MTCDs. However, the proposed algorithm provides an acceptable level of fairness to this MTCD group, as the total achievable data rate when the number of delay-sensitive MTCDs is largest (15) meets the minimum data rate requirements, $\rho R^{\mathrm{mt,min}}$.

The effect of the fairness factor, ρ, on the performance of delay-tolerant MTCDs is shown in Figure 2.12. This indicates that larger values of ρ result in a higher total achievable data rate since increasing ρ corresponds to increasing the minimum data rate requirements, $\rho R^{\mathrm{mt,min}}$, as given in (2.27).

To illustrate the scalability of the proposed algorithm, the performance is evaluated for a larger cell with radius $r = 2$ km and $N = 100$ RBs. The other simulation parameters are as in Table II. Figure 2.13 presents the total achievable data rate of

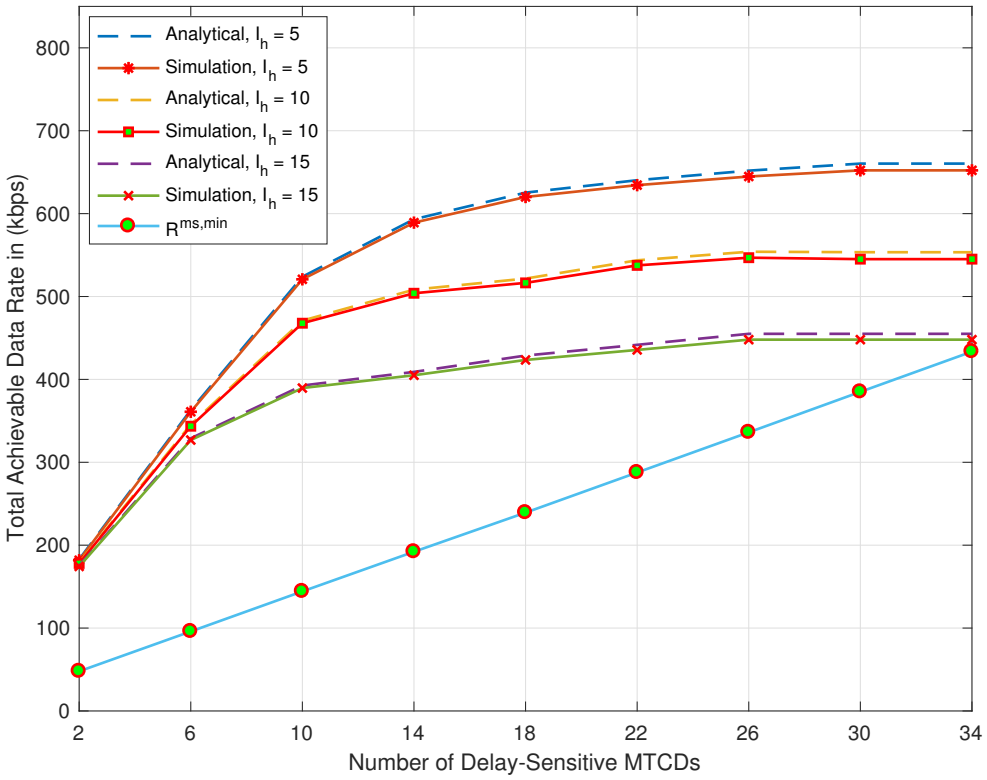

Figure 2.9: Total achievable data rate of delay-sensitive M2M traffic versus the number of delay-sensitive MTCDs, I_{ms}, for different numbers of H2H UE, I_{h}.

H2H traffic versus the number of H2H UE, I_{h}, for different numbers of MTCDs. This shows that the QoS requirements of the H2H UE are satisfied with a total achievable data rate exceeding the minimum data rate threshold, $R^{\mathrm{h,min}}$, in all cases. Comparing these results with those in Figure 2.5 for a cell radius $r = 500$ m shows that with a larger cell, a higher total achievable data rate for H2H traffic is achieved and more H2H UE (up to 65) are allocated resources as the number of available RBs N is larger. In both cases, the total achievable data rate increases with the number of H2H UE and it is always greater than the minimum data rate threshold, $R^{\mathrm{h,min}}$. Hence, the QoS requirements of H2H traffic are satisfied.

Figure 2.14 shows the total achievable data rate of delay-sensitive M2M traffic versus the number of delay-sensitive MTCDs, I_{ms}, for different numbers of H2H UE, I_{h}. For a given I_{h}, the total achievable data rate increases with the number of delay-sensitive MTCDs. On the other hand, as I_{h} increases, the total achievable data rate decreases. However, this is greater than the minimum data rate threshold, $R^{\mathrm{ms,min}}$.

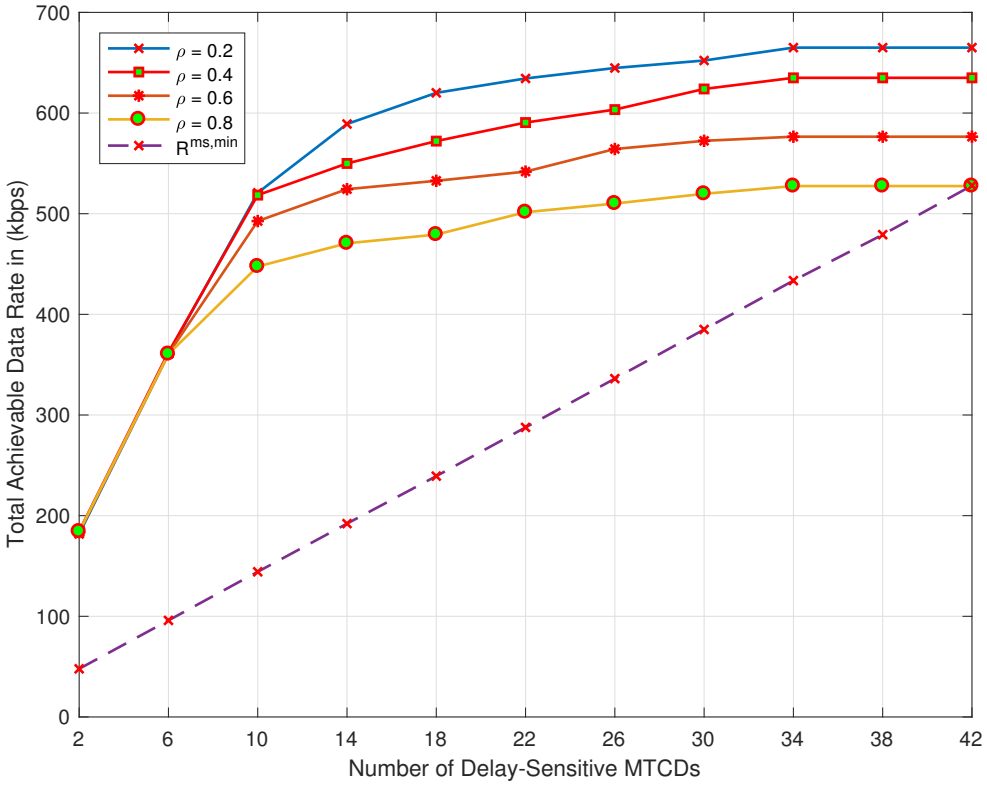

Figure 2.10: Total achievable data rate of delay-sensitive M2M traffic versus the number of delay-sensitive MTCDs, I_{ms}, for different values of the fairness factor, ρ.

Hence, the proposed algorithm is capable of allocating resources to meet the QoS requirements of this MTCD group. Comparing these results with those in Figure 2.9 for a cell radius of $r = 500$ m shows that with a larger cell, more delay-sensitive MTCDs (up to 50) are allocated resources resulting in a higher total achievable data rate. In both cases, the total achievable data rate increases with the number of delay-sensitive MTCDs and it is always greater than the minimum data rate threshold, $R^{\text{ms,min}}$. Thus, the proposed algorithm is scalable as it is able to allocate resources effectively in larger cells.

2.8 Conclusion

In this chapter, the problem of resource allocation in cellular networks with H2H/M2M coexistence and both delay-sensitive and delay tolerant MTCDs was considered. A

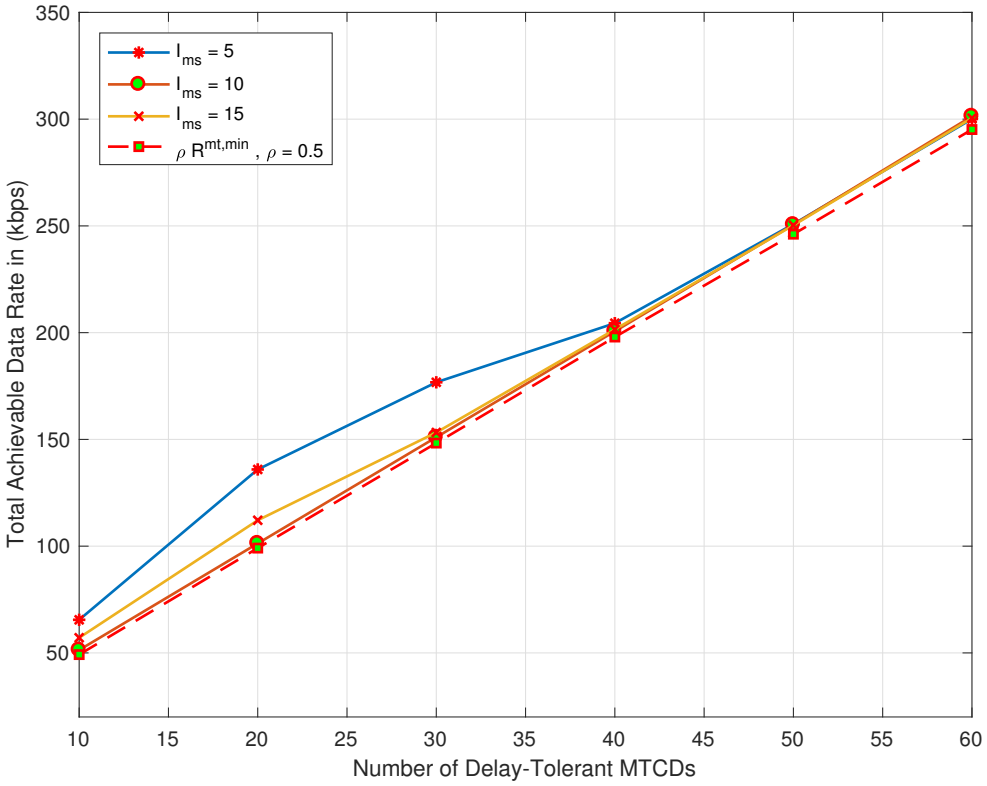

Figure 2.11: Total achievable data rate of delay-tolerant M2M traffic versus the number of delay-tolerant MTCDs, I_{mt}, for different numbers of delay-sensitive MTCDs, I_{ms}.

two-phase resource allocation algorithm was proposed to meet the diverse QoS requirements of H2H and M2M traffic. For the first phase, a joint power-resource allocation problem was formulated with the objective of maximizing the total achievable data rate of H2H UE while considering the delay constraints of delay-sensitive MTCDs. In the second phase the total achievable data rate of delay-sensitive MTCDs is maximized while taking into account fairness for delay-tolerant MTCDs. Simulation results were presented which indicate that the proposed algorithm can control the impact of M2M traffic on H2H traffic while achieving the QoS requirements of both H2H UE and MTCDs. Moreover, fairness is ensured for delay-tolerant MTCDs.

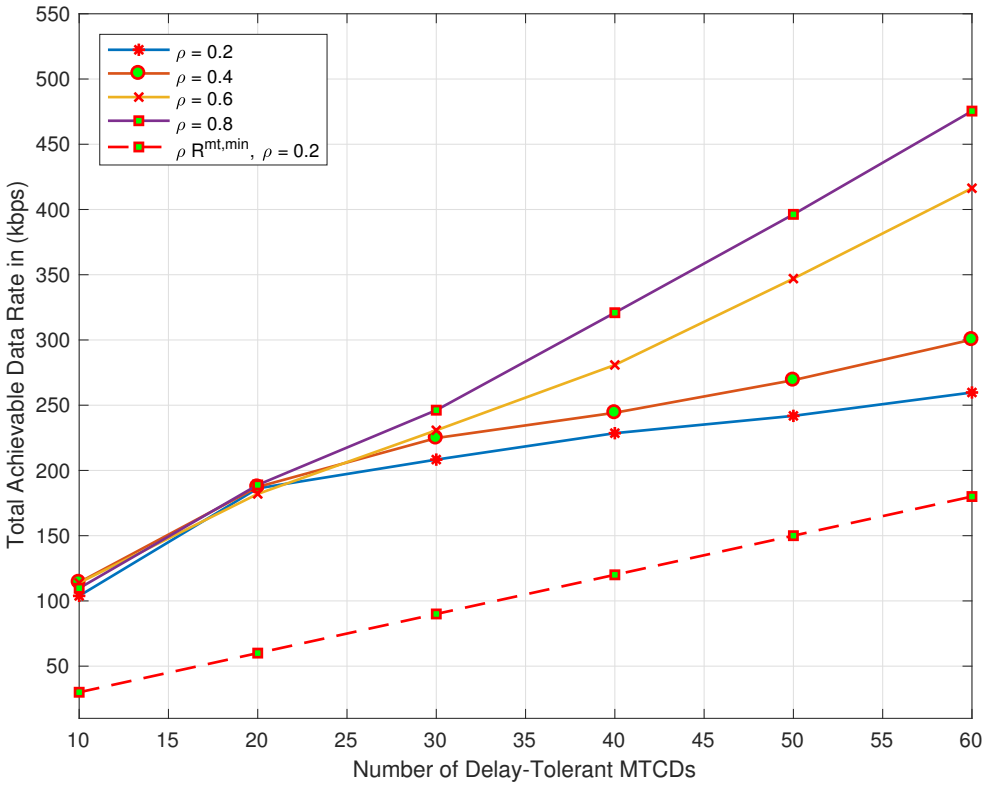

Figure 2.12: Total achievable data rate of delay-tolerant M2M traffic versus the number of delay-tolerant MTCDs, I_{mt}, for different values of the fairness factor, ρ.

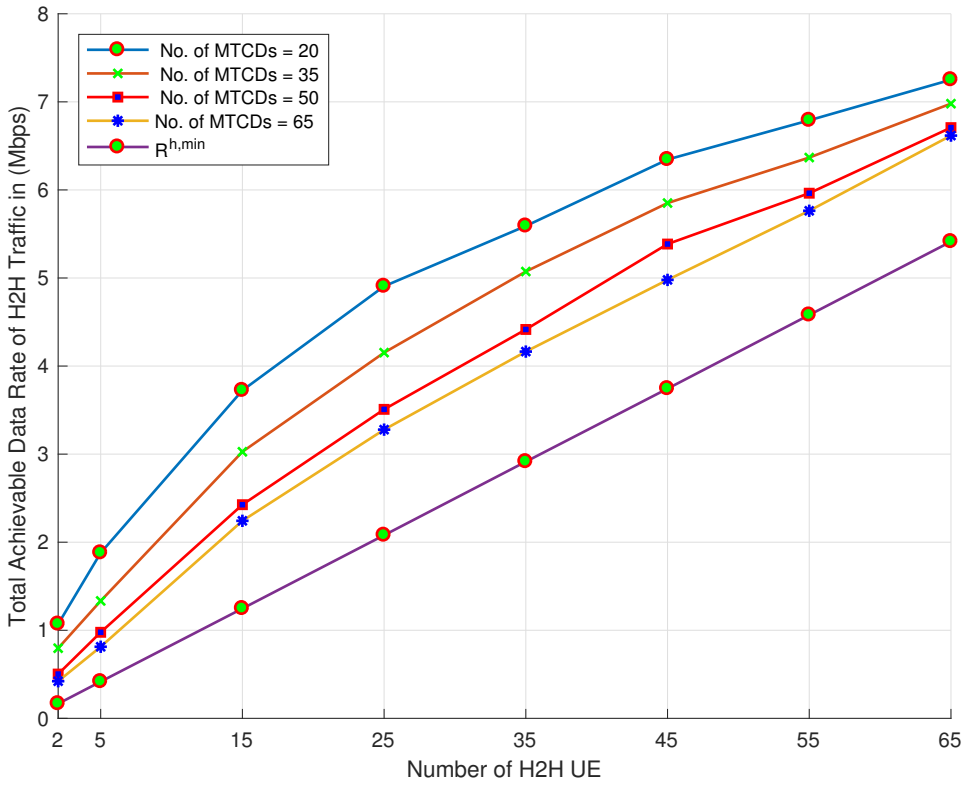

Figure 2.13: Total achievable data rate of H2H traffic versus the number of H2H UE, I_{h}, for $r = 2$ km and different numbers of MTCDs, I_{m2m}.

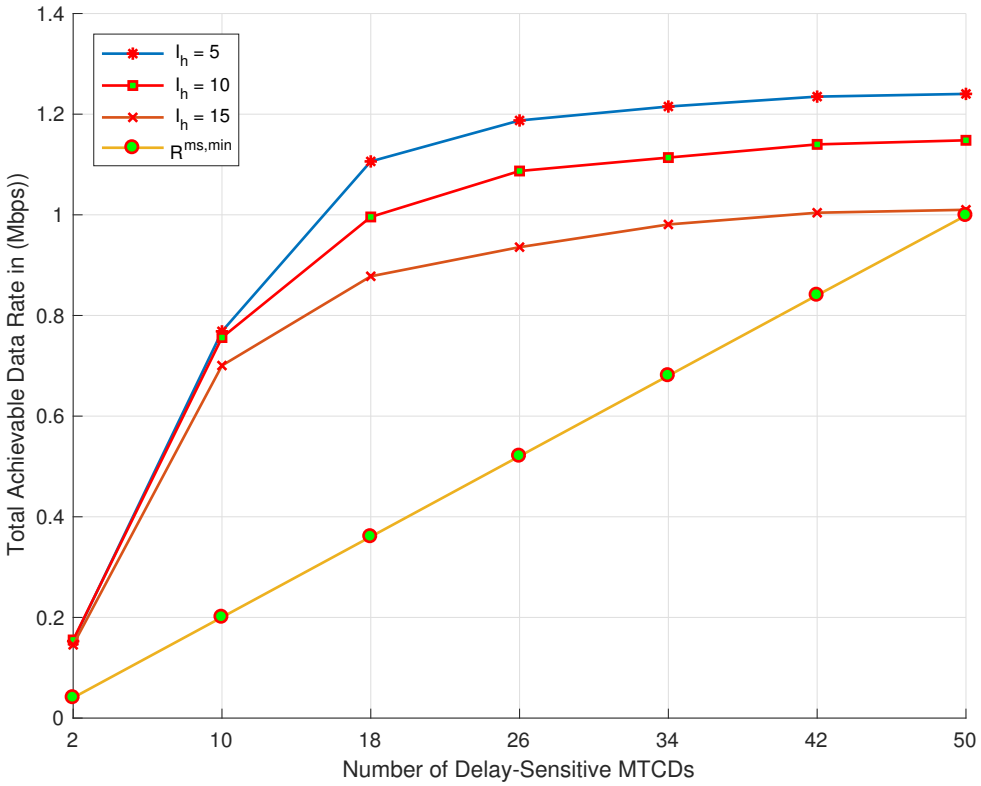

Figure 2.14: Total achievable data rate of delay-sensitive M2M traffic versus the number of delay-sensitive MTCDs, I_{ms}, for $r = 2$ km and different numbers of H2H UE, I_{h}.

Chapter 3

Joint Resource and Power Allocation for Clustered Cognitive M2M Communications Underlaying Cellular Networks

M2M communications play a prominent role in the areas of cloud computing, smart grid (demand/response management, grid control, smart metering), vehicular communications (vehicle-to-vehicle and vehicle-to-infrastructure communications), industrial monitoring, and e-health [3]. Non-cellular wireless technologies such as WiFi, Bluetooth, and Zigbee have been considered for enabling M2M communications, but they are inefficient and insufficient to support the growing demands of M2M applications [87]. On the other hand, cellular networks provide extended wireless connectivity, more flexibility for radio resource management, and higher capacity to support M2M deployment.

The characteristics of M2M communications are distinct from those of traditional cellular communications which cellular networks were originally designed for [43]. M2M communications are characterized by infrequent transmissions of small amounts of data, delay tolerance, and high traffic demands on the uplink channel. Moreover, the number of Machine-Type Communication Devices (MTCDs) is growing exponentially and is estimated to reach 29.3 billion by 2023 [86]. These characteristics pose significant challenges to cellular networks in terms of interference, congestion in the Random Access Channel (RACH), and spectrum scarcity [5], [44], [89].

To overcome the problem of spectrum scarcity while exploiting the benefits of M2M communications, Cognitive Radio (CR) can be integrated with M2M communications to provide Cognitive M2M (CM2M) communications [77], [92], [46]. In CM2M communications, MTCDs use the spectrum resources of Cellular User Equipment (CUE). This is achieved by sharing the spectrum resources of the CUE with the MTCDs where the CUE are considered to be Primary Users (PUs) and the MTCDs are Secondary Users (SUs).

Spectrum sharing (SS) models in CM2M communications include interweave, overlay, and underlay. In interweave SS, also known as opportunistic SS, SUs sense the spectrum of PUs and opportunistically access the unused spectrum resources, also called spectrum holes, over time and/or space [39]. In overlay SS, PUs and SUs simultaneously communicate as long as PU transmissions are not affected. To facilitate this, SUs can identify PU signals as they have prior knowledge regarding transmissions. Hence, a SU can transmit at any power level and employ coding to protect against PU interference [40]. In underlay SS, SUs share the PU spectrum while transmitting below an interference threshold [79], [94]. In M2M environments, underlay SS is the most suitable due to the high spectral and power efficiencies with low transmission delay [41].

Clustered M2M communications can be used to coordinate communications between the BS and MTCDs and hence alleviate the issue of RACH congestion. In this case, MTCDs are grouped into clusters based on their spatial locations or Quality-of-Service (QoS) requirements, and communicate with the Base Station (BS) via a Machine-Type Communication Gateway (MTCG). An MTCG collects data from the MTCDs within its cluster and forwards it to the BS.

3.1 Related Work

The problem of resource allocation for M2M communications has been considered in the literature. Some approaches focus on resource allocation for clustered M2M communications while others consider CM2M. However, little attention has been given to resource allocation for clustered CM2M.

A clustered M2M architecture was investigated in [50]. Congestion avoidance due to M2M traffic was proposed to reduce uplink transmissions in a Long Term Evolution (LTE) cellular network. This approach assumes that MTCDs can communicate with each other via a local network, but they access the BS through a group leader. In

[51], clustered M2M communications was studied where M2M traffic is directed to the BS under the control of a cluster head while CUE directly access the BS. This system was analyzed with respect to the RACH considering the relationship between the access success probability and the number of random access preambles allocated to the machines within a cluster.

Joint MTCD selection and power allocation was proposed in [52] for a buffer-aided CM2M network. The CM2M network decides on whether to be silent or select either a relay or an MCTD for data transmission. A buffer is used at the relay to reduce data loss. The goal is to maximize the sum-rate of the MTCDs while limiting interference to CUE. In [107], energy-efficient selective compressive sensing and resource allocation for CM2M networks was proposed where M2M traffic is classified into four service classes based on data rate and delay requirements. The MTCDs employ selective compressive sensing to detect spectrum holes where each MTCD senses a subset of the PU spectrum. The BS then allocates the available spectrum to the MTCDs in order to satisfy minimum data rate and latency requirements.

Energy-efficient energy harvesting-based CM2M communications was considered in [110]. In order to maximize the energy efficiency of the M2M transmitters (M2M-TXs), joint optimization of channel selection, peer discovery, power control, and time allocation is performed via a two-stage 3D matching algorithm. In the first stage, M2M-TXs, M2M receivers (M2M-RXs), and Resource Blocks (RBs) are temporally matched together, and then a joint power control and time allocation problem is solved. In the second stage, the joint channel selection and peer discovery problem is solved using a pricing-based matching algorithm based on the established preference lists. In [80], CM2M communications underlaying a cellular network was considered. In this system, M2M devices reuse the licensed spectrum of cellular users in an opportunistic and fair manner. Two power allocation strategies based on proportional fairness and max-min fairness were proposed to achieve global fairness objectives while satisfying an interference temperature constraint at the BS.

In [81], a CM2M paradigm which enables a massive number of Machine Type Devices (MTDs) to either opportunistically use licensed spectrum or exploit under-utilized unlicensed spectrum was considered. A channel selection problem with both licensed and unlicensed spectrum was formulated as an adversarial multi-armed bandit (MAB) problem. An exponential-weight algorithm for exploration and exploitation (EXP3) and Lyapunov optimization were combined to develop a context-aware channel selection algorithm named C2-EXP3.

A Quality-of-Experience (QoE) oriented rate control and resource allocation scheme for CM2M communications in OFDM networks was introduced in [54]. The resulting optimization problem was converted into two subproblems for admission control and resource allocation. A Mean Opinion Score (MOS) model is used to measure the QoE and stochastic optimization is employed to maximize the time-averaged MOS of all M2M pairs subject to network stability, admission rate control, and delay constraints. In [55], a CM2M network was considered where MTCDs generate packets and transfer them to a concentrator. The concentrators are considered as SUs in a CR network and they opportunistically access PU spectrum. A RACH mechanism based on queuing theory is used to analyze the performance of the network. An iterative approach is employed to reduce the complexity and determine the throughput and delay.

A cluster-based CM2M architecture was considered from both engineering and business perspectives in [56] to solve the RACH problem of M2M communications over cellular networks. In this architecture, MTCDs access the BS either directly or via a cluster head using a CR approach. In [57], clustered CM2M communications with joint CUE and MTCD selection was proposed. The goal is to reduce the outage probability of M2M communications and two ordering policies for the joint selection were proposed. The first is based on the MTCD locations while the second is based on the instantaneous channel gains.

In this chapter, clustered Cognitive Machine-to-Machine (CM2M) communications underlaying cellular networks is considered. In this system, Machine-Type Communication Devices (MTCDs) are grouped in clusters based on their spatial location and communicate with the Base Station (BS) via a Machine-Type Communication Gateway (MTCG). Underlay cognitive radio is employed so that the MTCDs within each cluster share the spectrum of the neighbouring Cellular User equipment (CUE). A joint resource-power allocation problem is formulated and solved using a two-phase resource and power allocation algorithm. The goal is to maximize the total achievable data rate of the CUE and clustered MTCDs while adhering to minimum data rate, power, and interference constraints. In the first phase, adaptive resource block allocation (ARBA) is employed to allocate RBs to MTCDs considering the constraints. In the second phase, iterative power allocation (IPA) is used for power allocation to the MTCDs over the set of RBs allocated in the first phase.

3.2 Contributions of the Chapter

The contributions of this chapter are as follows.

1. To overcome the problem of congestion, a clustered M2M model is proposed based on the spatial locations of the MTCDs. In this model, MTCDs connect with the closest MTCG to form a cluster where communications with the BS are managed by the MTCG. This reduces the number of direct accesses to the BS in order to mitigate congestion.

2. To address the problem of spectrum scarcity and improve spectrum utilization, CR is integrated into the clustered M2M communications system. Specifically, underlay CR is employed so that the MTCDs within each cluster share the spectrum of neighbouring CUE.

3. A two-phase resource and power allocation scheme for the clustered CM2M system is presented to maximize the uplink sum-rate of the neighbouring CUE and clustered MTCDs while considering interference, power, and minimum data rate constraints. To date, the problem of resource and power allocation in a clustered CM2M environment considering these constraints has not been addressed in the literature.

4. To protect neighbouring CUE from interference induced by the clustered MTCDs and satisfy the MTCD QoS requirements in terms of delay and minimum data rate, an Adaptive Resource Block Allocation (ARBA) algorithm is proposed as the first phase. This algorithm allocates available Resource Blocks (RBs) to the MTCDs within each cluster based on an adaptive metric that considers minimum data rate, power, and interference requirements.

5. An Iterative Power Allocation (IPA) algorithm is proposed as the second phase to allocate the MTCD transmit power over the allocated RBs. This algorithm maximizes the uplink sum-rate of the neighbouring CUE and clustered MTCDs while keeping the interference induced to the CUE below an interference threshold and satisfying the MTCD power and minimum data rate constraints.

6. To provide scalability and reduce signalling overhead at the BS, the MTCG in each cluster performs spectrum sensing to exploit spectrum opportunities. Moreover, resource allocation for the MTCDs within each cluster is managed by the corresponding MTCG.

7. Simulation results are presented which show that the proposed scheme significantly improves the sum-rate of the network compared to other schemes while satisfying the constraints.

3.3 System Model

3.3.1 Network Model

Figure 3.1 shows a single cell in an LTE-A cellular network with one BS, multiple CUE, multiple MTCGs, and multiple MTCDs. Denote the set of CUE by $\mathcal{U} = \{1, 2, \cdots, U\}$, the set of MTCGs by $\mathcal{G}_c = \{1, 2, \cdots, G_c\}$, and the set of MTCDs by $\mathcal{M} = \{1, 2, \cdots, M\}$. The BS is assumed to be located at the centre of the cell and MTCGs are distributed at fixed locations from the BS. Moreover, CUE and MTCDs are uniformly distributed according to spatial Binomial Point Processes (BPP) denoted by ψ_U and ψ_M, respectively. MTCDs within the cell are grouped into M2M clusters based on their spatial locations and communicate with the BS through an MTCG where each cluster comprises one MTCG and a number of MTCDs located within radius R_c from the MTCG. It is assumed that there are C M2M clusters in the cell with the set of M2M clusters denoted by $\mathcal{C} = \{1, 2, \cdots, C\}$. Moreover, the set of MTCDs within the cth cluster is denoted by $\mathcal{M}_c = \{1, 2, \cdots, M_c\}$.

Underlay cognitive radio scheme is considered where MTCDs within the cth cluster share the spectrum of neighbouring CUE. A neighbouring CUE to the cth cluster is a CUE located within radius R_s from MTCG g_c where R_s is the sensing radius. Let the set of neighbouring CUE to the cth cluster be $\mathcal{U}_c = \{1, 2, \cdots, U_c\}$. The maximum transmit power of the MTCG and MTCDs within the cth cluster are denoted by $P_{g_c}^{\max}$ and $P_{m_c}^{\max}$, respectively, and are chosen such that the interference to neighbouring CUE is kept below an interference threshold. The assumption in an underlay cognitive radio scheme is that the achievable data rate of the neighbouring CUE without the presence of the underlaying MTCDs is larger than the minimum data rate threshold. Therefore, MTCDs can share the spectrum of the CUE as long as the interference introduced to the CUE is kept below an interference threshold [83].

3.3.2 Channel Model

Uplink traffic is considered with Single Carrier Frequency Division Multiple Access (SC-FDMA) as the channel access scheme. Available resources are divided in time and frequency into RBs where each RB has a bandwidth of 180 kHz and a time duration of 0.5 ms [26]. Denote the set of available RBs by $\mathcal{N} = \{1, 2, \cdots, N\}$. It is assumed that an RB is assigned to only one CUE within the cell so there is no co-channel interference among the CUE. The allocation of RBs to CUE is performed by the BS and is not within the scope of this work. Moreover, each MTCD within a cluster shares the RBs of one CUE only and this is sufficient to fulfill the data rate requirements of the MTCD. Besides, MTCDs within the same cluster cannot share the same RB, implying that there will be no intra-cluster interference. The scenario where multiple MTCDs within a cluster share the same RB is left for future research.

The channel experiences path loss, large scale shadowing that follows a log-normal distribution with zero mean and standard deviation σ_s, and small scale Rayleigh fading with unit mean exponential distribution. The channel gains from the CUE $\rightarrow$ BS, CUE $\rightarrow$ MTCG, MTCG $\rightarrow$ BS, and MTCD[1] $\rightarrow$ MTCG[2] links over the nth RB are denoted by $|h^n_{u,b}|^2$, $|h^n_{u,g_c}|^2$, $|h^n_{g_c,b}|^2$, and $|h^n_{m_c,g_c}|^2$, respectively. It is assumed that the channel state information (CSI) of all links in addition to the CUE interference threshold are known at the BS and transmitted by the BS to the MTCGs at the beginning of each Transmission Time Interval (TTI). Note that the coherence time is assumed to be greater than the TTI. Therefore, these coefficients are constant within a given TTI [58].

3.3.3 Spectrum Sensing and Data Transmission

It is assumed that the MTCG within each cluster is capable of sensing the spectrum of the neighbouring CUE to exploit spatial spectrum opportunities. The MTCG then assigns available RBs to the clustered MTCDs such that the interference induced to neighbouring CUE is below the interference threshold. Note that decreasing this threshold provides better protection to the neighbouring CUE. However, this also reduces the number of spectrum sharing opportunities for the clustered MTCDs. Conversely, increasing this threshold provides more spectrum sharing opportunities but results in more interference to the CUE.

[1]Here MTCD refers to MTCD m_c in cluster c.
[2]Here MTCG refers to the MTCG g_c in cluster c.

The MTCG operates in half-duplex mode where the RBs allocated to MTCD m_c are used for the overall link $m_c \rightarrow g_c \rightarrow$ BS. In the first hop, data is transmitted from $m_c \rightarrow g_c$ while in the second hop data is forwarded from $g_c \rightarrow$ BS. Therefore, the achievable data rate of MTCD m_c over the nth RB in the first hop is

$$R_{m_c}^{n(1)} = B_{\text{RB}} \log_2(1 + \gamma_{m_c}^{n(1)}) , \quad \forall m_c \in \mathcal{M}_c, n \in \mathcal{N} \tag{3.1}$$

where B_{RB} is the RB bandwidth and $\gamma_{m_c}^{n(1)}$ is the Signal-to-Interference Plus Noise Ratio (SINR) of MTCD m_c over the nth RB in the first hop defined as

$$\begin{aligned} \gamma_{m_c}^{n(1)} &= \frac{p_{m_c}^n |h_{m_c,g_c}^n|^2 d_{m_c,g_c}^{-\alpha}}{p_u^n \, |h_{u,g_c}^n|^2 d_{u,g_c}^{-\alpha} + \sigma_0^2} \\ &= p_{m_c}^n \Gamma_{m_c,g_c}^n \end{aligned} \tag{3.2}$$

where $p_{m_c}^n$ is the transmit power of MTCD m_c over the nth RB, d_{m_c,g_c} is the distance between MTCD m_c and MTCG g_c, α is the path loss exponent, p_u^n is the transmit power of the uth neighbouring CUE over the nth RB, d_{u,g_c} is the distance between the uth neighbouring CUE and MTCG g_c, σ_0^2 is the Additive White Gaussian Noise (AWGN) power, and Γ_{m_c,g_c}^n is the channel gain-to-noise ratio between MTCD m_c and MTCG g_c over the nth RB.

The achievable data rate of MTCD m_c over the nth RB in the second hop is given by

$$R_{m_c}^{n(2)} = B_{\text{RB}} \log_2(1 + \gamma_{m_c}^{n(2)}) , \quad \forall m_c \in \mathcal{M}_c, n \in \mathcal{N} \tag{3.3}$$

where $\gamma_{m_c}^{n(2)}$ is the SINR in the second hop defined as

$$\begin{aligned} \gamma_{m_c}^{n(2)} &= \frac{p_{g_c}^n |h_{g_c,b}^n|^2 d_{g_c,b}^{-\alpha}}{p_u^n \, |h_{u,b}^n|^2 d_{u,b}^{-\alpha} + \sigma_0^2} \\ &= p_{g_c}^n \Gamma_{g_c,b}^n \end{aligned} \tag{3.4}$$

where $p_{g_c}^n$ is the transmit power of MTCG g_c, $d_{g_c,b}$ is the distance between MTCG g_c and the BS, $d_{u,b}$ is the distance between the uth neighbouring CUE and the BS, and $\Gamma_{g_c,b}^n$ is the channel gain-to-noise ratio between MTCG g_c and the BS over the nth RB. Thus, the overall end-to-end achievable data rate of MTCD m_c over the nth RB is

$$R_{m_c}^n = \frac{1}{2} \min(R_{m_c}^{n(1)}, R_{m_c}^{n(2)}) , \tag{3.5}$$

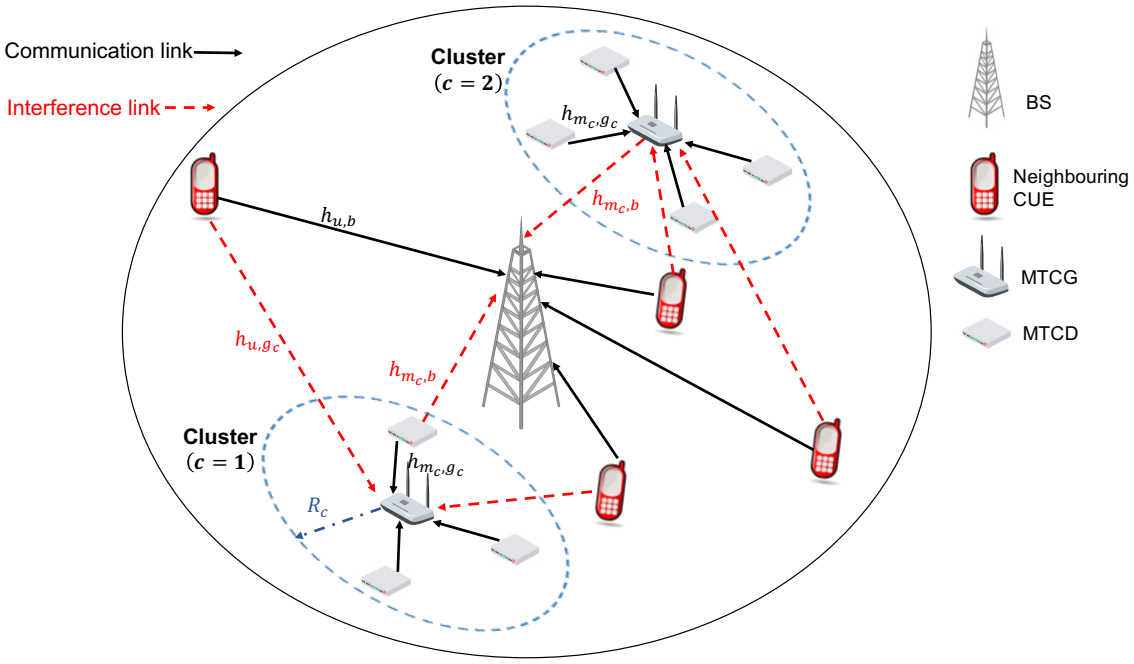

Figure 3.1: The proposed network model.

The maximum achievable data rate is achieved when $R_{m_c}^{n(1)} = R_{m_c}^{n(2)}$ and hence $p_{m_c}^n \Gamma_{m_c,g_c}^n = p_{g_c}^n \Gamma_{g_c,b}^n$ [19]. Replacing $p_{g_c}^n \Gamma_{g_c,b}^n$ with $p_{m_c}^n \Gamma_{m_c,g_c}^n$, the overall end-to-end achievable data rate of MTCD m_c over the nth RB can be expressed as

$$R_{m_c}^n = \frac{1}{2} B_{\text{RB}} \log_2(1 + p_{m_c}^n \Gamma_{m_c,g_c}^n) , \quad \forall m_c \in \mathcal{M}_c, n \in \mathcal{N} \tag{3.6}$$

The achievable data rate of the uth neighbouring CUE over the nth RB is given by

$$R_u^n = B_{\text{RB}} \log_2(1 + \gamma_u^n) , \quad \forall u \in \mathcal{U}_c, n \in \mathcal{N} \tag{3.7}$$

where γ_u^n is the SINR over the nth RB given by

$$\begin{aligned} \gamma_u^n &= \frac{p_u^n \, |h_{u,b}^n|^2 d_{u,b}^{-\alpha}}{\sum_{m_c=1}^{M_c} x_{m_c}^n \, p_{m_c}^n |h_{m_c,b}^n|^2 d_{m_c,b}^{-\alpha} + \sigma_0^2} \\ &= p_u^n \, \Gamma_{u,b}^n \end{aligned} \tag{3.8}$$

where $\Gamma_{u,b}^n$ is the channel gain-to-noise ratio between the uth neighbouring CUE and the BS over the nth RB.

3.4 Problem Formulation

The objective is to perform resource block and power allocation for the clustered CM2M network in order to maximize the uplink sum-rate of the neighbouring CUE and clustered MTCDs given interference and power constraints as well as minimum data rate requirements. This can be formulated as the following optimization problem

$$\max_{x^n_{m_c},\, p^n_{m_c}} \quad \sum_{n=1}^{N}\left(\sum_{u=1}^{U_c} R^n_u(p^n_{m_c}) + \sum_{m_c=1}^{M_c} x^n_{m_c} R^n_{m_c}(p^n_{m_c})\right) \tag{3.9}$$

$$\text{s.t.} \quad \sum_{m_c=1}^{M_c} x^n_{m_c} \leq 1,\ x^n_{m_c} \in \{0,1\},\ \forall\, m_c \in \mathcal{M}_c,$$

$$\forall\, n \in \mathcal{N} \tag{3.9a}$$

$$x^n_{m_c} p^n_{m_c} \leq \hat{P}^{\max}_{\mathrm{RB}},\ \ \forall\, m_c \in \mathcal{M}_c, \forall n \in \mathcal{N} \tag{3.9b}$$

$$\sum_{n=1}^{N} x^n_{m_c} p^n_{m_c} \leq P^{\max}_{m_c},\ \ \forall\, m_c \in \mathcal{M}_c \tag{3.9c}$$

$$\sum_{n=1}^{N} x^n_{m_c} p^n_{m_c} |h^n_{m_c,b}|^2 d^{-\alpha}_{m_c,b} \leq I_{\mathrm{th}},\ \ \forall m_c \in \mathcal{M}_c \tag{3.9d}$$

$$\sum_{n=1}^{N} x^n_{m_c} R^n_{m_c} \geq R^{\min}(d^{\max}_{m_c}, \varepsilon^{\max}_{m_c}),\ \ \forall\, m_c \in \mathcal{M}_c, \tag{3.9e}$$

where the objective function to be maximized is the uplink sum-rate of the neighbouring CUE and MTCDs within the cth cluster with $x^n_{m_c}$ and $p^n_{m_c}$ as the RB and power allocation decision variables, respectively, If the nth RB is allocated to MTCD m_c then $x^n_{m_c} = 1$, otherwise $x^n_{m_c} = 0$. Constraint (3.9a) indicates that MTCDs within the cth cluster cannot share the same RB. The peak power constraint over the nth RB is given in (3.9b) where the transmit power of MTCD m_c over the nth RB should not exceed the RB peak power, $\hat{P}^{\max}_{\mathrm{RB}}$. Constraint (3.9c) ensures that the total transmit power of MTCD m_c over the allocated RBs is less than the maximum transmit power, $P^{\max}_{m_c}$. Constraint (3.9d) ensures that the interference to neighbouring CUE from the MTCDs within the cth cluster is less than the interference threshold, I_{th}. The QoS requirements of MTCD m_c are considered in constraint (3.9e) where its achievable data rate should be higher than the minimum data rate threshold, $R^{\min}(d^{\max}_{m_c}, \varepsilon^{\max}_{m_c})$. This is the minimum data rate that provides a Probability of Delay-Bound Violation (PDBV) equal to the maximum delay-violation threshold, $\varepsilon^{\max}_{m_c}$, when the maximum delay-bound is $d^{\max}_{m_c}$ (see Appendix B for details).

3.5 Proposed Resource and Power Allocation Scheme

In this section, a two-phase resource and power allocation scheme is proposed to solve (3.9). In the first phase, an Adaptive Resource Block Allocation (ARBA) algorithm employs an adaptive metric considering power, interference, and minimum data rate requirements to allocate available RBs to the MTCDs within each cluster. In the second phase, an Iterative Power Allocation (IPA) algorithm is used to allocate the power of the MTCDs over the corresponding allocated RBs.

3.5.1 Adaptive Resource Block Allocation (ARBA) Algorithm

The ARBA algorithm allocates available RBs to MTCDs within each cluster based on an adaptive metric that considers the minimum data rate, power, and interference requirements. Let $\mathcal{N}_{m_c}$ be the set of RBs allocated to MTCD m_c with $N_{m_c} = |\mathcal{N}_{m_c}|$, i.e. the cardinality of set $\mathcal{N}_{m_c}$. Then, the total achievable date rate of MTCD m_c when allocated N_{m_c} RBs is given by

$$\begin{aligned} R_{m_c}^{N_{m_c}} &= \sum_{n=1}^{N_{m_c}} R_{m_c}^{n} \\ &= \sum_{n=1}^{N_{m_c}} \frac{1}{2} B_{\mathrm{RB}} \log_2(1 + p_{m_c}^{n} \Gamma_{m_c,g_c}^{n}) \, , \forall m_c \in \mathcal{M}_c, n \in \mathcal{N}_{m_c} \end{aligned} \tag{3.10}$$

Now assume RB n' is assigned to MTCD m_c. Then, the corresponding increment in its achievable data rate can be written as

$$\Delta R_{m_c}^{n'} = R_{m_c}^{N_{m_c}+1} - R_{m_c}^{N_{m_c}} \, , \qquad \forall m_c \in \mathcal{M}_c, n' \in \mathcal{N} \tag{3.11}$$

The increment in the transmit power of MTCD m_c due to assigning RB n' can be expressed as

$$\Delta p_{m_c}^{n'} = \frac{2^{R_{m_c}^{N_{m_c}}} [2^{\Delta R_{m_c}^{n'}} - 1]}{\Gamma_{m_c,g_c}} \, , \tag{3.12}$$

The corresponding increase in the interference induced by MTCD m_c due to assigning RB n' can be written as

$$\Delta I_{m_c}^{n'} = \Delta p_{m_c}^{n'} |h_{m_c,b}|^2 d_{m_c,b}^{-\alpha} \, , \tag{3.13}$$

The data rate gain, power gain, and interference gain due to assigning RB n' to MTCD m_c can be expressed as

$$G_{m_c}^{(R)} = \frac{\Delta R_{m_c}^{n'}}{R^{\min}(d_{m_c}^{\max}, \varepsilon_{m_c}^{\max})} , \tag{3.14}$$

$$G_{m_c}^{(p)} = \frac{\Delta p_{m_c}^{n'}}{P_{m_c}^{\max}} , \tag{3.15}$$

$$G_{m_c}^{(I)} = \frac{\Delta I_{m_c}^{n'}}{I_{\text{th}}} , \tag{3.16}$$

respectively. Now we introduce an adaptive metric for assigning RB n' to MTCD m_c based on the above gains as

$$\Theta_{m_c}^{n'} = \frac{G_{m_c}^{(R)}}{G_{m_c}^{(p)} \times G_{m_c}^{(I)}} , \quad \forall m_c \in \mathcal{M}_c, n' \in \mathcal{N} \tag{3.17}$$

For a given power allocation, $p_{m_c}^n = \frac{P_{m_c}^{\max}}{N}$, the data rate, power, and interference gains for assigning RB n' are calculated for all MTCDs within the cluster using (3.14), (3.15), and (3.16), respectively. Then, the corresponding value of the adaptive metric $\Theta_{m_c}^{n'}$ is obtained. According to the proposed ARBA algorithm, RB n' is assigned to the MTCD m_c' that achieves the maximum value of $\Theta_{m_c}^{n'}$. This implies that n' provides a higher data rate gain with respect to the power and interference gains when allocated to MTCD m_c' compared with other MTCDs in the cluster. The proposed ARBA algorithm is summarized in Algorithm 3.1 and a flowchart of the algorithm is given in Figure 3.2.

3.5.2 Iterative Power Allocation (IPA) Algorithm

After RB allocation is completed for the MTCDs within the cth cluster using the proposed ARBA algorithm, power allocation is performed for the MTCD over the allocated RBs. This is given by the following optimization problem

$$\max_{p_{m_c}^n} \sum_{n=1}^{N} \left(\sum_{u=1}^{U_c} R_u^n(p_{m_c}^n) + \sum_{m_c=1}^{M_c} x_{m_c}^n R_{m_c}^n(p_{m_c}^n) \right) \tag{3.18}$$

$$\text{s.t. } (3.9\text{b}) - (3.9\text{e}).$$

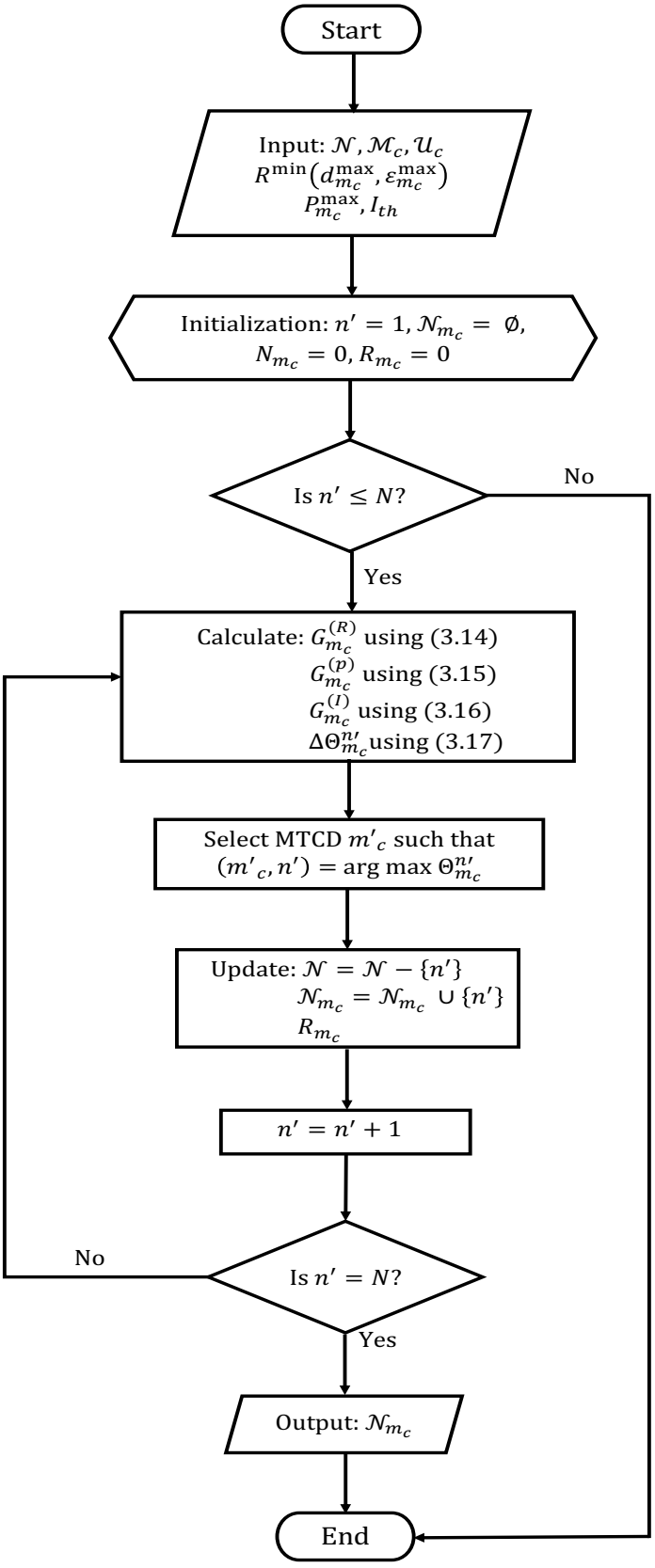

Figure 3.2: Flowchart of the proposed ARBA algorithm.

Algorithm 3.1 : Adaptive Resource Block Allocation(ARBA) Algorithm

1 : **Input** : $\mathcal{N} = \{1, 2, \cdots, N\}, \mathcal{M}_c = \{1, 2, \cdots, M_c\}$,
$\mathcal{U}_c = \{1, 2, \cdots, U_c\}, I_{\text{th}}$,
$R^{\min}(d_{m_c}^{\max}, \varepsilon_{m_c}^{\max}), P_{m_c}^{\max}, p_{m_c}^{n} = \frac{P_{m_c}^{\max}}{N},\ \forall m_c \in \mathcal{M}_c$

2 : *Initialization* : $\mathcal{N}_{m_c} = \phi, N_{m_c} = 0$,
$R_{m_c} = 0,\ \forall m_c \in \mathcal{M}_c$

3 : **for** $n' = 1$ to N **do**

4 : Obtain $G_{m_c}^{(R)}$ using (3.14)

5 : Obtain $G_{m_c}^{(p)}$ using (3.15)

6 : Obtain $G_{m_c}^{(I)}$ using (3.16)

7 : Obtain $\Theta_{m_c}^{n'}$ using (3.17)

8 : $(m_c', n') = \arg\max \Theta_{m_c}^{n'}$

9 : $\mathcal{N} = \mathcal{N} - \{n'\}, \mathcal{N}_{m_c'} = \mathcal{N}_{m_c'} \cup \{n'\}$

10 : Update $R_{m_c},\ \ \forall m_c \in \mathcal{M}_c$

11 : **end**

The objective function in (3.18) is a function of $p^n_{m_c}$ only. However, (3.18) is a non-convex problem with respect to $p^n_{m_c}$ and can be rewritten as

$$\max_{p^n_{m_c}} \ \sum_{n=1}^{N} F^n(p^n_{m_c}) - G^n(p^n_{m_c}) \tag{3.19}$$

$$\text{s.t. } p^n_{m_c} \leq \hat{P}^n_{m_c}, \ \ \forall\, m_c \in \mathcal{M}_c, \forall n \in \mathcal{N}_{m_c} \tag{3.19a}$$

$$\sum_{n=1}^{N_{m_c}} p^n_{m_c} \leq P^{\max}_{m_c}, \ \ \forall\, m_c \in \mathcal{M}_c \tag{3.19b}$$

$$\sum_{n=1}^{N_{m_c}} p^n_{m_c} |h^n_{m_c,b}|^2 d^{-\alpha}_{m_c,b} \leq I_{\text{th}}, \ \ \forall m_c \in \mathcal{M}_c \tag{3.19c}$$

$$\sum_{n=1}^{N_{m_c}} p^n_{m_c} \Gamma^n_{m_c,g_c} \geq N_{m_c}\ (2^{R^{\min}(d^{\max}_{m_c}, \varepsilon^{\max}_{m_c})} - 1),\ \forall\, m_c \in \mathcal{M}_c, \tag{3.19d}$$

where constraint (3.19d) is equivalent to constraint (3.9e) (see Appendix C for details), and $F^n(p^n_{m_c})$ and $G^n(p^n_{m_c})$ are given by

$$F^n(p^n_{m_c}) = \sum_{u=1}^{U_c} \log_2 \left(\sigma_0^2 + p^n_u |h^n_{u,b}|^2 d^{-\alpha}_{u,b} + \sum_{m_c=1}^{M_c} x^n_{m_c} p^n_{m_c} |h^n_{m_c,b}|^2 d^{-\alpha}_{m_c,b} \right)$$
$$+ \frac{1}{2} \sum_{m_c=1}^{M_c} x^n_{m_c} \log_2 \left(\sigma_0^2 + p^n_{m_c} |h^n_{m_c,g_c}|^2 d^{-\alpha}_{m_c,g_c} + \sum_{u=1}^{U_c} p^n_u |h^n_{u,g_c}|^2 d^{-\alpha}_{u,g_c} \right), \tag{3.20}$$

and

$$G^n(p^n_{m_c}) = \sum_{u=1}^{U_c} \log_2 \left(\sigma_0^2 + \sum_{m_c=1}^{M_c} x^n_{m_c} p^n_{m_c} |h^n_{m_c,b}|^2 d^{-\alpha}_{m_c,b} \right)$$
$$+ \frac{1}{2} \sum_{m_c=1}^{M_c} x^n_{m_c} \log_2 \left(\sigma_0^2 + \sum_{u=1}^{U_c} p^n_u |h^n_{u,g_c}|^2 d^{-\alpha}_{u,g_c} \right). \tag{3.21}$$

Both $F^n(p^n_{m_c})$ and $G^n(p^n_{m_c})$ are concave functions of $p^n_{m_c}$, so (3.19) is a Difference of two Convex functions (D.C.) programming problem [82]. This type of problem can be solved by solving a sequence of concave optimization problems [62]. For a given power allocation $\hat{p}^n_{m_c}$, an improved feasible solution to (3.19) can be obtained

by solving the following lower bound optimization problem

$$\max_{p^n_{m_c}} \sum_{n=1}^{N} F^n(p^n_{m_c}) - G^n(\hat{p}^n_{m_c}) - \langle \nabla G^n(\hat{p}^n_{m_c}), p^n_{m_c} - \hat{p}^n_{m_c} \rangle \quad \text{s.t. (3.19a)} - \text{(3.19d)}, \tag{3.22}$$

where $\nabla G^n(\hat{p}^n_{m_c})$ is the gradient of $G^n(p^n_{m_c})$ given by

$$\nabla G^n(\hat{p}^n_{m_c}) = \frac{\sum_{m_c=1}^{M_c} x^n_{m_c} |h_{m_c,b}|^2 d^{-\alpha}_{m_c,b}}{\ln(2)(\sigma_0^2 + \sum_{m_c=1}^{M_c} x^n_{m_c} \hat{p}^n_{m_c} |h_{m_c,b}|^2 d^{-\alpha}_{m_c,b})}. \tag{3.23}$$

For simplicity we denote the sum-rate function in (3.19) as

$$\sum_{n=1}^{N} R^n(p^n_{m_c}) = \sum_{n=1}^{N} F^n(p^n_{m_c}) - G^n(p^n_{m_c}),$$

and the lower bound sum-rate function in (3.22) as

$$\sum_{n=1}^{N} \underline{\mathrm{R}}^n(p^n_{m_c}, \hat{p}^n_{m_c}) = \sum_{n=1}^{N} F^n(p^n_{m_c}) - G^n(\hat{p}^n_{m_c}) - \langle \nabla G^n(\hat{p}^n_{m_c}), p^n_{m_c} - \hat{p}^n_{m_c} \rangle.$$

Proposition 3.1: Given $\hat{p}^n_{m_c}$ that satisfies (3.19a) − (3.19d), $\sum_{n=1}^{N} \underline{\mathrm{R}}^n(p^n_{m_c}, \hat{p}^n_{m_c})$ is a tight lower bound of $\sum_{n=1}^{N} R^n(p^n_{m_c})$.
Proof: See Appendix D.

■

Since (3.22) is convex, it can be efficiently solved using CVX [36]. Moreover, the following proposition proves that improved feasible solutions to (3.19) can be obtained by iteratively solving (3.22). That is, the sum-rate achieved in (3.19) using the optimal solution to (3.22), $p^{*n}_{m_c}$, is no worse than that achieved by $\hat{p}^n_{m_c}$.

Proposition 3.2: If $p^{*n}_{m_c}$ is the optimal solution to (3.22), then $\sum_{n=1}^{N} R^n(p^{*n}_{m_c}) \geq \sum_{n=1}^{N} R^n(\hat{p}^n_{m_c})$.
Proof: See Appendix E.

■

Starting with an initial power allocation $p_{m_c}^{n(0)}$ that satisfies (3.19a) – (3.19d), a solution for (3.22) is obtained. A suitable initial power allocation to MTCD m_c over the nth RB is $p_{m_c}^{n(0)} = \frac{P_{m_c}^{\max}}{N_{m_c}}, \forall m_c \in \mathcal{M}_c$. If the interference induced to the uth neighbouring CUE exceeds the interference threshold I_{th}, the power of MTCD m_c is decreased. This is repeated until the interference induced to the uth neighbouring CUE is less than I_{th}. Then using $p_{m_c}^{n(0)}$, an updated solution to (3.22) $p_{m_c}^{n(1)}$ is obtained. This procedure is repeated until $|| \ p_{m_c}^{n(l)} - p_{m_c}^{n(l-1)}|| \leq \epsilon$ where ϵ is the desired accuracy or the maximum number of iterations, L, is reached.

Proposition 3.3: In the proposed IPA algorithm, the next solution $p_{m_c}^{n(l)}$ is not worse than the previous solution $p_{m_c}^{n(l-1)}$, so it generates a sequence of feasible solutions to (3.22) that converge to a local optimal solution $p_{m_c}^{*n}$ of (3.19).
Proof: See Appendix F.

■

The proposed IPA algorithm is given in Algorithm 3.2 and a flowchart of the algorithm is depicted in Figure 3.3.

Algorithm 3.2 : Iterative Power Allocation (IPA) Algorithm

1 : **Input** : $\mathcal{N}_{m_c} = \{1, 2, \cdots, N_{m_c}\}, P_{m_c}^{\max}, \ \ \forall m_c \in \mathcal{M}_c$
$L, \epsilon, I_{\text{th}}$
2 : *Initialization* : $l = 0$,
$p_{m_c}^{n(0)} = \frac{P_{m_c}^{\max}}{N_{m_c}}, \ \forall m_c \in \mathcal{M}_c, \forall n \in \mathcal{N}_{m_c}$
3 : **repeat**
4 : $l = l + 1$
5 : $\hat{p}_{m_c}^{n} = p_{m_c}^{n(l-1)}$
6 : Solve (3.22) to obtain $p_{m_c}^{n(l)}$
7 : **until** $|| \ p_{m_c}^{n(l)} - p_{m_c}^{n(l-1)}|| \leq \epsilon$ or $l \geq L$
8 : **return** $p_{m_c}^{n(l)}$

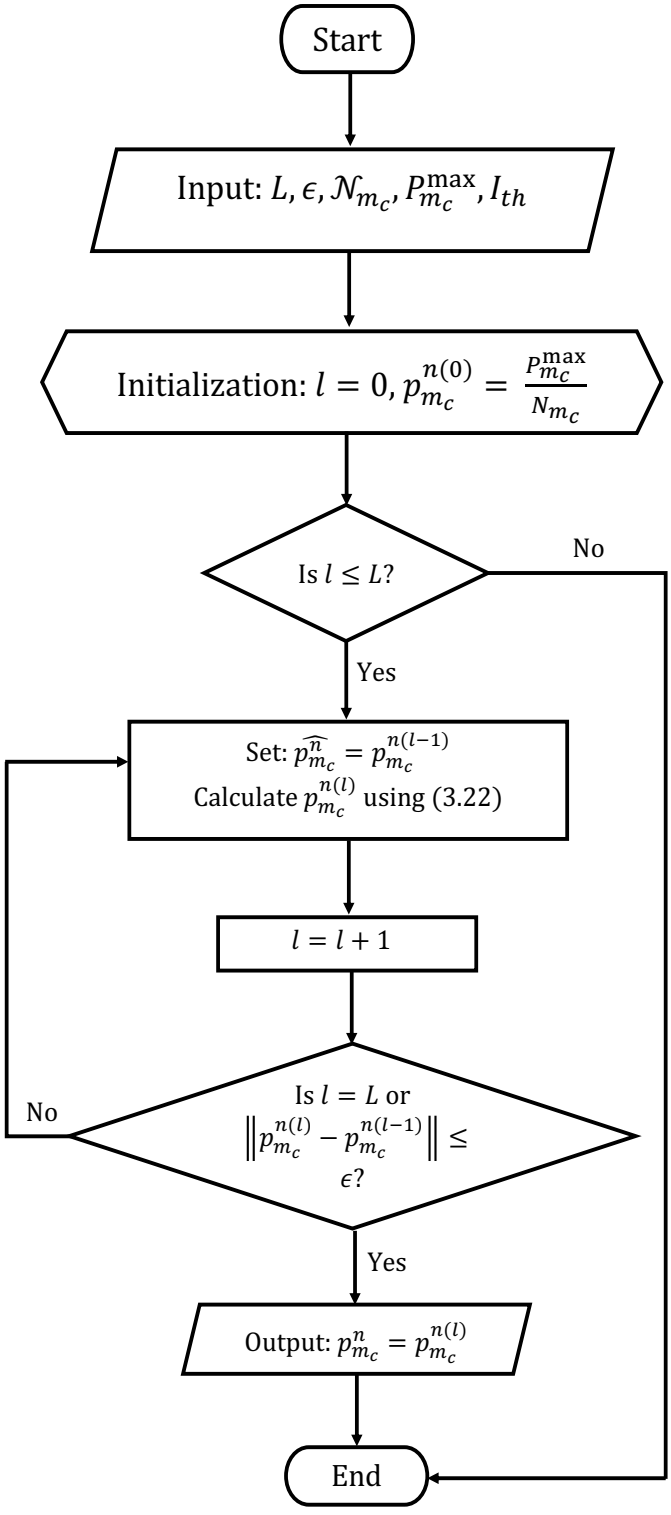

Figure 3.3: Flowchart of the proposed IPA algorithm.

3.6 Performance Evaluation

In order to evaluate the performance of the proposed two-phase resource and power allocation scheme, a cell with radius 500 m in an LTE-A system is considered. The BS is located at the centre of the cell and all CUE are distributed according to a spatial Binomial Point Process (BPP) with the number of neighbouring CUE between 15 and 45. Since there is no interference between clusters, the proposed scheme can be independently applied to each cluster. Therefore, results are presented for the proposed scheme with an MTCG located 175 m from the BS, and this MTCG forms a cluster with radius 100 m which is denoted as cluster 1. MTCDs within the cluster are distributed according to a BPP distribution with the number of MTCDs between 10 and 50. The maximum delay-violation threshold of the clustered MTCDs is $\varepsilon_{m_1}^{\max} = 0.1$ with maximum delay-bound $d_{m_1}^{\max} = 0.2$ ms [59].

A bandwidth of 20 MHz is considered with a total of 100 RBs, each of which has a bandwidth of 180 kHz and a time duration of 0.5 ms [59]. The RB peak power is $\hat{P}_{\mathrm{RB}}^{\max} = 10$ mW [20]. It is assumed that each CUE is assigned one RB. Therefore, the number of available resources for the clustered MTCDs is equal to the number of neighbouring CUE. The channel experiences path loss with path loss constant 10^{-2} and path loss exponent $\alpha = 3$. The large scale shadowing follows a log-normal distribution with zero mean and standard deviation $\sigma_s = 8$ dB, and the small scale Rayleigh fading follows an exponential distribution with unit mean [120]. The Additive White Gaussian Noise (AWGN) power is $\sigma_0^2 = -100$ dBm [59].

The performance of the proposed scheme is compared with two other schemes. Scheme 1 is referred to as ARBA-Uniform PA. In this scheme, the ARBA algorithm is used in the first phase to allocate available RBs to the clustered MTCDs while a uniform power allocation is used in the second phase such that the transmit power of each MTCD is equally distributed over the set of allocated RBs. Scheme 2 is called Random RBA-IPA. In this scheme, the available RBs are randomly allocated to the clustered MTCDs in the first phase while IPA is used to allocate the transmit power of each MTCD over the set of allocated RBs in the second phase.

Figure 3.4 presents the sum-rate of the neighbouring CUE and clustered MTCDs versus the number of clustered MTCDs for the three schemes. These results show that the sum-rate increases with the number of MTCDs for all schemes. However, the proposed scheme outperforms the other schemes. The reason is that it considers the power, interference, and minimum data rate constraints. The ARBA algorithm

Table 3.1: Simulation Parameters.

Parameter	**Value**
Cell radius	500 m
Cluster radius (R_c)	100 m
Distribution of CUE/MTCDs	BPP
Bandwidth	20 MHz
No. of RBs (N)	100
No. of neighbouring CUE (U_c)	$15 - 45$
No. of clustered MTCDs (M_c)	$10 - 50$
Path loss constant (G)	10^{-2}
Path loss exponent (α)	3
Shadowing standard deviation (σ_s)	8 dB
Rayleigh fading (mean of the exponential distribution)	1
MTCD maximum transmit power ($P_{m_c}^{\max}$)	100 mW
RB peak power ($\hat{P}_{\mathrm{RB}}^{\max}$)	10 mW
Noise power (σ^2)	-100 dBm
Interference threshold (I_{th})	$0.2 - 1.4$ mW
Maximum delay-bound ($d_{m_c}^{\max}$)	0.2 ms
Maximum delay-violation threshold ($\varepsilon_{m_c}^{\max}$)	10%

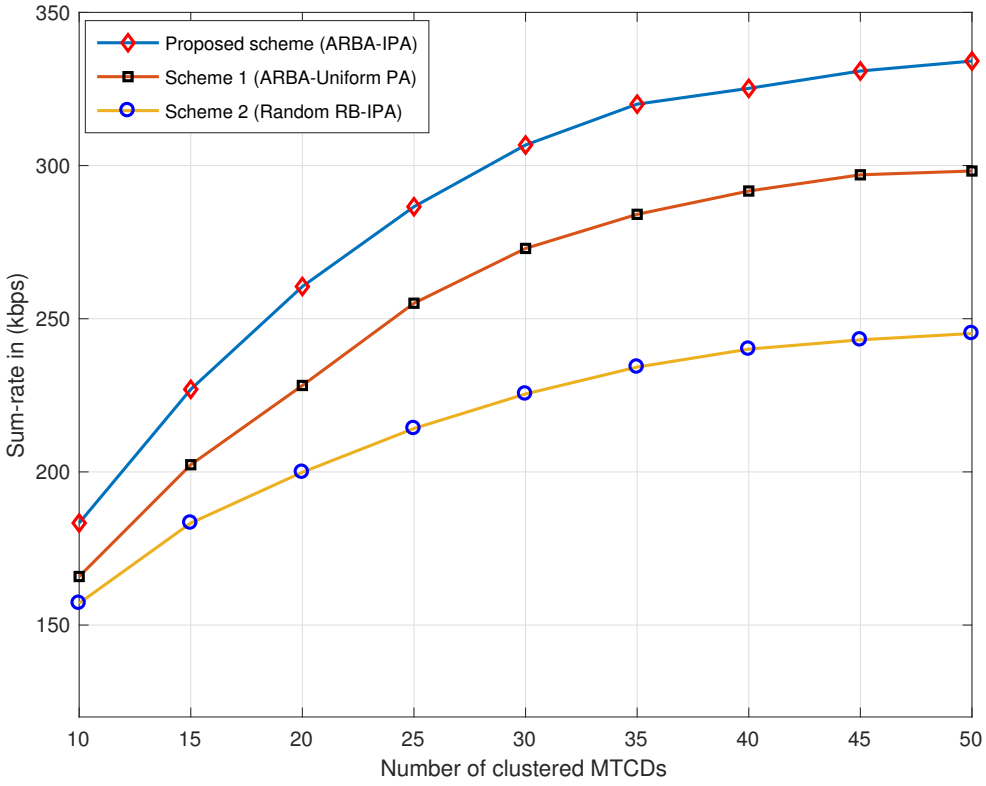

Figure 3.4: Sum-rate versus the number of clustered MTCDs, M_1, for the three schemes.

in the first phase ensures that each MTCD is assigned resources that provide a higher data rate gain with respect to the power and interference gains. Moreover, the IPA algorithm in the second phase ensures that the interference and minimum data rate constraints are satisfied.

Figure 3.5 gives the sum-rate versus the number of clustered MTCDs for different numbers of neighbouring CUE for the proposed scheme. These results show that the sum-rate increases with the number of neighbouring CUE. This is because the number of RBs available for sharing with the clustered MTCDs increases with the number of neighbouring CUE.

The effect of interference induced by clustered MTCDs to neighbouring CUE on the sum-rate with the proposed scheme is illustrated in Figure 3.6. The sum-rate is given versus the number of clustered MTCDs for different interference thresholds. These results indicate that a higher sum-rate is achieved as the interference threshold

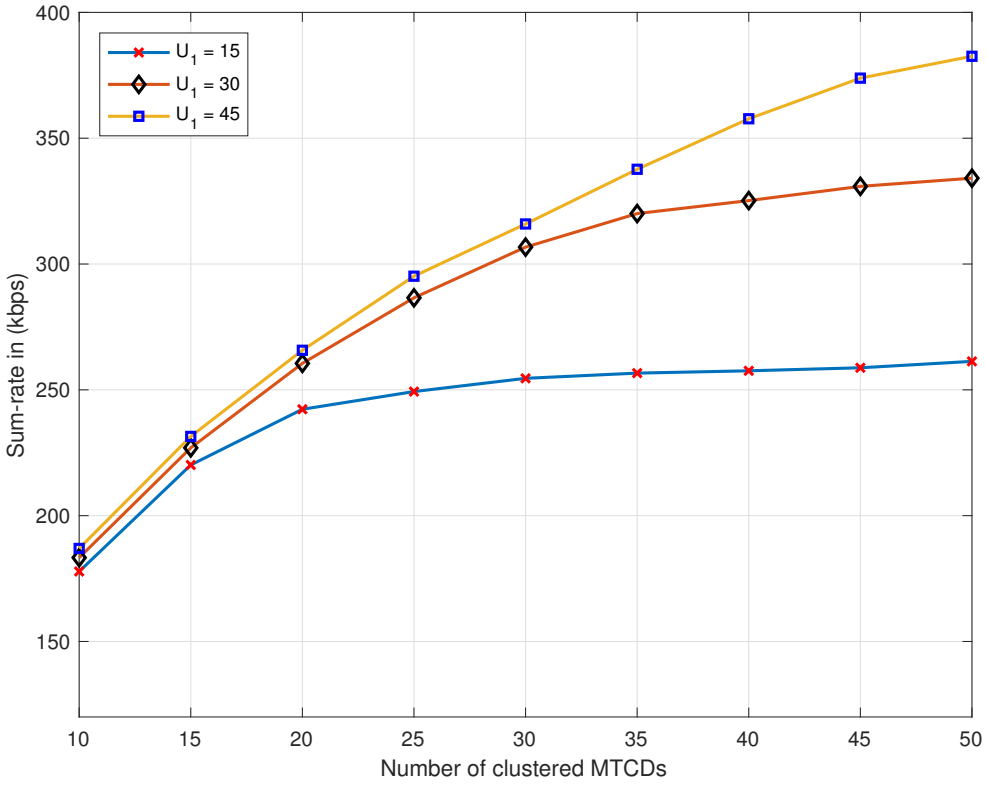

Figure 3.5: Sum-rate versus the number of clustered MTCDs, M_1, for different numbers of neighbouring CUE, U_1, for the proposed scheme.

increases from 0.2 to 1.4 mW. Since the proposed algorithm considers the interference threshold, higher threshold values allow more power to be allocated to the clustered MTCDs, so a higher sum-rate is achieved.

Figure 3.7 illustrates the effect of the MTCD maximum power threshold on the interference induced to neighbouring CUE for the three schemes. This shows that this interference increases with the maximum power threshold. However, the proposed scheme results in less interference than the other schemes. The reason is that the IPA algorithm used in the second phase allocates power to the clustered MTCDs such that the sum-rate is maximized while limiting the interference to neighbouring CUE below the threshold. On the other hand, Scheme 1 results in significant interference since it employs a uniform power allocation which does not consider the interference constraint.

Figure 3.8 presents the sum-rate versus the interference threshold for the three

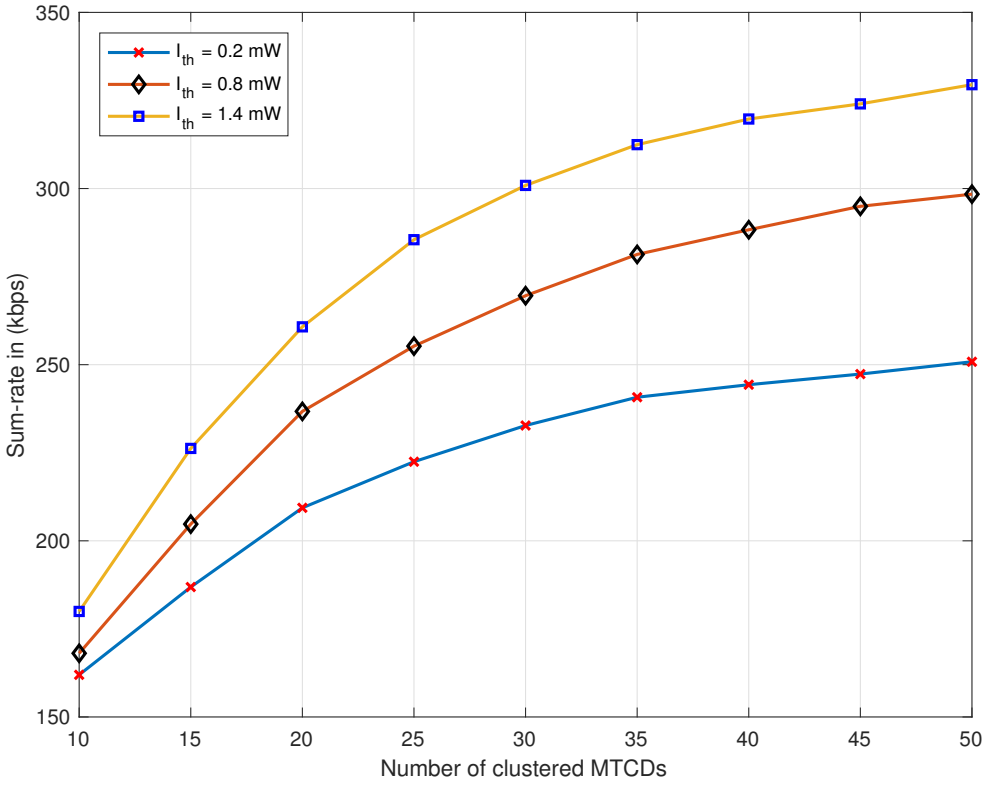

Figure 3.6: Sum-rate versus the number of clustered MTCDs, M_1, for different interference thresholds, I_{th}, for the proposed scheme.

schemes. This shows that the sum-rate increases with the interference threshold. Furthermore, the proposed scheme outperforms the other schemes. This is because the IPA algorithm in the second phase considers the interference and power thresholds while satisfying the minimum data rate requirements. Note that Scheme 1 provides the lowest sum-rate due to the use of a uniform power allocation that does not take these thresholds into account. The sum-rate of the clustered MTCDs versus the interference threshold is given in Figure 3.9 for the three schemes. The sum-rate of the clustered MTCDs increases with the interference threshold. Moreover, the proposed scheme outperforms the other schemes. This is because the IPA algorithm in the second phase considers the interference threshold where higher threshold values allow more power to be allocated to the clustered MTCDs. Therefore, a higher data rate is achieved using the proposed scheme.

Figure 3.10 presents the sum-rate of the neighbouring CUE versus the interference

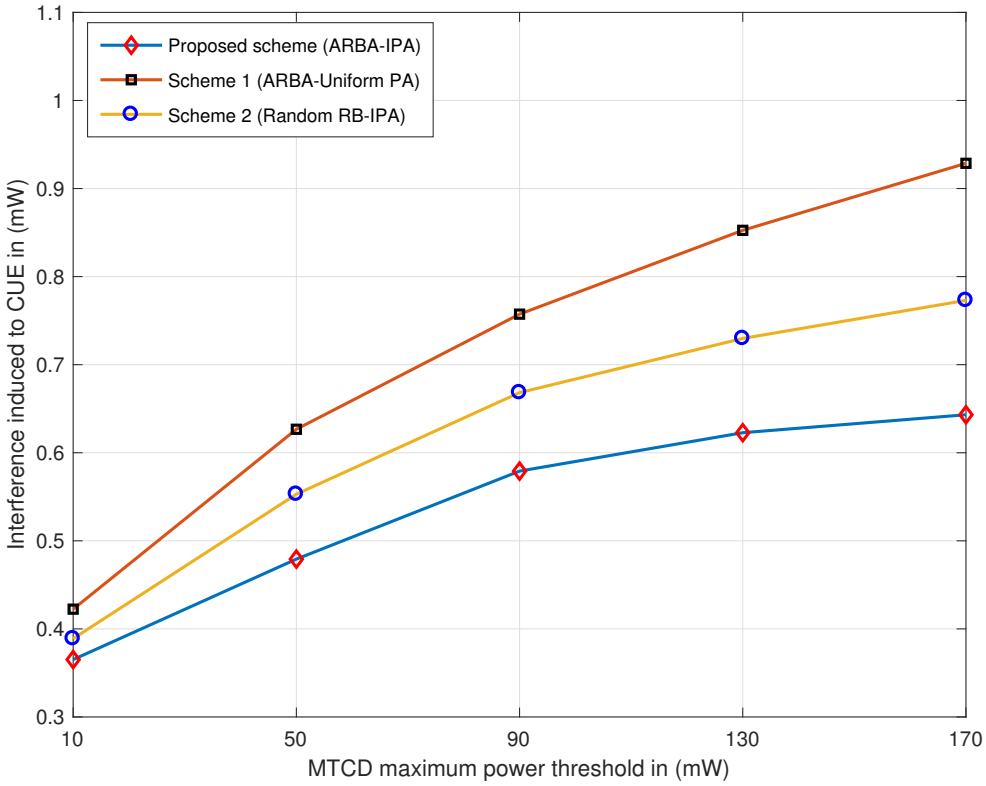

Figure 3.7: Interference induced to neighbouring CUE versus the maximum power threshold, $P_{m_1}^{\max}$, for the three schemes.

threshold for the three schemes. These results indicate that the proposed scheme outperforms Schemes 1 and 2 with a higher sum-rate of the neighbouring CUE. This is because the IPA algorithm in the second phase considers the interference threshold such that the interference induced to the neighbouring CUE is less than this threshold. Hence, a higher data rate is achieved using the proposed scheme compared with the other two schemes. Figure 3.11 presents the sum-rate versus the number of available RBs for the three schemes. These results show that the sum-rate increases with the number of available RBs. Moreover, the proposed scheme provides higher sum-rate than Schemes 1 and 2. The reason is that the proposed ARBA algorithm used in the first phase considers the interference, power, and data rate gains through the adaptive metric. This results in better utilization of the available RBs and demonstrates that the proposed scheme can achieve higher data rates.

The sum-rate of the clustered MTCDs versus the number of available RBs for the

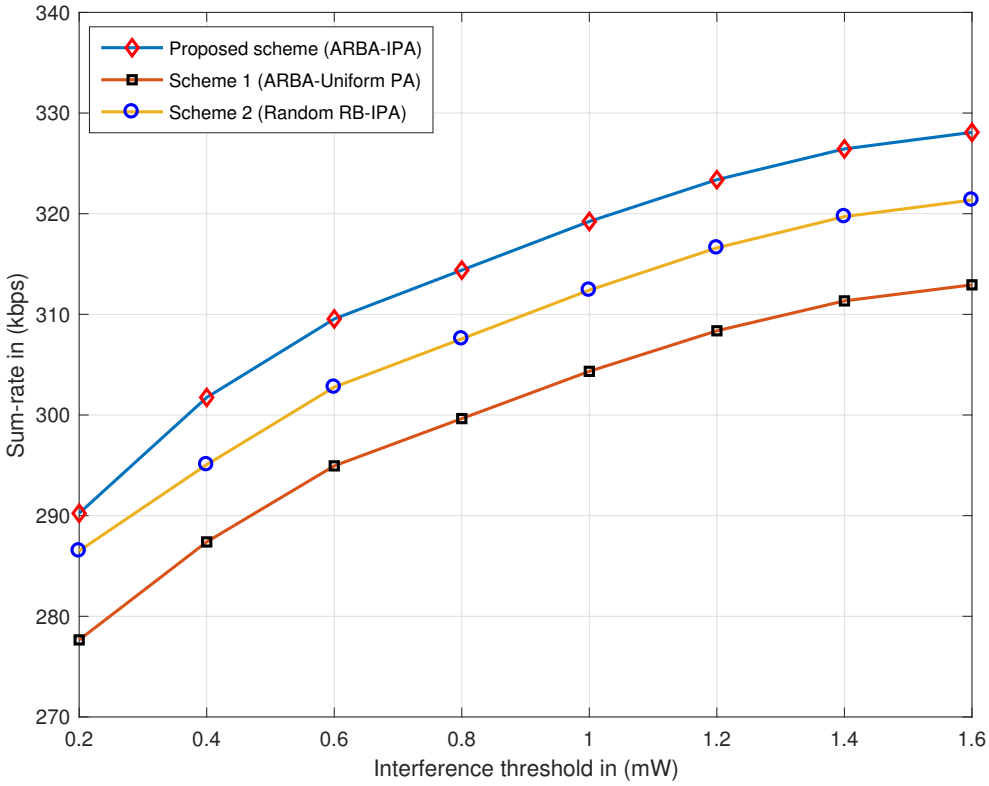

Figure 3.8: Sum-rate versus the interference threshold, I_{th}, for the three schemes.

three schemes is given in Figure 3.12. This shows that the sum-rate of the clustered MTCDs increases with the number of the available RBs. Furthermore, the proposed scheme results in higher data rate than the other two schemes. This is because the proposed ARBA algorithm in the first phase allocates available resources such that higher data rate gain is achieved with respect to the power and interference gains. Figure 3.13 presents the sum-rate of the neighbouring CUE versus the number of available RBs. The results indicate that the proposed scheme outperforms Schemes 1 and 2 in terms of the sum-rate. Since the proposed IPA algorithm in the second phase considers the interference induced to the neighbouring CUE such that it is below the interference threshold. Higher data rate is achieved using the proposed scheme compared with the other two schemes.

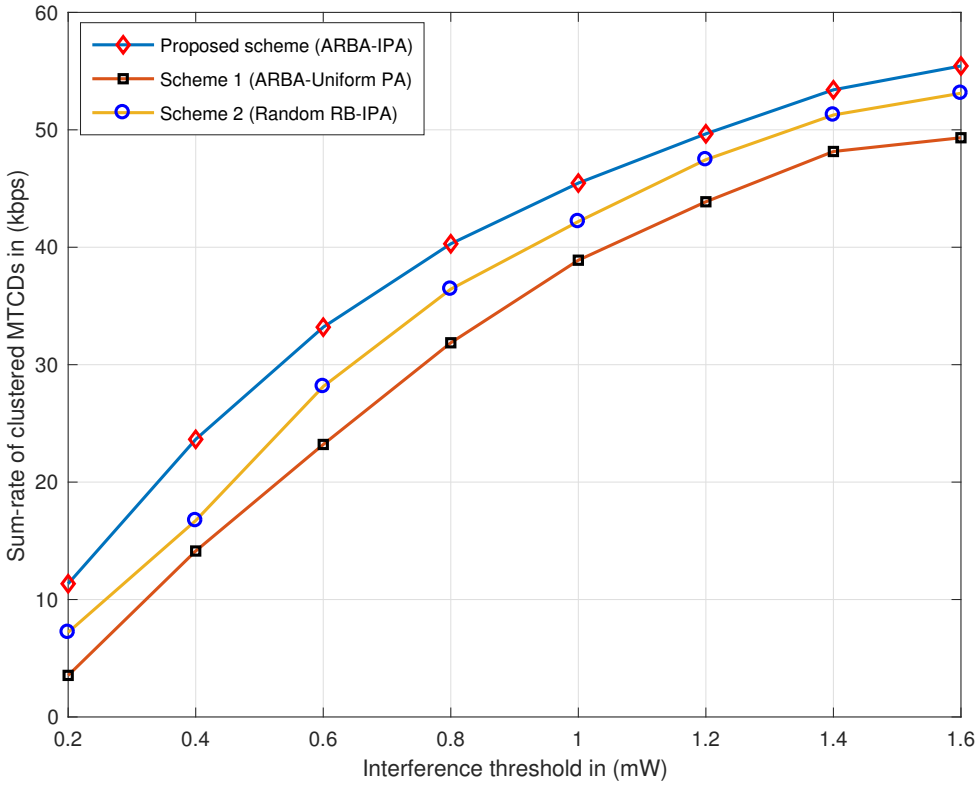

Figure 3.9: Sum-rate of clustered MTCDs versus the interference threshold, I_{th}, for the three schemes.

3.7 Conclusion

In this chapter, clustered CM2M communications underlaying cellular network were considered. In this network, MTCDs are grouped into clusters based on their spatial locations and communicate with the BS via a MTCG in order to accommodate a massive number of MTCDs coexisting with the CUE. Moreover, to exploit spatial spectrum opportunities and overcome spectrum scarcity, underlay cognitive radio is employed so the MTCDs within each cluster can share the spectrum of neighbouring CUE. In order to meet the QoS requirements of the clustered MTCDs while protecting the neighbouring CUE from interference, a joint resource-power allocation problem was formulated and solved using a two-phase resource and power allocation scheme. The goal is to maximize the uplink sum-rate of the neighbouring CUE and clustered MTCDs while adhering to interference, power, and minimum data rate constraints. In

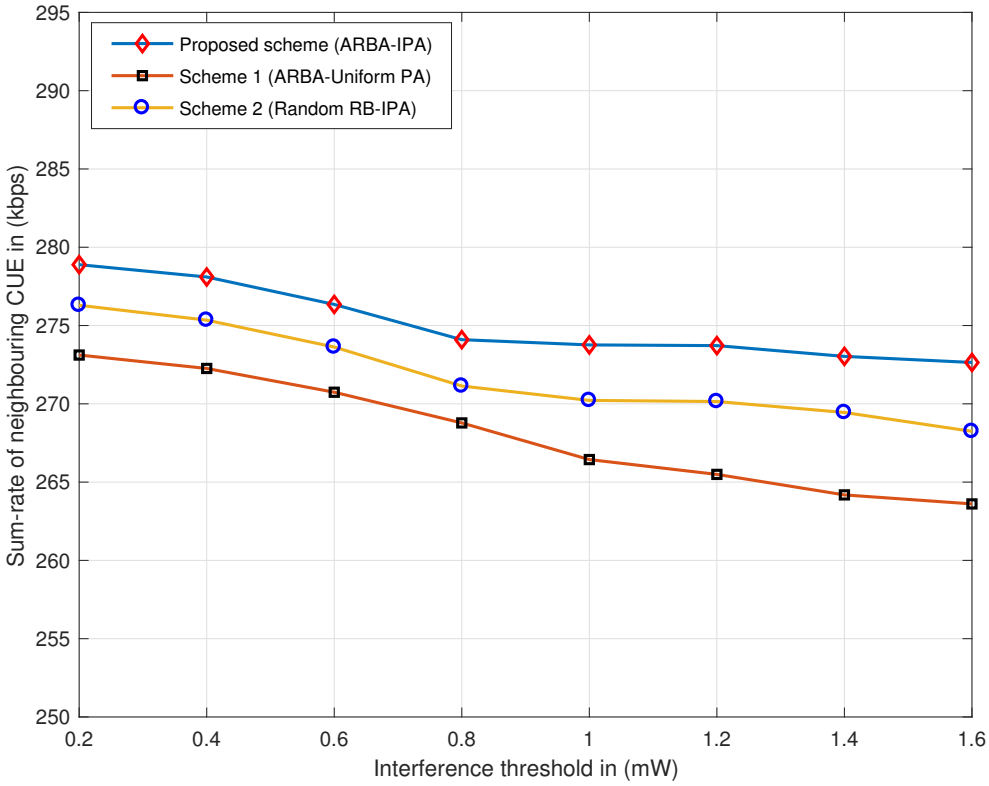

Figure 3.10: Sum-rate of neighbouring CUE versus the interference threshold, I_{th}, for the three schemes.

the first phase, an Adaptive Resource Block Allocation (ARBA) algorithm is employed to allocate available Resource Blocks (RBs) to clustered MTCDs considering these constraints. In the second phase, an Iterative Power Allocation (IPA) algorithm is used to allocate transmit power to the MTCDs over the set of RBs allocated in the first phase. Simulation results were presented which show that the proposed scheme outperforms other schemes in terms of the sum-rate of the network while satisfying the constraints.

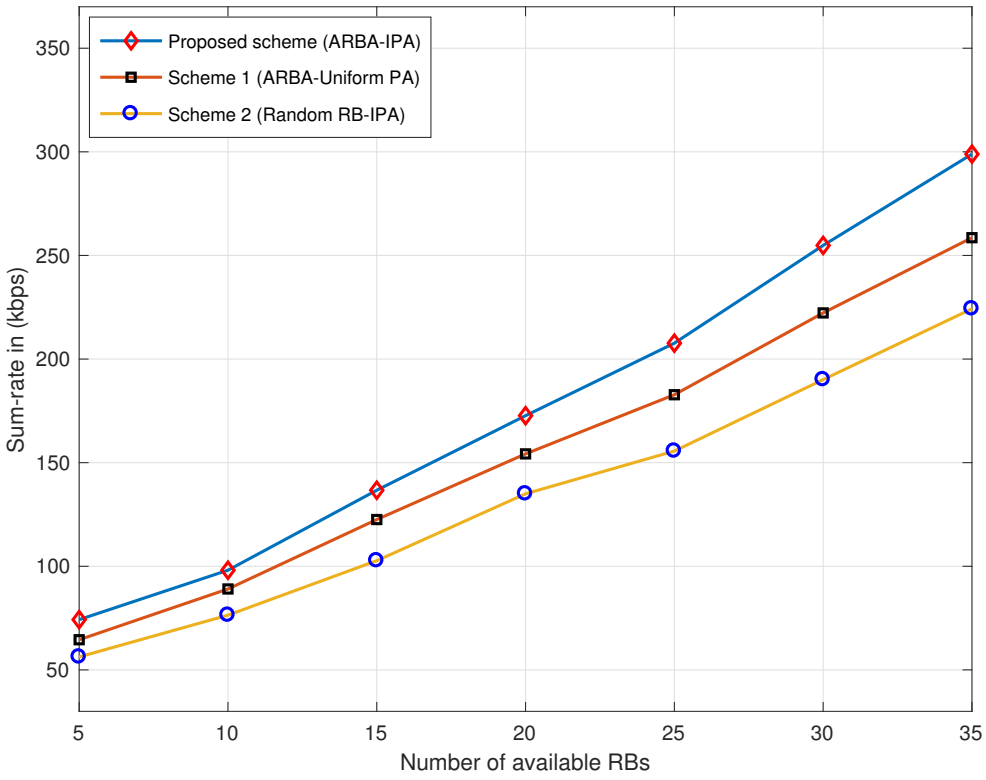

Figure 3.11: Sum-rate versus the number of available RBs for the three schemes.

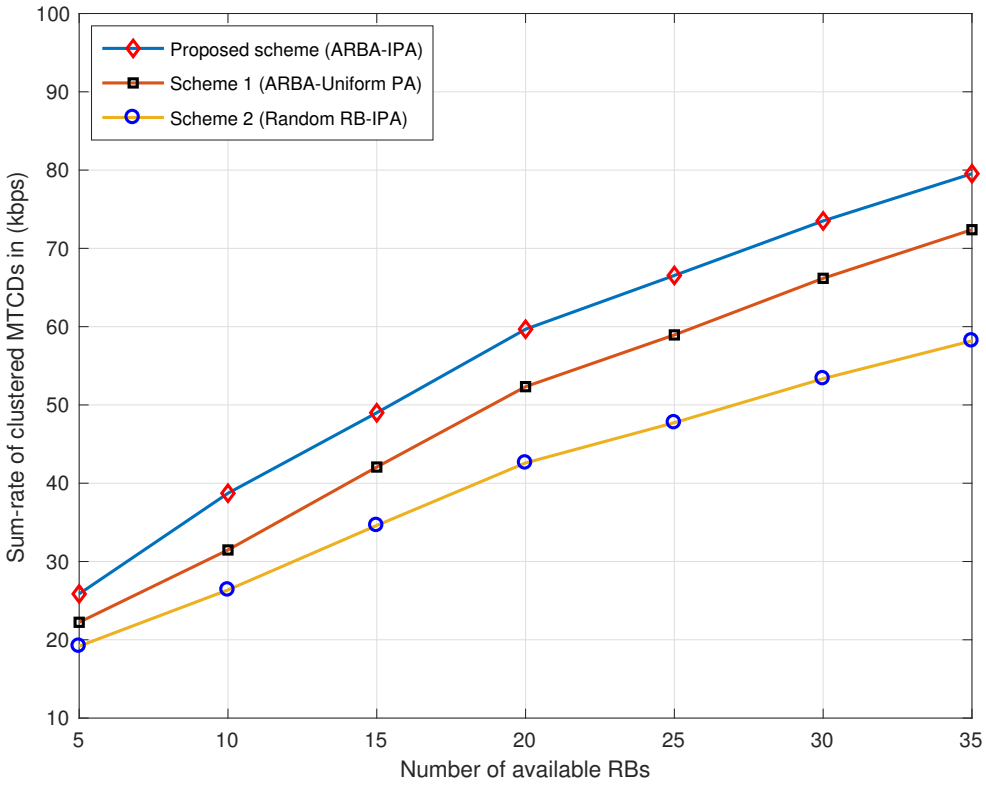

Figure 3.12: Sum-rate of clustered MTCDs versus the number of available RBs for the three schemes.

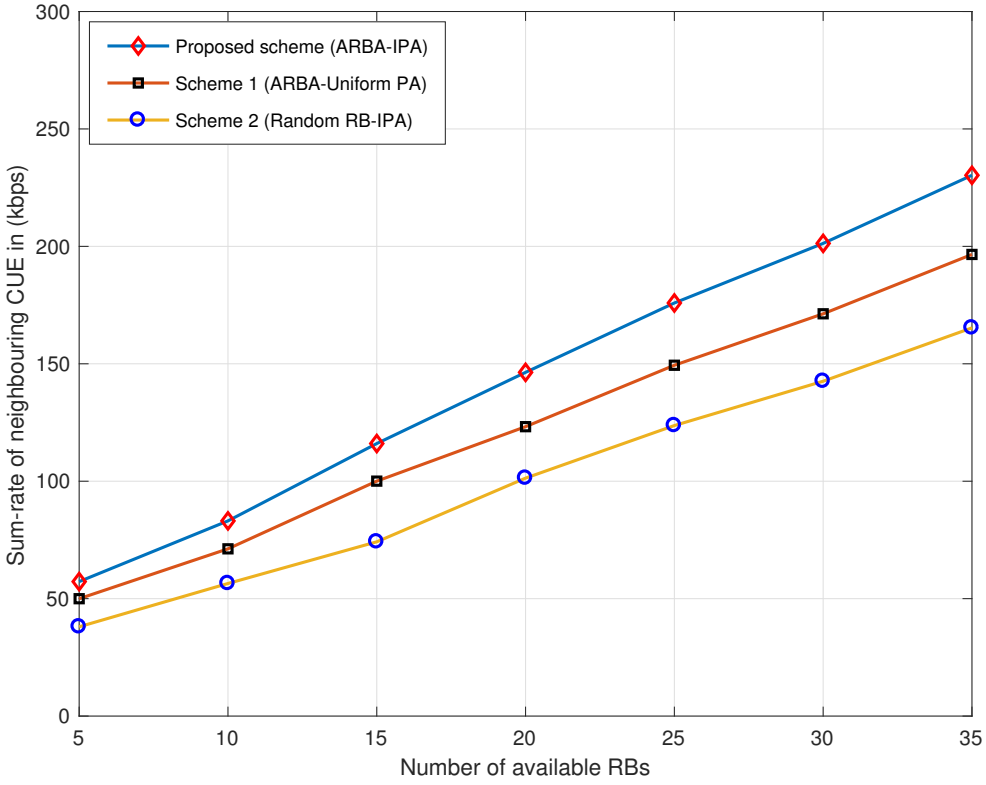

Figure 3.13: Sum-rate of neighbouring CUE versus the number of available RBs for the three schemes.

Chapter 4

Energy-Efficient Cognitive Machine-to-Machine Communications Underlaying Cellular Networks

Information and Communication Technology (ICT) is responsible for more than 3% of energy consumption worldwide [84]. Moreover, it contributes more than 5% of global CO_2 emissions and 37% of these emissions are due to Mobile and Wireless Communications (MWC) [85]. It is projected that more than 29 billion Machine Type Communication Devices (MTCDs) will be connected in the Internet-of-Things (IoT) by 2023 [86]. Several non-cellular wireless technologies including WiFi, Bluetooth, and Zigbee have been considered for enabling Machine-to-Machine (M2M) communications, but they are inefficient and insufficient to support the growing demands of M2M applications [87]. On the other hand, cellular networks provide extended wireless connectivity, more flexible radio resource management, and higher capacity to support M2M deployment. However, the characteristics and Quality-of-Service (QoS) requirements of M2M communications are distinct from those of traditional cellular communications that these networks were originally designed for [59]. While the majority of cellular traffic is on the downlink, M2M traffic is mainly on the uplink with infrequent transmissions under delay limits [88].

The massive number of MTCDs communicating with Base Stations (BSs) and coexisting with cellular users in addition to the diverse QoS requirements pose challenges

for cellular networks in terms of energy and spectral efficiencies. These challenges cannot be overcome by just increasing spectrum occupancy and MTCD transmit power due to the increased interference and limited battery capacity [89] as the majority of MTCDs are battery-operated and may be deployed in remote areas where access for battery replacement is difficult [90]. Hence, Energy Efficiency (EE) in M2M communications is crucial to prolonging network lifetime [92]. It is defined as the ratio between the network sum-rate and consumed power [91], i.e. the information transmitted per joule of energy. The EE of M2M communications can be optimized by adjusting MTCD transmit power such that the power consumption is minimized while achieving performance requirements.

The growing number of MTCDs and increasing data rate demands and QoS requirements create spectrum scarcity. Cognitive M2M (CM2M) communications can be employed to overcome this problem while exploiting the benefits of M2M communications [92, 46]. In CM2M communications, Cognitive Radio (CR) [93] is integrated with M2M communications to increase spectrum utilization where MTCDs are Secondary Users (SUs) that share the spectrum of the Cellular User Equipment (CUE) which are Primary Users (PUs). Conventional CR can improve spectrum utilization from the time and frequency perspectives but ignores the space dimension. Therefore, underlay CR is proposed to utilize spatial spectrum opportunities where MTCDs share the spectrum of CUE in their vicinity while transmitting below an interference threshold [94]. The advantage of this approach is that no coordination is required between the PUs and SUs [95, 96, 97]. Moreover, underlay CR is an efficient solution for M2M communications due to its high spectral and power efficiencies with low transmission delay [57]. Thus, underlay CR can be used to provide spectral and energy efficient M2M communications to realize the IoT vision [77].

4.1 Related Work

There has been research on energy-efficient resource allocation algorithms for CR networks [98]-[103]. Subchannel and power allocation that maximizes the EE of a cognitive Base Station (BS) operating in TV white space was considered in [98]. The proposed two-step solution maximizes the EE while satisfying user minimum data rate requirements and limiting the interference to neighbouring PUs. In [99], energy-efficient resource allocation in Orthogonal Frequency Division Multiplexing (OFDM)-based CR networks was considered under a system power budget constraint, PU in-

terference thresholds, SU data rate requirements, and fairness among users. Two resource allocation schemes for energy-efficient CR systems were proposed in [100]. These schemes employ spectrum sharing combined with soft-sensing, adaptive sensing thresholds, and adaptive power to achieve EE. Energy-efficient CR-based systems were investigated in [101]. Reinforcement learning was used for power-bandwidth efficient modulation where the modulation level is adjusted by determining the available spectrum cognitively. In [102], energy-efficient downlink power control for two-tier macrocell and femtocell cognitive heterogeneous networks was considered. Joint game theoretic power control and interference mitigation under QoS constraints for energy-efficient CR networks was presented in [103].

EE for M2M communications has been investigated in the literature. In [26], a QoS-driven energy-efficient design for the uplink of Long Term Evolution (LTE) networks with M2M/H2H coexistence was proposed. This design considers resource allocation to maximize the capacity under statistical QoS provisioning. Two multiple access strategies, namely Non-orthogonal Multiple Access (NOMA) and Time Division Multiple Access (TDMA) were considered in [111] with the goal of minimizing the total energy consumption of the network via joint power control and time allocation for MTCDs. In [104], scheduling M2M traffic in LTE systems was investigated and a cross-layer resource allocation scheme that minimizes the energy consumption of MTCDs was proposed considering QoS requirements. A joint resource allocation and clustering algorithm to maximize the EE of MTCDs was presented in [105]. In [106], an iterative energy-efficient game-theoretic random access algorithm was developed for joint BS selection and power allocation for M2M devices. The random access success probability weighted by the EE was used as the utility for these devices.

Energy-efficient CM2M communications has had little consideration in the literature [107]-[110]. A selective compressive sensing and resource allocation framework to maximize the EE was introduced in [107]. In this framework, MTCDs sense the assigned channels and report their decisions to a BS. The BS then determines the availability of channels and allocates resources based on local sensing results while maximizing the EE. In [108], an energy-efficient data aggregation scheme to minimize the average total energy consumption of a M2M network was proposed. In this scheme, stochastic geometry is employed to achieve a tradeoff between energy consumption and coverage for different transmission modes. Energy Harvesting (EH) can be integrated with CM2M communications in what is known as EH-based CM2M (EH-CM2M) communications [109]. The downlink energy-efficient resource allocation

problem for EH-CM2M communications was investigated in [110]. A two-stage joint channel selection, peer discovery, power control, and time allocation optimization algorithm was proposed to maximize the EE.

Considering the results above, this chapter investigates the power allocation problem for energy-efficient CM2M communications underlaying cellular networks. Underlay CR is employed to manage the coexistence of MTCDs and CUE and exploit spatial spectrum opportunities to improve spectrum utilization. Two power allocation problems are proposed where the first targets MTCD power consumption minimization while the second considers MTCD EE maximization subject to MTCD transmit power constraints, MTCD minimum data rate requirements, and CUE interference limits. The proposed power consumption minimization problem is transformed into a Geometric Programming (GP) problem and solved iteratively. The proposed EE maximization problem is a nonconvex fractional programming problem. Hence, parametric transformation is used to convert it into an equivalent convex form and this is solved using an iterative approach. To the best of our knowledge, the problem of power allocation of energy-efficient CM2M communications underlaying cellular networks considering MTCD power consumption minimization and MTCD EE maximization subject to the aforementioned constraints has not been considered in the literature.

4.2 Contributions of the Chapter

The contributions of this chapter are as follows.

1. A CM2M communications model is proposed in which underlay CR is employed so that the MTCDs share the spectrum of surrounding CUE subject to the interference they introduce to the CUE being below an interference threshold.

2. A power allocation problem to minimize MTCD power consumption considering the aforementioned constraints is formulated. This problem is transformed into a Geometric Programming (GP) problem and solved iteratively using Algorithm 4.1.

3. A power allocation problem to maximize MTCD EE subject to the constraints above is derived. This is a nonconvex fractional programming problem, so a

parametric transformation is used to convert it into an equivalent convex form which can be solved iteratively using Algorithm 4.2.

4. Simulation results are presented which indicate that the proposed algorithms provide MTCD power allocation with lower power consumption and higher EE than the EPA scheme while satisfying the constraints.

4.3 System Model

Consider the uplink traffic of a cell in an LTE-A cellular network with a BS and U CUE coexisting with M MTCDs. The BS is assumed to be located at the centre of the cell and the CUE and MTCDs are uniformly distributed in the geographic area of the cell. It is assumed that the CUE occupy a total of N RBs where each RB has a bandwidth of 180 kHz and a time duration of 0.5 ms [26]. Denote the set of available RBs by $\mathcal{N} = \{1, 2, \cdots, N\}$, the set of CUE by $\mathcal{U} = \{1, 2, \cdots, U\}$, and the set of MTCDs by $\mathcal{M} = \{1, 2, \cdots, M\}$.

Underlay CR is considered to improve spectrum utilization and manage the coexistence of the CUE and MTCDs by exploiting spatial spectrum opportunities. In this approach, the MTCDs distributively sense the transmissions of surrounding CUE, make an intelligent decision on using CUE resources, and adjust their transmit power such that both the CUE and MTCDs meet their Quality-of-Service (QoS) requirements. That is, the mth MTCD can share the resources of the uth CUE as long as the interference it induces to the CUE is below a threshold, I_{th}. Moreover, it is assumed that multiple MTCDs can share the resources of a CUE but an MTCD can use the resources of only one CUE

The channel model accounts for the distance-based path loss, slow fading due to shadowing, and fast fading due to multipath propagation. Hence, the channel gain between the uth CUE and BS over the nth RB can be expressed as

$$|h_{u,b}^n|^2 = G\, \Gamma_{u,b}^n \beta_{u,b}^n d_{u,b}^{-\alpha}, \tag{4.1}$$

where G is the path loss constant, $\Gamma_{u,b}^n$ is the slow fading gain which has a log-normal distribution with zero mean and standard deviation σ_s, $\beta_{u,b}^n$ is the fast fading gain which has a unit mean exponential distribution, and $d_{u,b}$ is the transmission distance between the uth CUE and BS with path loss exponent α [112]. Similarly, the channel gain between the mth MTCD and BS over the nth RB is denoted by $|h_{m,b}^n|^2$.

The Signal to Interference plus Noise Ratio (SINR) of the uth CUE over the nth RB is given by

$$\gamma_u^n = \frac{p_u^n |h_{u,b}^n|^2}{\sum_{m=1}^{M} p_m^n |h_{m,b}^n|^2 + \sigma_1}, \tag{4.2}$$

where p_u^n and p_m^n are the transmit power of the uth CUE and mth MTCD over the nth RB, respectively, and σ_1 denotes the Additive White Gaussian Noise (AWGN) at the uth CUE. Likewise, the SINR of the mth MTCD over the nth RB can be expressed as

$$\gamma_m^n = \frac{p_m^n |h_{m,b}^n|^2}{p_u^n |h_{u,b}^n|^2 + \sum_{m' \neq m}^{M} p_{m'}^n |h_{m',b}^n|^2 + \sigma_2}, \tag{4.3}$$

where $|h_{m',b}^n|^2$ is the interference channel gain from other MTCDs over the nth RB, and σ_2 is the AWGN for the mth MTCD. Hence, the achievable data rate in b/s/Hz of the uth CUE and mth MTCD over the nth RB can be written as

$$R_u^n = \log_2 \left(1 + \gamma_u^n\right) \tag{4.4}$$

$$R_m^n = \log_2 \left(1 + \gamma_m^n\right), \tag{4.5}$$

respectively. The total power consumption of the mth MTCD over the nth RB can be written as [113],[114]

$$P_m^n = \frac{1}{\zeta} p_m^n + p_{\text{cir}}, \tag{4.6}$$

where $0 \leq \zeta \leq 1$ is the power amplifier efficiency and p_{cir} is the circuit power consumption that depends on blocks for functions such as Digital Signal Processing (DSP), Analog-to-Digital Conversion (ADC), and Digital-to-Analog Conversion (DAC). From (4.5) and (4.6), the EE in b/J/Hz of the mth MTCD over the nth RB can be expressed as

$$\psi_m^n = \frac{R_m^n}{P_m^n}. \tag{4.7}$$

Hence, the total energy efficiency of all MTCDs is given by

$$\psi = \sum_{m=1}^{M} \sum_{n=1}^{N} \psi_m^n = \sum_{m=1}^{M} \sum_{n=1}^{N} \frac{R_m^n}{P_m^n}. \tag{4.8}$$

4.4 Power Allocation Problems and Corresponding Algorithms

In this section, two power allocation problems considering the MTCD transmit power constraints, MTCD minimum data rate requirements, and CUE interference limits are formulated. The first problem addresses power allocation for MTCD power consumption minimization and is solved iteratively using Algorithm 4.1, while the second problem is power allocation for MTCD EE maximization and is solved using Algorithm 4.2. While Algorithm 4.1 focuses on minimizing MTCD total transmit power consumption and Algorithm 4.2 aims at maximizing the MTCD EE, the goal of both algorithms is energy-efficient CM2M communications.

4.4.1 Power Allocation for Power Consumption Minimization

One goal of power allocation is to prolong the lifetime of battery-operated MTCDs to ensure uninterrupted and reliable data transmission in order to meet their QoS requirements. Therefore, minimizing the MTCD power consumption is important. The power consumption minimization problem subject to MTCD power constraints, MTCD minimum data rate requirements, and CUE interference limits is formulated as

$$\min_{p_m^n} \quad \sum_{m=1}^{M}\sum_{n=1}^{N} p_m^n \tag{4.9}$$

$$\text{s.t.} \quad p_m^{\min} \leq p_m^n \leq p_m^{\max}, \qquad \forall m \in \mathcal{M}, n \in \mathcal{N} \tag{4.9a}$$

$$\sum_{m=1}^{M} p_m^n |h_{m,b}^n|^2 \leq I_{\text{th}}, \qquad \forall n \in \mathcal{N} \tag{4.9b}$$

$$R_m^n(p_m^n) \geq R_m^{\min}, \qquad \forall m \in \mathcal{M}, n \in \mathcal{N}, \tag{4.9c}$$

where the objective function to be minimized is the total MTCD transmit power. Constraint (4.9a) is the MTCD power constraint with $p_m^{\min}$ and $p_m^{\max}$ as the minimum and maximum power limits, respectively. Constraint (4.9b) ensures that the interference introduced to the CUE by the underlaying MTCDs is below the interference threshold, I_{th}. The MTCD QoS requirements are considered in constraint (4.9c) which ensures that the achievable data rate of an MTCD is higher than its minimum

data rate threshold, $R_m^{\min}$. The power allocation problem in (4.9) can be reformulated as

$$\min_{p_m^n} \quad \sum_{m=1}^{M}\sum_{n=1}^{N} p_m^n \tag{4.10}$$

$$\text{s.t.} \quad \frac{p_m^{\min}}{p_m^n} \leq 1, \qquad \forall m \in \mathcal{M}, n \in \mathcal{N} \tag{4.10a}$$

$$\frac{p_m^n}{p_m^{\max}} \leq 1, \qquad \forall m \in \mathcal{M}, n \in \mathcal{N} \tag{4.10b}$$

$$\sum_{m=1}^{M} p_m^n |h_{m,b}^n|^2 \leq I_{\text{th}}, \qquad \forall n \in \mathcal{N} \tag{4.10c}$$

$$\log_2(1+\gamma_m^n) \geq R_m^{\min}, \qquad \forall m \in \mathcal{M}, n \in \mathcal{N}. \tag{4.10d}$$

The power allocation problem in (4.10) is a Geometric Programming (GP) problem [34]. After some manipulation, (4.10d) can be expressed as

$$\gamma_m^n = \frac{p_m^n |h_{m,b}^n|^2}{p_u^n |h_{u,b}^n|^2 + \sum_{m' \neq m} p_{m'}^n |h_{m',b}^n|^2 + \sigma_2} \geq 2^{R_m^{\min}} - 1. \tag{4.11}$$

Hence, (4.10) can be written as

$$\max_{p_m^n} \quad \sum_{m=1}^{M}\sum_{n=1}^{N} p_m^n \tag{4.12}$$

$$\text{s.t.} \quad \frac{p_m^{\min}}{p_m^n} \leq 1, \qquad \forall m \in \mathcal{M}, n \in \mathcal{N} \tag{4.12a}$$

$$\frac{p_m^n}{p_m^{\max}} \leq 1, \qquad \forall m \in \mathcal{M}, n \in \mathcal{N} \tag{4.12b}$$

$$\sum_{m=1}^{M} p_m^n |h_{m,b}^n|^2 \leq I_{\text{th}}\ , \qquad \forall n \in \mathcal{N} \tag{4.12c}$$

$$\frac{p_m^n |h_{m,b}^n|^2}{p_u^n |h_{u,b}^n|^2 + \sum_{m' \neq m} p_{m'}^n |h_{m',b}^n|^2 + \sigma_2} \geq 2^{R_m^{\min}} - 1, \ \forall m \in \mathcal{M}, n \in \mathcal{N}, \tag{4.12d}$$

which can be solved iteratively using the following power update function (See Appendix G for the proof)

$$p_m^n(l+1) = \max\left\{ p_m^{\min}, \min\left[p_m^{\max}, \frac{2^{R_m^{\min}} - 1}{\gamma_m^n(l)} p_m^n(l) \right] \right\}. \tag{4.13}$$

A solution for (4.12) is obtained using (4.13). Let $p_m^n(0)$ be an initial power allocation that satisfies (4.12a)-(4.12d). If the interference induced to the CUE exceeds the interference threshold I_{th}, the power of the MTCD is decreased until this interference is less than I_{th}. Then using $p_m^n(0)$, an updated solution to (4.12), $p_m^n(1)$, is obtained. This process is iterated until $|p_m^n(l+1) - p_m^n(l)| \leq \epsilon$ where ϵ is the desired accuracy, or the maximum number of iterations, $L_{\max}$, is reached. The proposed iterative power allocation algorithm to solve (4.12) based on (4.13) is given in Algorithm 4.1.

4.4.2 Power Allocation for Energy Efficiency Maximization

Since data transmission consumes significant MTCD energy, energy-efficient design to minimize power consumption and prolong MTCD life is important. However, MTCD QoS requirements in terms of the achievable data rate should be guaranteed. Hence, EE must consider the tradeoff between performance and power consumption. The power allocation problem for MTCD EE maximization subject to MTCD power constraints, MTCD minimum data rate requirements, and CUE interference limits is then formulated as

$$\max_{p_m^n} \ \psi = \sum_{m=1}^{M}\sum_{n=1}^{N} \psi_m^n = \sum_{m=1}^{M}\sum_{n=1}^{N} \frac{R_m^n}{P_m^n} \tag{4.14}$$

$$\text{s.t.} \ \ p_m^{\min} \leq p_m^n \leq p_m^{\max}, \quad \forall m \in \mathcal{M}, n \in \mathcal{N} \tag{4.14a}$$

$$\sum_{m=1}^{M} p_m^n |h_{m,b}^n|^2 \leq I_{\text{th}}, \quad \forall n \in \mathcal{N} \tag{4.14b}$$

$$R_m^n \geq R_m^{\min}, \quad \forall m \in \mathcal{M}, n \in \mathcal{N}. \tag{4.14c}$$

After some manipulation, (4.14c) can be expressed as

$$\frac{p_m^n}{p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b}^n + \bar{\sigma}_2} \geq 2^{R_m^{\min}} - 1, \tag{4.15}$$

where $H_u^n = \frac{|h_{u,b}^n|^2}{|h_{m,b}^n|^2}$, $H_{m'}^n = \frac{|h_{m',b}^n|^2}{|h_{m,b}^n|^2}$, and $\bar{\sigma}_2 = \frac{\sigma_2}{|h_{m,b}^n|^2}$. The problem in (4.14) is a nonlinear fractional programming problem [115]. To obtain a subtractive form of the objective function, let

$$X_m^n = R_m^n - \delta_m^n P_m^n \,, \tag{4.16}$$

Algorithm 4.1 : Iterative Power Allocation for Power Consumption Minimization

1 : **Input** : $N, L_{\max}, \epsilon$

$R_m^{\min}, p_m^{\min}, p_m^{\max}, p_m^n(0),\ \forall m \in \mathcal{M}$

2 : If the interference induced to the CUE exceeds I_{th}, then the MTCD power is decreased until this interference is less than I_{th}

3 : **for** $m = 1 : M$ **do**

4 : **for** $l = 0 : L_{\max}$ **do**

5 : Each MTCD calculates its SINR over the nth RB using (4.3)

$$\gamma_m^n(l) = \frac{p_m^n(l)|h_{m,b}^n|^2}{p_u^n|h_{u,b}^n|^2+\sum_{m'\neq m}^{M} p_{m'}^n(l)|h_{m',b}^n|^2+\sigma_2}$$

6 : Each MTCD updates its transmit power over the nth RB using (4.13)

$$p_m^n(l+1) = \max\left\{p_m^{\min}, \min\left[p_m^{\max}, \frac{2^{R_m^{\min}}-1}{\gamma_m^n(l)} p_m^n(l)\right]\right\}$$

7 : **if** $|p_m^n(l+1) - p_m^n(l)| \leq \epsilon$ **or** $l \geq L_{\max}$ **then**

8 : $p_m^{*n} = p_m^n(l+1)$

9 : **break**

10 : **else**

11 : $l = l + 1$, go to **Step 5**

12 : **end if**

13 : **end for**

14 : **end for**

where $\delta_m^n = \frac{R_m^n}{P_m^n}$ is an auxiliary variable. Since $R_m^n \geq 0$ and $P_m^n \geq 0$, maximizing

$$\sum_{m=1}^{M}\sum_{n=1}^{N} \psi_m^n = \sum_{m=1}^{M}\sum_{n=1}^{N} \frac{R_m^n}{P_m^n},$$

is equivalent to maximizing

$$\sum_{m=1}^{M}\sum_{n=1}^{N} X_m^n.$$

Hence, the power allocation problem in (4.14) can be expressed as

$$\max_{p_m^n} \sum_{m=1}^{M}\sum_{n=1}^{N} X_m^n = \sum_{m=1}^{M}\sum_{n=1}^{N} R_m^n - \delta_m^n P_m^n \tag{4.17}$$

$$\text{s.t. } p_m^{\min} \le p_m^n \le p_m^{\max}, \quad \forall m \in \mathcal{M}, n \in \mathcal{N} \tag{4.17a}$$

$$\sum_{m=1}^{M} p_m^n |h_{m,b}^n|^2 \le I_{\text{th}}, \quad \forall n \in \mathcal{N} \tag{4.17b}$$

$$\frac{p_m^n}{p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b}^n + \bar{\sigma}_2} \ge 2^{R_m^{\min}} - 1, \ \forall m \in \mathcal{M}, n \in \mathcal{N}. \tag{4.17c}$$

Lemma: If p_m^{*n} is the optimal solution of (4.17) corresponding to $\delta_m^{*n} = \frac{R_m^{*n}}{P_m^{*n}}$, then p_m^{*n} is the optimal solution of (4.14).

Proof: Let p_m^{*n} maximize (4.17) so that

$$X_m^{*n} = R_m^n - \delta_m^{*n} P_m^n \le R_m^{*n} - \delta_m^{*n} P_m^{*n}. \tag{4.18}$$

From the definition of δ_m^{*n}, we have

$$R_m^{*n} - \delta_m^{*n} P_m^{*n} = 0, \tag{4.19}$$

and hence $R_m^n - \delta_m^{*n} P_m^n \le 0$ or $\frac{R_m^n}{P_m^n} \le \delta_m^{*n}$. This shows that $\delta_m^{*n} = \frac{R_m^{*n}}{P_m^{*n}}$ is the maximum of $\frac{R_m^n}{P_m^n}$, so p_m^{*n} is the optimal solution of (4.14).

■

The achievable data rate in (4.17)

$$R_m^n = \log_2\left(1 + \frac{p_m^n}{p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b}^n + \bar{\sigma}_2}\right)$$

is a logarithm of a fraction with respect to p_m^n, which is difficult to solve, so a para-

metric transformation is used [117]. Let

$$\theta_m^n = \frac{p_m^n}{p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b}^n + \bar{\sigma}_2} \,, \tag{4.20}$$

be an auxiliary variable. Then R_m^n can be expressed as

$$R_m^n = \log_2 \left(1 + \theta_m^n\right) + Y\left(\theta_m^n\right) + Z\left(p_m^n\right), \tag{4.21}$$

where $Y(\theta_m^n)$ and $Z(p_m^n)$ are auxiliary functions. To determine $Y(\theta_m^n)$ and $Z(p_m^n)$, a differential method is used so that maximizing

$$\left[\log_2 \left(1 + \frac{p_m^n}{p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b}^n + \bar{\sigma}_2}\right)\right],$$

is equivalent to maximizing

$$\left[\log_2(1 + \theta_m^n) + Y(\theta_m^n) + Z(p_m^n)\right].$$

Taking the derivative of R_m^n in (4.21) with respect to θ_m^n for fixed p_m^n yields

$$\frac{\partial R_m^n}{\partial \theta_m^n} = \frac{1}{1 + \theta_m^n} + Y'(\theta_m^n), \tag{4.22}$$

and setting this to zero gives

$$\begin{aligned} Y'(\theta_m^n) &= -\frac{1}{1 + \theta_m^n} \\ &= -\frac{p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b} + \bar{\sigma}_2}{p_m^n + p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b} + \bar{\sigma}_2}. \end{aligned} \tag{4.23}$$

Hence, $Y(\theta_m^n)$ is given by

$$\begin{aligned} Y(\theta_m^n) &= -\frac{p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b} + \bar{\sigma}_m^2}{p_m^n + p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b} + \bar{\sigma}_2} \theta_m^n \\ &= -\theta_m^n + \frac{p_m^n}{p_m^n + p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b} + \bar{\sigma}_2} \theta_m^n. \end{aligned} \tag{4.24}$$

The auxiliary function $Z(p_m^n)$ is obtained by setting $Y(\theta_m^n) + Z(p_m^n) = 0$ for fixed θ_m^n

which gives

$$Z\left(p_m^n\right) = \frac{p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b} + \bar{\sigma}_2}{p_m^n + p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b} + \bar{\sigma}_2} \theta_m^n$$
$$= \frac{p_m^n}{p_m^n + p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b} + \bar{\sigma}_2}. \tag{4.25}$$

Hence, R_m^n in (4.21) can be expressed as

$$R_m^n = \log_2\left(1 + \theta_m^n\right) - \theta_m^n + \frac{p_m^n}{p_m^n + p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b} + \bar{\sigma}_2} \left(\theta_m^n + 1\right), \tag{4.26}$$

and the power allocation problem in (4.17) becomes

$$\max_{p_m^n} \sum_{m=1}^{M} \sum_{n=1}^{N} \left\{ \log_2\left(1 + \theta_m^n\right) - \theta_m^n - \delta_m^n P_m^n + \frac{p_m^n}{p_m^n + p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b} + \bar{\sigma}_2} \left(\theta_m^n + 1\right) \right\} \tag{4.27}$$

$$\text{s.t. } p_m^{\min} \leq p_m^n \leq p_m^{\max}, \quad \forall m \in \mathcal{M}, n \in \mathcal{N} \tag{4.27a}$$

$$\sum_{m=1}^{M} p_m^n |h_{m,b}^n|^2 \leq I_{\text{th}}, \quad \forall n \in \mathcal{N} \tag{4.27b}$$

$$\theta_m^n \geq 2^{R_m^{\min}} - 1, \quad \forall m \in \mathcal{M}, n \in \mathcal{N}. \tag{4.27c}$$

To obtain a subtractive form of the objective function in (4.27), let

$$S_m^n = \frac{1}{2}\sqrt{\lambda_m^n}\left[p_m^n\left(\theta_m^n + 1\right)\right]^2 - \lambda_m^n \left[p_m^n + p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b} + \bar{\sigma}_2\right], \tag{4.28}$$

where λ_m^n is an auxiliary variable. Taking the derivative of S_m^n with respect to λ_m^n

gives

$$\frac{\partial S_m^n}{\partial \lambda_m^n} = \frac{1}{4\sqrt{\lambda_m^n}} \left[p_m^n \left(\theta_m^n + 1\right)\right]^2 - \left[p_m^n + p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b} + \bar{\sigma}_2\right], \tag{4.29}$$

and setting this to zero results in

$$\lambda_m^n = \left\{ \frac{\left[p_m^n \left(\theta_m^n + 1\right)\right]^2}{4\left[p_m^n + p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b} + \bar{\sigma}_2\right]} \right\}^2, \tag{4.30}$$

Then substituting (4.30) in (4.28) gives

$$S_m^n = \frac{\left[p_m^n \left(\theta_m^n + 1\right)\right]^4}{16\left[p_m^n + p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b} + \bar{\sigma}_2\right]}. \tag{4.31}$$

Since both the numerator and denominator in (4.31) are greater than zero, maximizing S_m^n is equivalent to maximizing

$$\left[\frac{p_m^n}{p_m^n + p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b} + \bar{\sigma}_2} \left(\theta_m^n + 1\right)\right].$$

Using (4.6) and (4.28), the power allocation problem (4.27) can be expressed as

$$\max_{p_m^n} \sum_{m=1}^{M} \sum_{n=1}^{N} \left\{ \log_2(1 + \theta_m^n) - \theta_m^n + \frac{1}{2}\sqrt{\lambda_m^n}[p_m^n(\theta_m^n + 1)]^2 - \lambda_m^n \left[p_m^n + p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n H_{m',b} + \bar{\sigma}_2\right] - \delta_m^n \left(\frac{1}{\zeta} p_m^n + p_{\text{cir}}\right) \right\} \tag{4.32}$$

$$\text{s.t.} \quad p_m^{\min} \leq p_m^n \leq p_m^{\max}, \qquad \forall m \in \mathcal{M}, n \in \mathcal{N} \tag{4.32a}$$

$$\sum_{m=1}^{M} p_m^n |h_{m,b}^n|^2 \leq I_{\text{th}}, \qquad \forall n \in \mathcal{N} \tag{4.32b}$$

$$\theta_m^n \geq 2^{R_m^{\min}} - 1, \qquad \forall m \in \mathcal{M}, n \in \mathcal{N}. \tag{4.32c}$$

Taking the derivative of the objective function in (4.32) with respect to p_m^n and

setting it to zero gives

$$p_m^n = \frac{\lambda_m^n(\sum_{m' \neq m} H_{m',b} + 1) + \frac{\delta_m^n}{\zeta}}{\sqrt{\lambda_m^n}\,(\theta_m^n + 1)^2}. \tag{4.33}$$

The proposed power allocation algorithm to solve (4.32) based on (4.33) is given in Algorithm 4.2. Starting with an initial power allocation $p_m^n(0)$ that satisfies (4.32a)-(4.32c), an updated solution to (4.32), $p_m^n(1)$, is obtained using (4.33). This process is iterated until $|p_m^n(l+1) - p_m^n(l)| \leq \epsilon$ where ϵ is the desired accuracy, or the maximum number of iterations, $L_{\max}$, is reached. Since (4.32) is a maximization problem, θ_m^n, λ_m^n, and δ_m^n increase or at least do not decrease in each iteration [117]. Hence, Algorithm 4.2 will converge by alternately updating θ_m^n, λ_m^n, and δ_m^n and then solving for p_m^n in (4.33). Note that the power allocation problem (4.14) is a nonconvex fractional programming maximization problem. However, it is transformed to the equivalent convex maximization problem in (4.32) which optimizes p_m^n in (4.33) for fixed θ_m^n, λ_m^n, and δ_m^n. Therefore, Algorithm 4.2 will converge to the globally optimal solution of (4.33).

4.4.3 Complexity Analysis

The time complexity of the proposed iterative power allocation algorithms is now examined. The complexity of Algorithm 4.1 is $O(ML_{\max})$ where M is the number of MTCDs and $L_{\max}$ is the maximum number of iterations. As will be discussed in Section V, Algorithm 4.1 converges after about 8 iterations. Hence, it efficiently performs power allocation to achieve minimum MTCD power consumption. Algorithm 4.2 has complexity $O(ML_{\max})$ and converges after about 10 iterations providing an efficient power allocation for MTCD EE maximization.

4.5 Performance Evaluation

In this section, the performance of the proposed power allocation algorithms for MTCD power consumption minimization and EE maximization is evaluated. The performance of these algorithms is compared with that of the Equal Power Allocation (EPA) scheme [118], [119]. With EPA, the transmit power is allocated equally among all MTCDs. A single cell with radius 500 m in an LTE-A system is considered. The BS is located at the centre and the CUE and MTCDs are uniformly distributed

Algorithm 4.2: Iterative Power Allocation for Energy Efficiency Maximization

1 : **Input** : $N, L_{\max}, \epsilon, R_m^{\min}, p_m^n(0), \ \ \forall m \in \mathcal{M}$

2 : **for** $m = 1 : M$ **do**

3 : **for** $l = 0 : L_{\max}$ **do**

4 : Each MTCD calculates $\theta_m^n(l)$ using (4.20)

$$\theta_m^n(l) = \frac{p_m^n(l)}{p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n(l) H_{m',b}^n + \bar{\sigma}_2}$$

5 : Each MTCD calculates $\lambda_m^n(l)$ using (4.30)

$$\lambda_m^n(l) = \left\{ \frac{[p_m^n(l)(\theta_m^n(l)+1)]^2}{4\left[p_m^n(l) + p_u^n H_u^n + \sum_{m' \neq m} p_{m'}^n(l) H_{m',b} + \bar{\sigma}_2\right]} \right\}^2$$

6 : Each MTCD calculates $\delta_m^n(l)$ using

$$\delta_m^n(l) = \frac{R_m^n(l)}{P_m^n(l)}$$

7 : Each MTCD updates its transmit power using (4.33)

$$p_m^n(l+1) = \frac{\lambda_m^n(l)(\sum_{m' \neq m} H_{m',b} + 1) + \frac{\delta_m^n(l)}{\zeta}}{\sqrt{\lambda_m^n(l)}(\theta_m^n(l)+1)^2}$$

8 : **if** $|p_m^n(l+1) - p_m^n(l)| \leq \epsilon$ **or** $l \geq L_{\max}$ **then**

9 : $p_m^{*n} = p_m^n(l+1)$

$\delta_m^{*n} = \frac{R_m^{*n}}{P_m^{*n}}$

10 : **break**

11 : **else**

12 : $l = l + 1$, go to **Step 4**

13 : **end if**

14 : **end for**

15 : **end for**

in the geographic area of the cell. The number of CUE is set to 5 and the number of underlaying MTCDs is varied between 2 and 20. The path loss exponent is $\alpha = 3$ and the path loss constant is $G = 10^{-2}$. The small scale Rayleigh fading follows an exponential distribution with unit mean, and the large scale shadowing follows

a log-normal distribution with zero mean and standard deviation $\sigma_s = 8$ dB [120]. The Additive White Gaussian Noise (AWGN) power is $\sigma_0^2 = -100$ dBm [59]. The circuit power consumption is $p_{\text{cir}} = 5$ dBm and the power amplifier efficiency ζ is 30%. The QoS requirements of the MTCDs are expressed in terms of the minimum data rate threshold $R_m^{\min}$ which is set to a percentage of the MTCD achievable data rate without any interference from the CUE or other MTCDs, R_{o}.

4.5.1 Power Consumption Minimization Performance Evaluation

The performance of Algorithm 4.1 that allocates power for MTCD power consumption minimization is presented in Figures 4.1 to 4.4. MTCD total power consumption for different MTCD maximum power limits $p_m^{\max}$ is given in Figure 4.1 for $R_m^{\min} = 75\%R_{\text{o}}$, and $I_{\text{th}} = 1$ mW. These results show that the MTCD total power consumption increases as $p_m^{\max}$ increases from 10 dBm to 20 dBm. This is because higher values of $p_m^{\max}$ allow more power to be allocated to the MTCDs. Moreover, Algorithm 4.1 converges after 8 iterations in all cases with an average computation time of 0.053 s. Figure 4.2 gives the MTCD total power consumption for different minimum data rate thresholds $R_m^{\min}$ for $p_m^{\max} = 15$ dBm and $I_{\text{th}} = 1$ mW. As $R_m^{\min}$ increases from $50\%R_{\text{o}}$ to $100\%R_{\text{o}}$, the total power consumption increases. The reason is that a higher minimum data rate threshold requires more power to be allocated to the MTCDs to achieve the desired data rate.

The effect of the CUE interference threshold I_{th} on the total power consumption of the underlaying MTCDs is presented in Figure 4.3 where the total power consumption is given for different interference thresholds, $p_m^{\max} = 15$ dBm, and $R_m^{\min} = 75\%R_{\text{o}}$. This shows that higher interference thresholds result in higher total power consumption. This is because Algorithm 4.1 allocates power to the MTCDs considering the CUE interference threshold, hence a higher interference threshold allows more power to be allocated to the MTCDs. Figure 4.4 presents the interference introduced to the CUE versus the MTCD maximum power limit $p_m^{\max}$ for Algorithm 4.1 and the EPA scheme. This shows that Algorithm 4.1 results in lower interference levels. This is because Algorithm 4.1 ensures that the interference introduced to the CUE does not exceed the interference threshold when allocating power to the underlaying MTCDs. However, the EPA scheme allocates power to the MTCDs based on the maximum power limit regardless of the resulting interference.

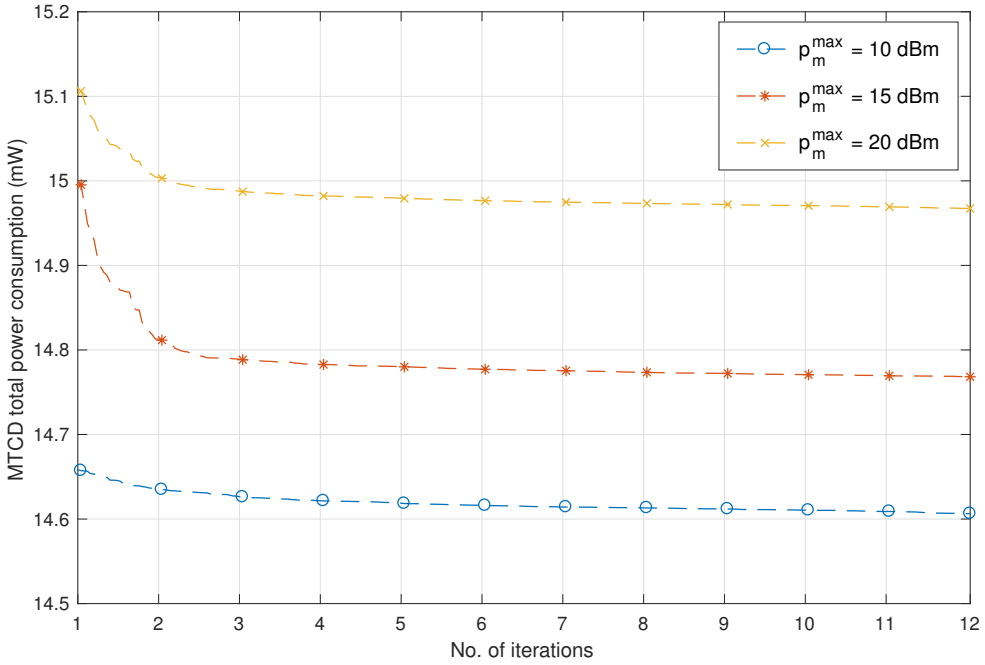

Figure 4.1: MTCD total power consumption versus the number of iterations for different MTCD maximum power limits, $p_m^{\max}$, with Algorithm 4.1.

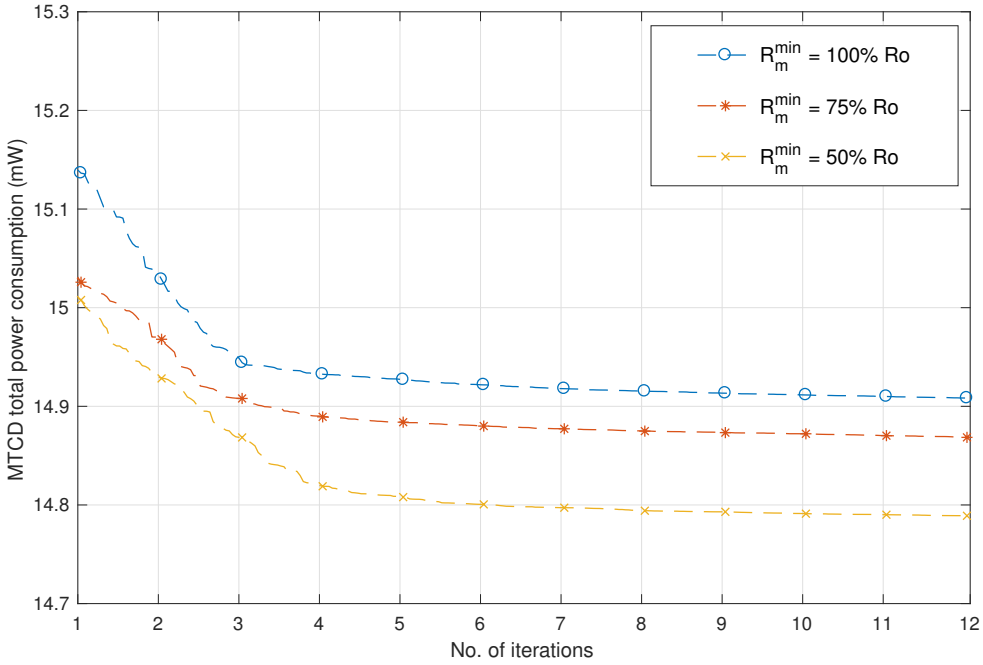

Figure 4.2: MTCD total power consumption versus the number of iterations for different MTCD minimum data rate thresholds, $R_m^{\min}$, with Algorithm 4.1.

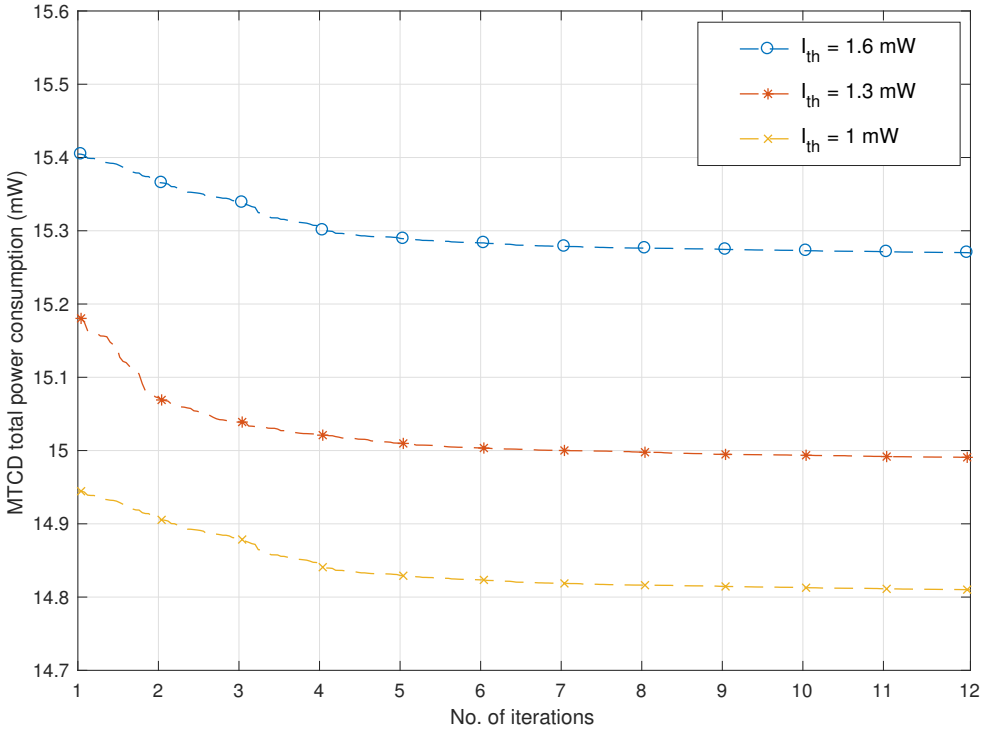

Figure 4.3: MTCD total power consumption versus the number of iterations for different CUE interference thresholds, $I_{\max}$, with Algorithm 4.1.

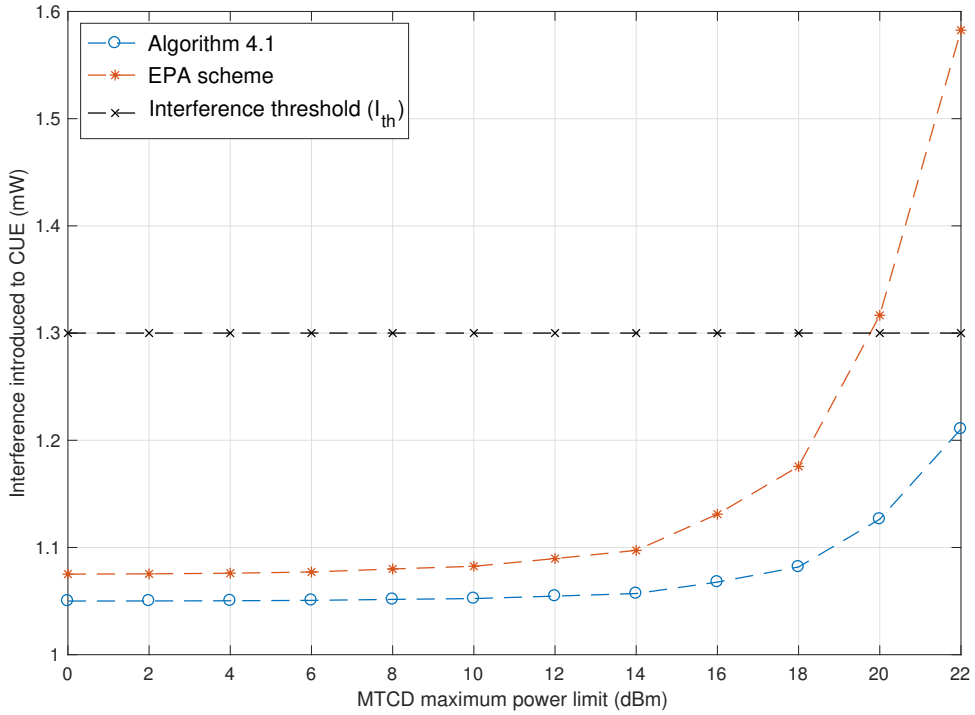

Figure 4.4: Interference introduced to the CUE versus the MTCD maximum power limit, $p_m^{\max}$, with Algorithm 4.1 and the EPA scheme.

4.5.2 Energy Efficiency Maximization Performance Evaluation

The performance of Algorithm 4.2 that allocates power for MTCD EE maximization is presented in Figures 4.5 to 4.10. The EE of the underlaying MTCDs versus the number of iterations for $M = 10$ with Algorithm 4.2 is given in Figure 4.5. This shows that Algorithm 4.2 converges for all MTCDs after about 10 iterations with an average computation time of 0.106 s. Figure 4.6 presents the EE of the MTCDs versus the number of MTCDs for Algorithm 4.2 and the EPA scheme with $p_m^{\max} = 15$ dBm, $I_{\text{th}} = 1$ mW, and $R_m^{\min} = 75\%R_{\text{o}}$. This shows that as the number of MTCDs grows from 2 to 20, the EE decreases because more MTCDs are sharing the same amount of resources. Hence, more power is allocated to overcome the interference from other MTCDs and achieve the QoS requirements. Moreover, as the number of MTCDs increases the gap between Algorithm 4.2 and the EPA scheme increases as Algorithm 4.2 provides a higher EE because it ensures that the MTCDs meet their minimum data rate thresholds.

The EE of the MTCDs versus the number of MTCDs for different minimum data rate thresholds, $R_m^{\min}$ with Algorithm 4.2 is given in Figure 4.7 for $p_m^{\max} = 15$ dBm and $I_{\text{th}} = 1$ mW. This shows that the EE decreases with the number of MTCDs. Moreover, as the minimum data rate threshold increases, the EE of the MTCDs decreases. This is because a higher minimum data rate threshold requires more power to be allocated to the MTCDs. The effect of the CUE transmit power on the EE of the underlaying MTCDs is shown in Figures 4.8 and 4.9. The EE of the MTCDs versus CUE transmit power with Algorithm 4.2 and the EPA scheme is given in Figure 4.8 for $p_m^{\max} = 15$ dBm, $I_{\text{th}} = 1$ mW, and $R_m^{\min} = 75\%R_{\text{o}}$. These results show that the EE decreases as the CUE transmit power increases because more interference is introduced to the underlaying MTCDs. Consequently, more power is allocated to the MTCDs to overcome the interference. Further, as the CUE transmit power increases, the gap between Algorithm 4.2 and the EPA scheme increases as Algorithm 4.2 provides a higher EE because it ensures that the MTCD minimum data rate requirements are satisfied. Figure 4.9 presents the EE of the MTCDs versus CUE transmit power for different minimum data rate thresholds, $R_m^{\min}$, with Algorithm 4.2. This shows that a higher MTCD minimum data rate threshold results in a lower EE since more power is allocated to achieve the required data rate.

Figure 4.10 presents the interference introduced to the CUE versus the MTCD

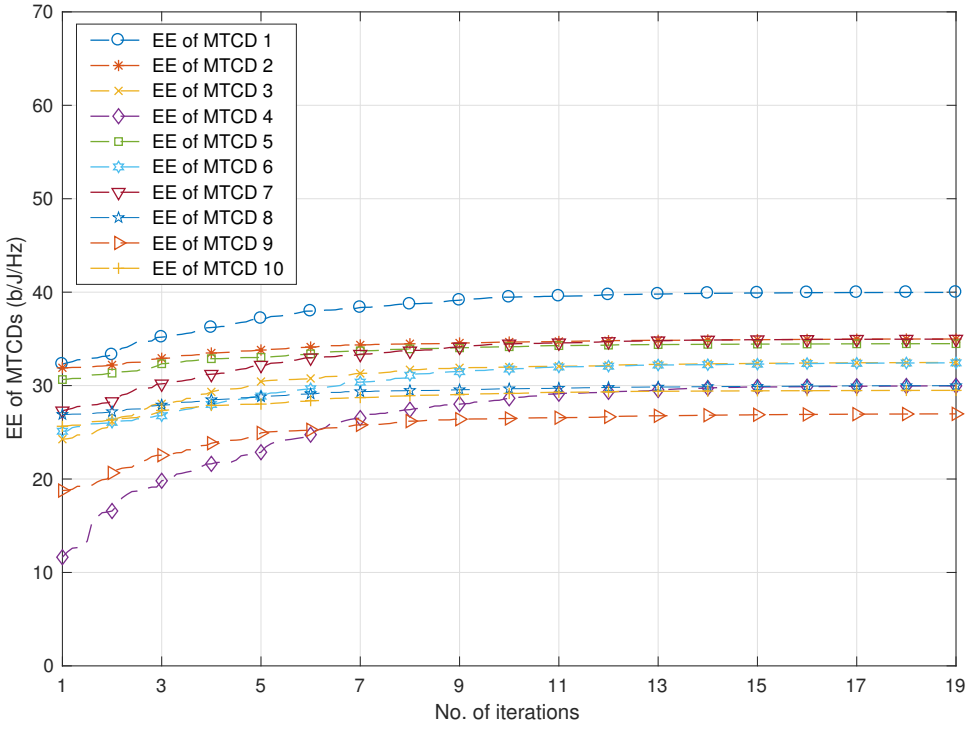

Figure 4.5: MTCD EE versus the number of iterations with Algorithm 4.2.

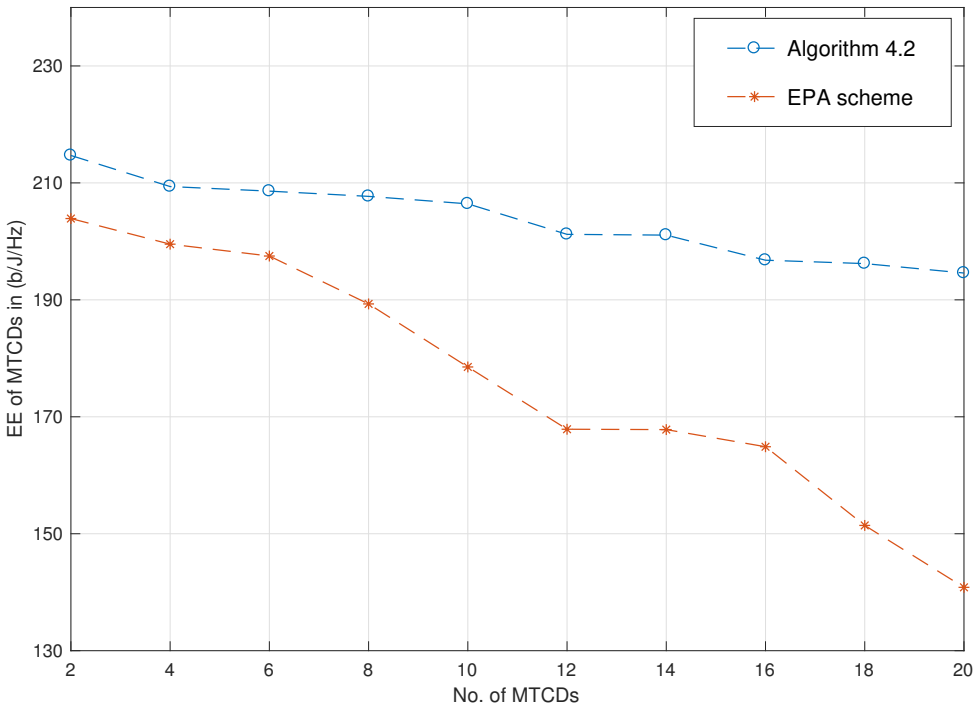

Figure 4.6: MTCD EE versus the number of MTCDs with Algorithm 4.2 and the EPA scheme.

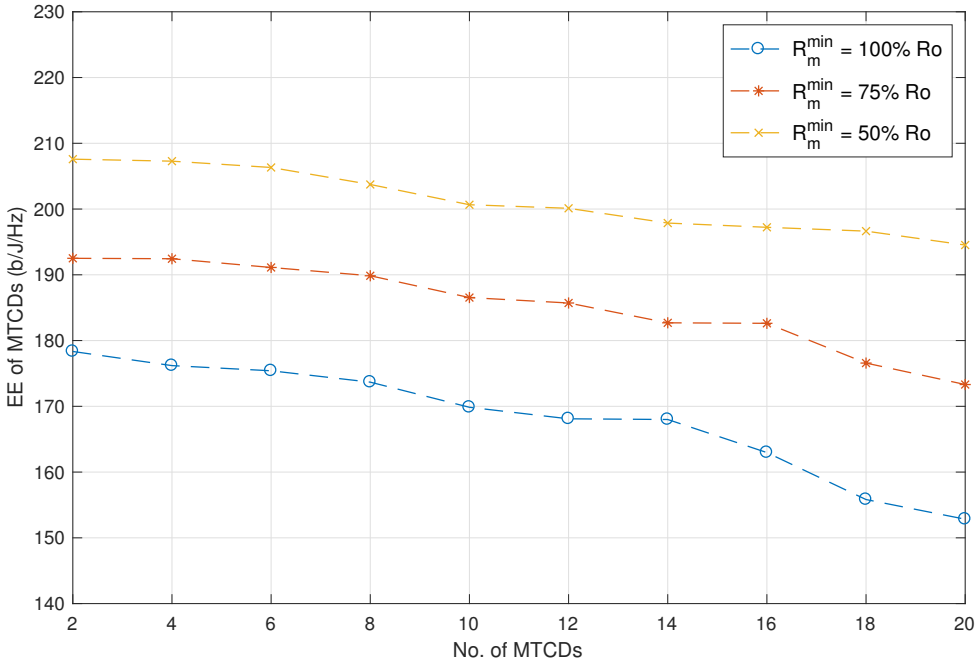

Figure 4.7: MTCD EE versus the number of MTCDs for different minimum data rate thresholds, $R_m^{\min}$, with Algorithm 4.2.

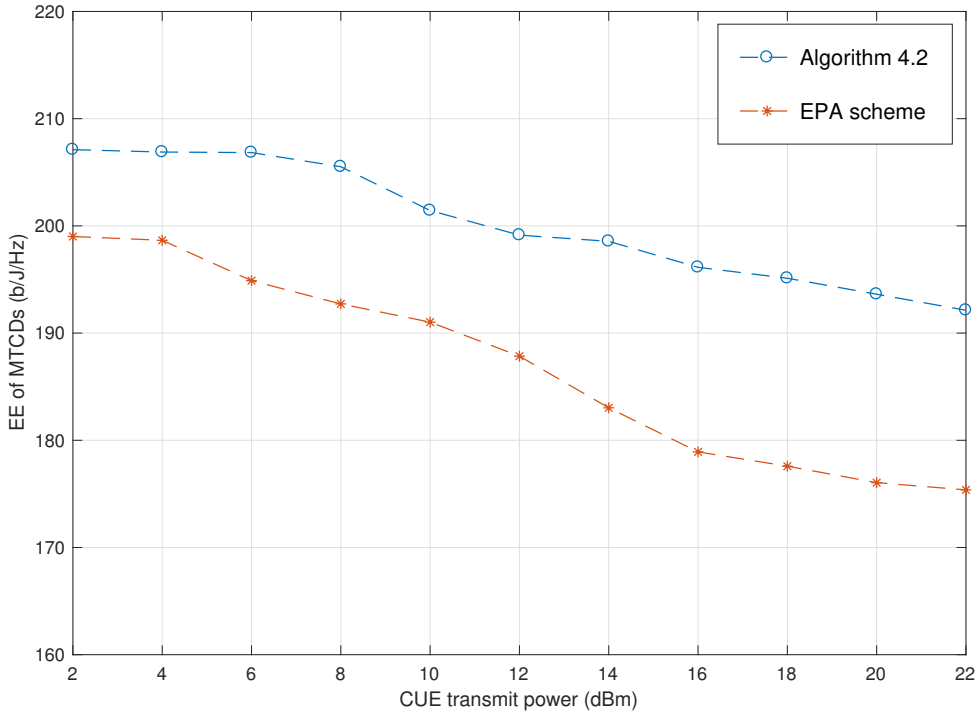

Figure 4.8: MTCD EE versus CUE transmit power with Algorithm 4.2 and the EPA scheme.

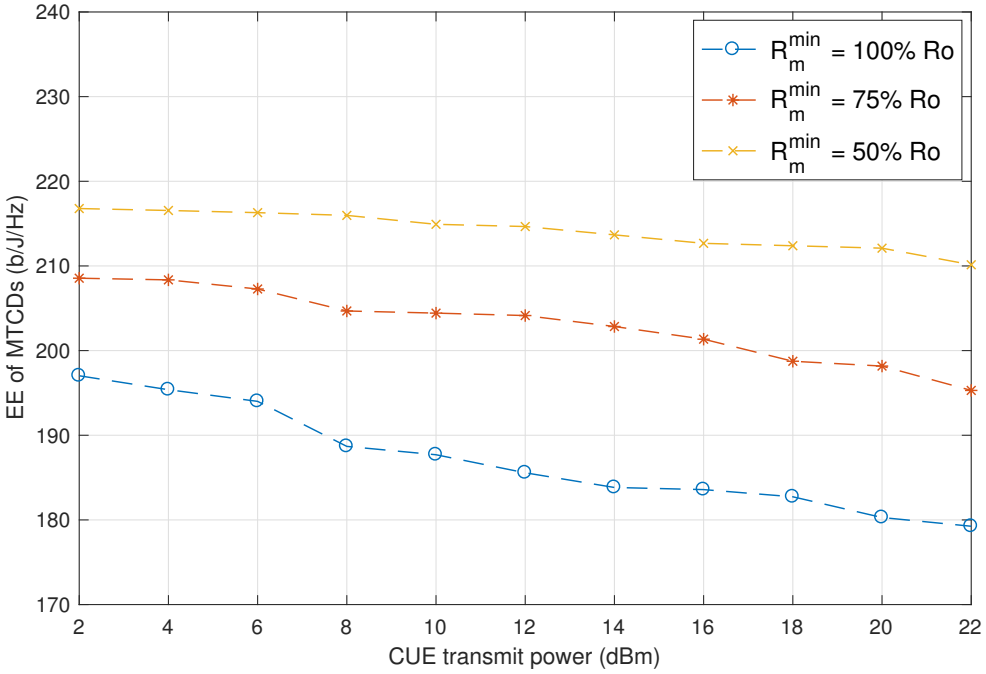

Figure 4.9: MTCD EE versus CUE transmit power for different minimum data rate thresholds, $R_m^{\min}$, with Algorithm 4.2.

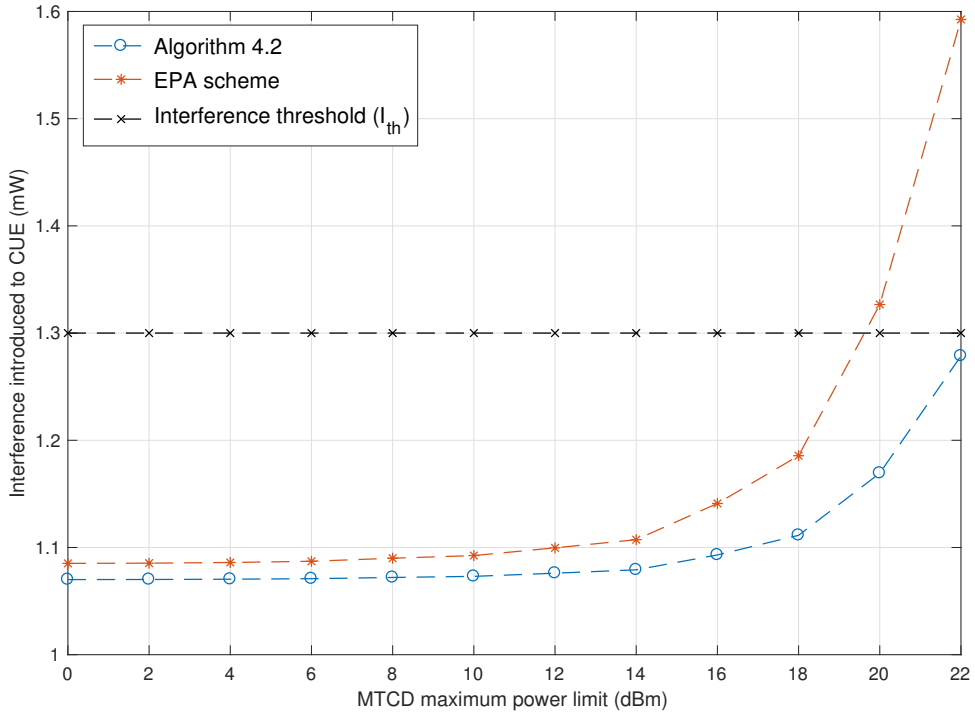

Figure 4.10: Interference introduced to the CUE versus the MTCD maximum power limit, $p_m^{\max}$, with Algorithm 4.2 and the EPA scheme.

maximum power limit $p_m^{\max}$ for Algorithm 4.2 and the EPA scheme. This shows that Algorithm 4.2 results in lower interference levels. This is because Algorithm 4.2 ensures that the interference introduced to the CUE does not exceed the interference threshold when allocating power to the underlaying MTCDs, whereas the EPA scheme allocates power to the MTCDs based on the maximum power limit regardless of the resulting interference.

4.6 Conclusion

In this chapter, energy-efficient CM2M communications underlaying cellular networks was considered. In this system, underlay CR is employed to manage the coexistence of MTCDs and CUE and exploit spatial spectrum opportunities to improve spectrum utilization. Two power allocation problems were proposed considering MTCD transmit power constraints, MTCD minimum data rate requirements, and CUE interference limits. The first problem targets MTCD power consumption minimization while the second addresses MTCD EE maximization. The proposed power consumption minimization problem was transformed into a Geometric Programming (GP) problem and solved iteratively. The proposed EE maximization problem is a non-convex fractional programming problem. Hence, a parametric transformation was used to convert it into an equivalent convex form and this was solved using an iterative approach. Simulation results were presented which indicate that the proposed power allocation algorithms are effective as they provide lower power consumption and higher EE compared with the EPA scheme while satisfying the constraints.

Chapter 5

Conclusion and Future Work

M2M communications is an emerging communication paradigm that provides ubiquitous connectivity between intelligent devices and the ability to communicate autonomously without direct human intervention. M2M communications acts as an enabling technology for the practical realization of the IoT. However, M2M communications differs from conventional H2H communications due to characteristics such as massive numbers of connected devices, small data transmissions, delay limits, and high energy efficiency requirements. These features create challenges for existing cellular networks which are primarily designed for H2H communications. These challenges include interference, congestion, spectrum scarcity and energy efficiency. This makes the process of radio resource allocation complex. Hence, efficient algorithms are required to implement M2M communications and meet their key QoS requirements.

In Chapter 2, a two-phase resource allocation algorithm for H2H/M2M coexistence in cellular networks was proposed. The goal is to meet the QoS requirements of H2H traffic and delay-sensitive M2M traffic while ensuring fairness for the delay-tolerant M2M traffic. Joint power-resource allocation was performed in the first phase in order to satisfy the QoS requirements of H2H traffic while considering the delay constraints of delay-sensitive M2M traffic. Then, the second phase focused on meeting the QoS requirements of both the delay-sensitive and delay-tolerant M2M traffic. Simulation results were presented which demonstrate that the proposed algorithm can balance the demands of M2M and H2H traffic, meet their diverse QoS requirements, and guarantee fairness for delay-tolerant M2M traffic.

Clustered Cognitive M2M (CM2M) communications underlaying cellular networks were considered in Chapter 3. Underlay CR was utilized to manage spectrum sharing between the clustered MTCDs and CUE while limiting the interference induced by the

MTCDs below an interference threshold. The goal is to maximize the uplink sum-rate of the neighbouring CUE and clustered MTCDs while satisfying interference, power, and minimum data rate constraints. A joint resource-power allocation problem was formulated and solved using a two-phase resource and power allocation scheme. In the first phase, Adaptive Resource Block Allocation (ARBA) was employed to allocate the available RBs to clustered MTCDs considering these constraints. In the second phase, an Iterative Power Allocation (IPA) algorithm was proposed to allocate the MTCD transmit power over the RBs allocated in the first phase. Simulation results were presented which show that the proposed scheme significantly improves the sum-rate of the network compared to other schemes while satisfying the constraints.

Energy-efficient CM2M communications underlaying cellular networks was considered in Chapter 4. In this system, underlay CR was employed to manage the coexistence of MTCDs and CUE and exploit spatial spectrum opportunities to improve spectrum utilization. Two power allocation problems were proposed considering MTCD transmit power constraints, MTCD minimum data rate requirements, and CUE interference limits. The first problem targets MTCD power consumption minimization while the second addresses MTCD EE maximization. Performance results were presented which indicate that the proposed power allocation algorithms are effective as they provide lower power consumption and higher EE compared to the Equal Power Allocation (EPA) scheme while satisfying the constraints.

This dissertation considered some of the key challenges for M2M communications in cellular networks and algorithms were developed to enable M2M communications. The primary focus was on the problem of resource allocation in order to accommodate the growing number of MTCDs and meet the diverse QoS requirements of coexisting M2M and H2H communications. A number of challenges remain that need to be addressed. The following section highlights some future research directions.

5.1 Future Work

5.1.1 Device-to-Device (D2D) Assisted M2M Communications

M2M communications in cellular networks can be realized in three scenarios [65] as shown in Figure 5.1. The first scenario is direct communications between MTCDs and the BS, similar to H2H UE communicating with the BS. In the second scenario,

MTCDs form a capillary network where communications with the BS is performed via an MTCG. The third scenario is peer-to-peer (P2P) communications between neighbouring MTCDs where MTCDs communicate directly with each other without using the BS to coordinate end-to-end communications between MTCDs. This type of communications is similar to D2D communications between cellular users.

The Third Generation Partnership Project (3GPP) introduced D2D communications to enable direct communications between cellular users in proximity [66]. D2D communications in cellular networks have several advantages such as offloading traffic, reducing communication delay, and increasing spectral efficiency [67]. Moreover, D2D communications is a suitable technology to support M2M communications due to the reduced power consumption. However, D2D communications impose new challenges in terms of resource allocation and interference management.

Resource allocation for D2D communications coexisting with H2H communications should be performed without degrading the performance of H2H UE. The majority of existing studies investigate D2D for H2H communications [69, 70, 71, 72, 110]. Therefore, it would be interesting to explore the potential of D2D communications for P2P communication between MTCDs. Towards this end, D2D communications can be implemented to increase the energy and spectral efficiency of MTCDs operating in P2P mode. To achieve this, resource allocation for H2H/M2M coexistence in cellular networks can be designed to manage resource sharing between H2H UE and the underlaying MTCDs communicating in P2P mode such that the QoS requirements of both H2H UE and MTCDs are satisfied.

5.1.2 Cooperative CM2M Communications

Cooperative communications is a promising technology in which users cooperate with each other for data delivery to combat issues such as signal fading due to multipath propagation in wireless channels [42]. The idea of cooperative communications comes from the relay model that consists of three terminals: a transmitter (Tx), a relay (R), and a receiver (Rx) where the relay forwards the signal received from the transmitter to the receiver. This allows users to relay signals to each other to mitigate the impact of multipath fading [73].

In cognitive networks, cooperative communications can be utilized by allowing SUs and/or PUs to use relaying. This can be realized in two paradigms: SU-SU cooperation and PU-SU cooperation, as shown in Figure 5.2. In SU-SU cooperation,

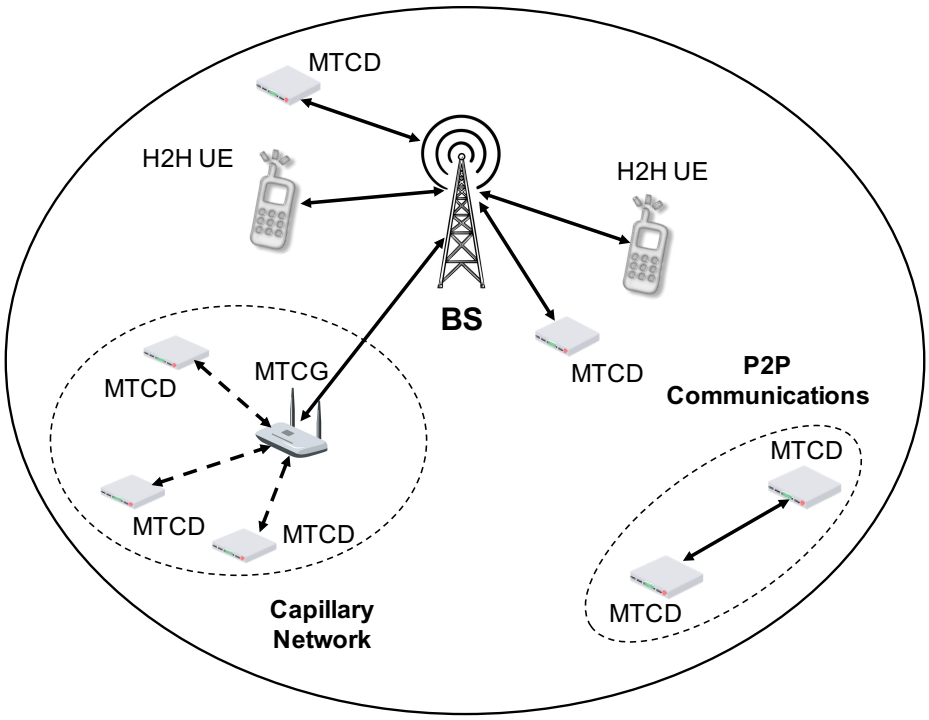

Figure 5.1: H2H/M2M coexistence scenarios in cellular networks.

SUs use cooperative communication to benefit from the channel gain obtained due to relaying [74]. Cooperation diversity in terms of spatial and time diversity can improve communication efficiency and combat the effect of fading. Moreover, relaying can reduce the power consumption of SUs as the transmit power of the source and relay can be less than that with no relaying. On the other hand, in PU-SU cooperation, PUs with weak transmission links can cooperate with SUs through relaying [75]. The PUs can reward these SUs with time or frequency resources.

Cooperative communications is also a promising technique for CM2M communications as then users can cooperate in data transmission [42] to improve network performance. It would be interesting to investigate cooperative communications for CM2M communications where the CUE are considered to be PUs and the MTCDs are SUs. MTCDs can utilize SU-SU cooperation through relaying to overcome interference from CUE, and meet their QoS requirements. Another possible direction is to implement PU-SU cooperation by allowing CUE to relay data through the MTCDs to improve the CUE performance. The spectrum efficiency will also be improved as cooperation with CUE can lead to more transmission opportunities for MTCDs [75].

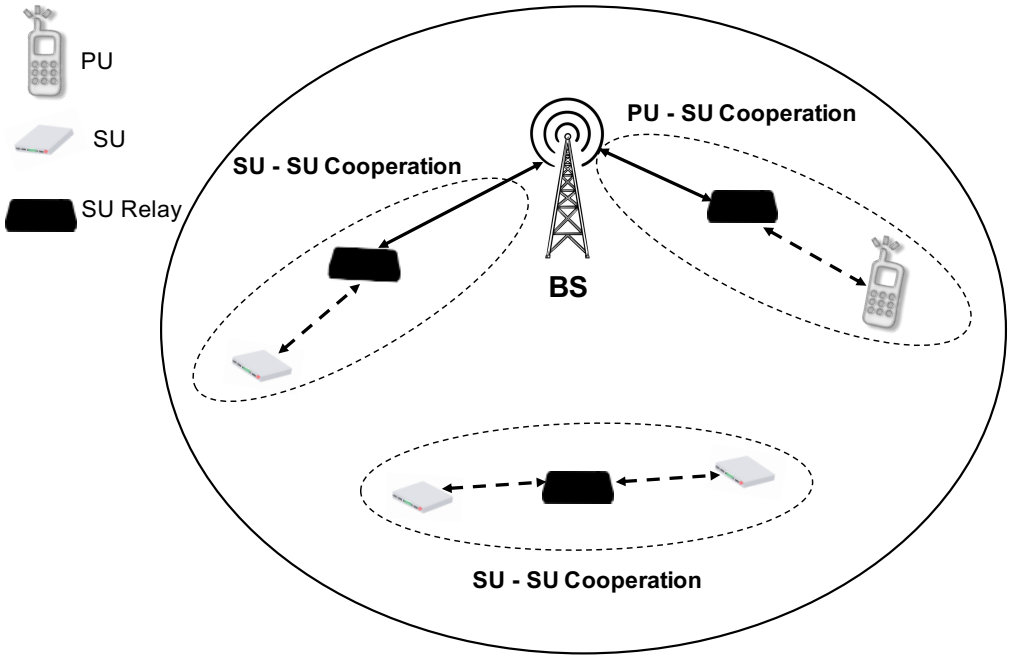

Figure 5.2: Cooperative CM2M communications.

5.1.3 Energy Harvesting for CM2M Communications

Energy Harvesting (EH) has emerged as an attractive technique to achieve energy efficiency and prolong the lifetime of battery-operated devices. Integrating EH with CM2M communications can improve both spectral and energy efficiencies in what is known as energy harvesting-based cognitive M2M (EH-CM2M) communications [110]. In EH-CM2M communications, MTCDs can harvest energy from the ambient environment. In particular, they can harvest energy from external energy sources, such as the sun, wind, and/or vibrations. Moreover, MTCDs can harvest energy from CUE Radio Frequency (RF) transmissions.

Since MTCDs in EH-CM2M communications are allowed to use the spectrum resources of CUE in an opportunistic manner and harvest energy from interference induced by CUE transmission, EH-CM2M communications can play a key role in achieving spectral and energy-efficient communications. However, resource allocation in EH-CM2M communications is crucial to achieving MTCD energy efficiency while fulfilling the QoS requirements of both CUE and MTCDs. Towards this end, the CM2M communications model proposed in Chapter 3 can be integrated with EH where MTCDs within each cluster harvest energy from the RF transmissions of neighbouring CUE. This model can be investigated from an energy efficiency per-

spective where resource allocation is performed with the objective of maximizing the total energy harvested by the MTCDs while considering the QoS requirements of the CUE and underlaying clustered MTCDs. Another possible direction is to design a joint resource-power allocation framework such that the sum-rate of the CUE and underlaying MTCDs is maximized considering the minimum data rate requirements of CUE and minimum energy harvesting threshold of MTCDs.

Appendix A

Geometric Programming Formulation of (2.19)

A standard form Geometric Programming (GP) problem can be expressed as the following optimization problem [34]

$$\min_{y} \ f_0(y) \tag{1.A}$$

$$\text{s.t.} \ \ f_i(y) \leq 1, \ i = 1, \cdots, m \tag{1.Aa}$$

$$g_i(y) = 1, \ i = 1, \cdots, p \, , \tag{1.Ab}$$

where f_0 is the objective function to be minimized which is a posynomial function with y as the optimization variable. The inequality constraint functions f_i are also posynomials, and the equality constraints functions g_i are monomials. A monomial is a function of the form

$$g(y) = c y_1^{\alpha_1} y_2^{\alpha_2} \cdots y_n^{\alpha_n}, \tag{2.A}$$

where the coefficient $c > 0$ and exponents $\alpha_1, \cdots, \alpha_n \in \mathbf{R}$. A posynomial function is a sum of monomials and is given by

$$f(y) = \sum_{k=1}^{K} g_k(y) = \sum_{k=1}^{K} c_k y_1^{\alpha_{1k}} y_2^{\alpha_{2k}} \cdots y_n^{\alpha_{nk}}. \tag{3.A}$$

Note that the reciprocal of a monomial is a monomial and any monomial is also a posynomial. The objective function in (2.19) is a monomial of the variable $\gamma_{i,n}$ with coefficient $c = 1$ and exponents $-x_{i,n}$. Hence, it is also a posynomial. This also holds

for constraints (2.19a) and (2.19b) which are monomials of $\gamma_{i,n}$ with coefficients $1/a_i$ and $1/b_i$, respectively, and exponents $-x_{i,n}$. Therefore, they are both posynomials and (2.19) is a standard form GP problem.

Appendix B

Derivation of Constraint (3.9e)

The arrival process of event-triggered M2M communications is modeled as a Poisson process. Hence, each MTCD can be modeled as an M/D/1 queue with arrival rate λ_{m_c} and service rate R_{m_c}. These queues are considered simultaneously during the resource allocation process [59]. With an M/D/1 queue, the wait time of MTCD m_c is given by

$$\Pr(W_{m_c} \leq t_{m_c}) = (1 - \lambda_{m_c}\tau_{m_c}) \sum_{v=0}^{z} \frac{[-\lambda_{m_c}(t_{m_c} - v\tau_{m_c})]^v}{v!} e^{\lambda_{m_c}(t_{m_c} - v\tau_{m_c})}, \tag{1.B}$$

where z is an integer satisfying $z\tau_{m_c} \leq t_{m_c} \leq (z+1)\tau_{m_c}$ and $\tau_{m_c} = 1/R_{m_c}$ is the deterministic service time. The total delay of MTCD m_c can be expressed as

$$d_{m_c} = W_{m_c} + \tau_{m_c}. \tag{2.B}$$

Replacing t_{m_c} by ($d_{m_c}^{\max}$ - τ_{m_c}) in (1.B), the Probability of Delay-Bound Violation (PDBV) of MTCD m_c can be expressed as

$$\Pr(d_{m_c} > d_{m_c}^{\max}) = 1 - (1 - \lambda_{m_c}\tau_{m_c}) \sum_{v=0}^{z} \frac{[-\lambda_i(d_{m_c}^{\max} - \tau_{m_c} - v\tau_{m_c})]^v}{v!} e^{\lambda_{m_c}(d_k^{\max} - \tau_{m_c} - v\tau_{m_c})}. \tag{3.B}$$

Let $\varepsilon_{m_c}^{\max}$ be the maximum delay-violation threshold of MTCD m_c and $R^{\min}(d_{m_c}^{\max}, \varepsilon_{m_c}^{\max})$ the minimum data rate of MTCD m_c that achieves a PDBV equal to $\varepsilon_{m_c}^{\max}$ when the delay-bound is $d_{m_c}^{\max}$.

A guarantee of an acceptable Quality-of-Service (QoS) of MTCD m_c requires that

$$\Pr(d_{m_c} > d_{m_c}^{\max}) \leq \varepsilon_{m_c}^{\max}. \tag{4.B}$$

This implies that

$$R_{m_c} \geq R^{\min}(d_{m_c}^{\max}, \varepsilon_{m_c}^{\max}), \tag{5.B}$$

which is equivalent to constraint (3.9e).

Appendix C

Derivation of Constraint (3.19d)

From Chapter 3, recall constraint (3.9e) where

$$\sum_{n=1}^{N} x_{m_c}^n R_{m_c}^n \geq R^{\min}(d_{m_c}^{\max}, \varepsilon_{m_c}^{\max}), \ \ \forall\, m_c \in \mathcal{M}_c. \tag{1.C}$$

After the resource allocation is completed, the RB allocation indicator $x_{m_c}^n = 1,\ \forall m_c \in \mathcal{M}_c, n \in \mathcal{N}_{m_c}$. Hence, (1.C) can be expressed as

$$\sum_{n=1}^{N_{m_c}} R_{m_c}^n = R_{m_c} \geq R^{\min}(d_{m_c}^{\max}, \varepsilon_{m_c}^{\max}), \ \ \forall\, m_c \in \mathcal{M}_c, \tag{2.C}$$

where

$$R_{m_c} = \log_2(1 + \gamma_{m_c}), \ \ \forall\, m_c \in \mathcal{M}_c. \tag{3.C}$$

This is equivalent to

$$\gamma_{m_c} \geq (2^{R^{\min}(d_{m_c}^{\max}, \varepsilon_{m_c}^{\max})} - 1), \ \ \forall\, m_c \in \mathcal{M}_c. \tag{4.C}$$

According to [60], SC-FDMA transmission spreads each data symbol over the entire bandwidth, so the SINR per subcarrier is not directly related to the data symbol. Hence, the SINR with SC-FDMA cannot be approximated using the models for Orthogonal Frequency-Division Multiple Access (OFDMA) such as Exponential Effective SNR Mapping (EESM) or Mutual Information Effective SINR Mapping (MIESM) [61]. Therefore, even though the channel gain experienced by each subcarrier is different, the channel gain differences averaged out over a sufficiently large bandwidth. Hence, the SINR of MTCD m_c can be computed as the average SINR over a set of

allocated RBs $\mathcal{N}_{m_c}$ and is given by

$$\gamma_{m_c} = \frac{1}{N_{m_c}} \sum_{n=1}^{N_{m_c}} \gamma_{m_c,g_c}^{n}, \quad \forall\, m_c \in \mathcal{M}_c. \tag{5.C}$$

Hence, (4.C) can be expressed as

$$\sum_{n=1}^{N_{m_c}} \gamma_{m_c,g_c}^{n} = \sum_{n=1}^{N_{m_c}} p_{m_c}^{n} \Gamma_{m_c,g_c}^{n} \geq N_{m_c}\, (2^{R^{\min}(d_{m_c}^{\max}, \varepsilon_{m_c}^{\max})} - 1), \tag{6.C}$$

which gives constraint (3.19d).

Appendix D

Proof of Proposition 3.1

Since $G^n(p^n_{m_c})$ is concave, based on the first-order condition for a concave function, we have

$$G^n(p^n_{m_c}) \leq G^n(\hat{p}^n_{m_c}) - \langle \nabla G^n(\hat{p}^n_{m_c}), p^n_{m_c} - \hat{p}^n_{m_c} \rangle. \tag{1.D}$$

Hence

$$R^n(p^n_{m_c}) \geq \underline{\mathrm{R}}^n(p^n_{m_c}, \hat{p}^n_{m_c}). \tag{2.D}$$

Moreover, when $\hat{p}^n_{m_c} = p^n_{m_c}$

$$R^n(p^n_{m_c}) = \underline{\mathrm{R}}^n(p^n_{m_c}, p^n_{m_c}), \tag{3.D}$$

which implies that $\underline{\mathrm{R}}^n(p^n_{m_c}, \hat{p}^n_{m_c})$ in (3.22) is a tight lower bound on $R^n(p^n_{m_c})$ in (3.19).

Appendix E

Proof of Proposition 3.2

Based on Proposition 3.1, we have

$$\sum_{n=1}^{N} R^n(p_{m_c}^{*n}) \overset{(a)}{\geq} \sum_{n=1}^{N} \underline{\mathrm{R}}^n(p_{m_c}^{*n}, \hat{p}_{m_c}^n) \overset{(b)}{\geq} \sum_{n=1}^{N} \underline{\mathrm{R}}^n(\hat{p}_{m_c}^n, \hat{p}_{m_c}^n) \overset{(c)}{=} \sum_{n=1}^{N} R^n(\hat{p}_{m_c}^n), \tag{1.E}$$

where (a) follows from the fact that $\underline{\mathrm{R}}^n(p_{m_c}^{*n}, \hat{p}_{m_c}^n)$ is a lower bound for $R^n(p_{m_c}^{*n})$, (b) follows from the fact that $p_{m_c}^{*n}$ is the optimal solution to (3.22), and (c) follows from (3.D).

Appendix F

Proof of Proposition 3.3

Since $G^n(p^n_{m_c})$ is concave, its gradient $\nabla G^n(\hat{p}^n_{m_c})$ is also its supergradient [62]. At a given iteration l

$$G^n(p^{n(l)}_{m_c}) \leq G^n(p^{n(l-1)}_{m_c}) - \langle \nabla G^n(p^{n(l-1)}_{m_c}), p^{n(l)}_{m_c} - p^{n(l-1)}_{m_c} \rangle. \tag{1.F}$$

Moreover, we have

$$\begin{aligned} R^n(p^{n(l)}_{m_c}) &= F^n(p^{n(l)}_{m_c}) - G^n(p^{n(l)}_{m_c}) \\ &\overset{(a)}{\geq} F^n(p^{n(l)}_{m_c}) - G^n(p^{n(l-1)}_{m_c}) - \langle \nabla G^n(p^{n(l-1)}_{m_c}), p^{n(l)}_{m_c} - p^{n(l-1)}_{m_c} \rangle \\ &\overset{(b)}{\geq} F^n(p^{n(l-1)}_{m_c}) - G^n(p^{n(l-1)}_{m_c}) - \langle \nabla G^n(p^{n(l-1)}_{m_c}), p^{n(l-1)} - p^{n(l-1)}_{m_c} \rangle \\ &\overset{(c)}{=} F^n(p^{n(l-1)}_{m_c}) - G^n(p^{n(l-1)}_{m_c}) = R^n(p^{n(l-1)}_{m_c}), \end{aligned} \tag{2.F}$$

where (a) follows from (2.D), and (b) and (c) follow from (1.E). Hence, the next solution $p^{n(l)}_{m_c}$ is never worse than the previous solution $p^{n(l-1)}_{m_c}$, so the objective function in (3.19) will improve. Moreover, since the constraint set is compact, the IPA algorithm will converge to a local optimal solution $p^{*n}_{m_c}$ of (3.19).

Appendix G

Proof of (4.13)

The solution to (4.12) must satisfy constraint (4.12d) with equality, that is

$$\log_2\left(1+\frac{p_m^n|h_{m,b}^n|^2}{p_u^n|h_{u,b}^n|^2+\sum_{m'\neq m}p_{m'}^n|h_{m',b}^n|^2+\sigma_2}\right)=R_m^{\min}. \tag{1.G}$$

Therefore, the power of the mth MTCD over the nth RB can be expressed as

$$p_m^n=\frac{p_u^n|h_{u,b}^n|^2+\sum_{m'\neq m}p_{m'}^n|h_{m',b}^n|^2+\sigma_2}{|h_{m,b}^n|^2}\left(2^{R_m^{\min}}-1\right). \tag{2.G}$$

Moreover, the power of the mth MTCD over the nth RB must satisfy the minimum and maximum power constraints, that is

$$p_m^{\min}\leq p_m^n\leq p_m^{\max}. \tag{3.G}$$

Define

$$P_m^{\max}\triangleq\min\left\{p_m^{\max},\frac{p_u^n|h_{u,b}^n|^2+\sum_{m'\neq m}p_{m'}^n|h_{m',b}^n|^2+\sigma_2}{|h_{m,b}^n|^2}\left(2^{R_m^{\min}}-1\right)\right\}. \tag{4.G}$$

From (3.G) and (4.G), we have

$$p_m^{\min}\leq p_m^n\leq P_m^{\max}. \tag{5.G}$$

Now, using (5.G) and [116], we have

$$p_m^n(l+1) = \max\left\{p_m^{\min}, \min\left[p_m^{\max}, \frac{2^{R_m^{\min}}-1}{\gamma_m^n(l)} p_m^n(l)\right]\right\}. \tag{6.G}$$

Bibliography

[1] INFSO D.4 Networked Enterprise & RFID INFSO G.2 Micro & Nanosystems, in: co-operation with the working group RFID of the ETP EPOSS, "Internet of Things, roadmap for the future," version 1.1, 27 May 2008.

[2] Cisco annual internet report (2018-2023) white paper, Mar. 2020. Accessed: Jul. 1, 2020. Available [Online]: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.pdf

[3] G. Wu, S. Talwar, K. Johnsson, N. Himayat, and K. D. Johnson, "M2M: From mobile to embedded internet," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 36–43, Apr. 2011.

[4] D. Boswarthick, O. Elloumi, and O. Hersent, "M2M communications: A systems approach," John Wiley, 2012.

[5] T. Taleb and A. Kunz, "Machine type communications in 3GPP networks: Potential, challenges, and solutions," *IEEE Commun. Mag.*, vol. 50, no. 3, pp. 178–184, Mar. 2012.

[6] N. Xia, H. H. Chen, and C. S. Yang, "Radio resource management in machine-to-machine communications – A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 791–828, 1st Qtr. 2017.

[7] Y. Zhang, R. Yu, S. Xie, W. Yao, Y. Xiao, and M. Guizani, "Home M2M networks: architectures, standards, and QoS improvement," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 44–52, Apr. 2011.

[8] I. Yaqoob, E. Ahmed, A. T. Hashem, A. I. A. Ahmed, A. Gani, M. Imran, and M. Guizani, "Internet of things architecture: Recent advances, taxonomy,

requirements, and open challenges," *IEEE Wireless Commun.*, vol. 24, no. 3, pp. 10–16, Jun. 2017.

[9] A. Amokrane, "Congestion control in the context of machine type communications in long term evolution networks," M.S. thesis, Univ. Rennes I, Rennes, France, Oct. 2011.

[10] A. M. Maia, D. Vieira, M. F. de Castro, and Y. Ghamri-Doudane, "A mechanism for uplink packet scheduler in LTE network in the context of machine-to-machine communication," in *Proc. IEEE Global Commun. Conf.*, Austin, TX, pp. 2776–2782, Dec. 2014.

[11] M. Chen, J. Wan, and F. Li, "Machine-to-machine communications: Architectures, standards and applications," *KSII transactions on Internet and Information Systems*, vol. 6, no. 2, pp. 480–497, Feb. 2012.

[12] R. Lu, X. Li, X. Liang, X. Shen, and X. Lin, "GRS: The green, reliability, and security of emerging machine to machine communications," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 28–35, Apr. 2011.

[13] K. Chang, A. Soong, M. Tseng, and Z. Xiang, "Global wireless machine-to-machine standardization," *IEEE Internet Computing*, vol. 15, no. 2, pp. 64–69, Mar. 2011.

[14] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. on Sel. Areas in Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[15] M. Nekovee, "Quantifying the availability of TV white spaces for cognitive radio operation in the UK," in *IEEE Int. Conf. on Commun. Workshops*, Dresden, Germany, pp. 1–5, Jun. 2009.

[16] J. Palicot, "Cognitive radio: an enabling technology for the green radio communications concept," in *Proc. ACM Int. Wireless Commun. and Mobile Computing Conf.*, Leipzig, Germany, pp. 489–494, Jun. 2009.

[17] A. Elhamy, and Y. Gadallah, "BAT: A balanced alternating technique for M2M uplink scheduling over LTE," in *Proc. IEEE Vehic. Technol. Conf.*, Glasgow, UK, May 2015.

[18] Z. Yang, Y. Pan, W. Xu, R. Guan, Y. Wang, and M. Chen, "Energy efficient resource allocation for machine-to-machine communications with NOMA and energy harvesting," in *Proc. IEEE Conf. on Comp. Commun. Workshops*, Atlanta, GA, pp. 145–150, May 2017.

[19] M. R. Mardani, S. Mohebi, B. Maham, and M. Bennis, "Delay-sensitive resource allocation for relay-aided M2M communication over LTE-advanced networks," in *Proc. IEEE Symp. on Comp. and Commun.*, Heraklion, Greece, pp. 1033–1038, Jul. 2017.

[20] F. Ghavimi, Y. W. Lu, and H. H. Chen, "Uplink scheduling and power allocation for M2M communications in SC-FDMA-based LTE-A networks with QoS guarantees," *IEEE Trans. Vehic. Technol.*, vol. 66, no. 7, pp. 6160–6170, Jul. 2016.

[21] J. Brown and J. Y. Khan, "A predictive resource allocation algorithm in the LTE uplink for event based M2M applications," *IEEE Trans. Mobile Computing*, vol. 14, no. 12, pp. 2433–2446, Dec. 2015.

[22] A. Ali, G. A. Shah, and J. Arshad, "Energy efficient resource allocation for M2M devices in 5G," *Sensors*, vol. 19, no. 8, art. 1830, Apr. 2019.

[23] G. Bartoli, R. Fantacci, and D. Marabissi, "Resource allocation approaches for two-tiers machine-to-machine communications in an interference limited environment," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 9112–9122, Oct. 2019.

[24] Y. Wu, N. Zhang, and K. Rong, "Non-orthogonal random access and data transmission scheme for machine-to-machine communications in cellular networks," *IEEE Access*, vol. 8, pp. 27687–27704, Feb. 2020.

[25] A. M. Atchome, H. H. R. Sherazi, R. Alahassa, F. Tossa, T. O. Edoh, L. A. Grieco, and A. C. Vianou, "A Hybrid network nodel embracing NB-IoT and D2D communications: Stochastic geometry analysis," in *e-Infrastructure and e-Services for Developing Countries*, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 331, pp. 154–165, Springer, Berlin, Dec. 2019.

[26] A. Aijaz, M. Tshangini, M. R. Nakhai, X. Chu, and A.-H. Aghvami, "Energy-efficient uplink resource allocation in LTE networks with M2M/H2H co-existence

under statistical QoS guarantees," *IEEE Trans. Commun.*, vol. 62, no. 7, pp. 2353–2365, Jul. 2014.

[27] A. G. Gotsis, A. S. Lioumpas, and A. Alexiou, "M2M scheduling over LTE: challenges and new perspectives," *IEEE Vehic. Technol. Mag.*, vol. 7, no. 3, pp. 34–39, Sep. 2012.

[28] S. Zhenqi, Y. Haifeng, C. Xuefen, and L. Hongxia, "Research on uplink scheduling algorithm of massive M2M and H2H services in LTE," in *Proc. IET Int. Conf. on Inform. and Commun. Technol.*, Beijing, China, pp. 365–369, Apr. 2013.

[29] 3GPP TR 25.814 Physical layer aspects for evolved Universal Terrestrial Radio Access (UTRA), v.7.0.0, Jun. 2006.

[30] 3GPP R1 120056, Analysis on traffic model and characteristics for MTC and text proposal, Tech. Rep. 3GPP TSG RAN Meeting WG1 68, R1-120056. Dresden, Germany, Feb. 2012.

[31] A. G. Gotsis, A. S. Lioumpas, and A. Alexiou, "Analytical modelling and performance evaluation of realistic time-controlled M2M scheduling over LTE cellular networks," *Trans. Emerging Telecommun. Technologies*, vol. 24, no. 4, pp. 378–388, Jun. 2013.

[32] A. K. Erlang, "The theory of probabilities and telephone conversations," *Nyt. Tidsskr. Mat. Ser. B*, vol. 20, pp. 33–39, 1909.

[33] J. C. Bezdek and R. J. Hathaway, "Some notes on alternating optimization," in *Advances in Soft Computing-AFSS 2002*, Lecture Notes in Computer Science, vol. 2275, pp. 288–300, Feb. 2002.

[34] M. Chiang, *Geometric Programming for Communication Systems*, Now Publishers, Hanover, MA, 2005.

[35] M. Chiang, C. W. Tan, D. P. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, Jul. 2007.

[36] M. Grant, S. Boyd, and Y. Ye, "CVX users' guide," 2009. Available [Online]*: http://www.stanford.edu/˜boyd/software.html*

[37] C. Kai, L. Xu, J. Zhang, and M. Peng, "Joint uplink and downlink resource allocation for D2D communication underlying cellular networks," in *Proc. IEEE Int. Conf. on Wireless Commun. and Signal Process.*, Hangzhou, China, Oct. 2018.

[38] M. Y. Abdelsadek, Y. Gadallah, and M. H. Ahmed, "An LTE-based optimal resource allocation scheme for delay-sensitive M2M deployments coexistent with H2H users," in *Proc. IEEE Conf. on Comp. Commun. Workshops*, Atlanta, GA, pp. 139–144, May 2017.

[39] V. Kouhdaragh, D. Tarchi, A. V. Coralli, and G. E. Corazza, "Cognitive radio based smart grid networks," in *Tyrrhenian Int. Workshop on Digital Communications-Green ICT*, Sep. 2013.

[40] A. Patel, M. Z. A. Khan, S. N. Merchant, U. B. Desai, and L. Hanzo, "The achievable rate of interweave cognitive radio in the face of sensing errors," *IEEE Access*, vol. 5, pp. 8579–8605, Sep. 2016.

[41] L. Liang, L. Xu, B. Cao, and Y. Jia, "A cluster-based congestion-mitigating access scheme for massive M2M communications in Internet of Things," *IEEE Internet of Things J.*, vol. 5, no. 3, pp. 2200–2211, Apr. 2018.

[42] P. Rawat, K. D. Singh, and J. M. Bonnin, "Cognitive radio for M2M and Internet of Things: A survey," *Comp. Commun.*, vol. 94, pp. 1–29, Nov. 2016.

[43] N. Abu-Ali, A.-E. M. Taha, M. Salah, and H. Hassanein, "Uplink scheduling in LTE and LTE-advanced: Tutorial, survey and evaluation framework," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1239–1265, 3rd Qtr. 2014.

[44] E. Soltanmohammadi, K. Ghavami, and M. Naraghi-Pour, "A survey of traffic issues in machine-to-machine communications over LTE," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 865–884, Feb. 2016.

[45] M. A. Shah, S. Zhang, and C. Maple, "Cognitive radio networks for Internet of Things: Applications, challenges and future," in *Proc. IEEE Int. Conf. on Automation and Computing*, Sep. 2013.

[46] K. Shuaib, E. Barka, N. Al Hussien, M. Abdel-Hafez, and M. Alahmad, "Cognitive radio for smart grid with security considerations," *Computers*, vol. 5, no. 2, Apr. 2016.

[47] K. D. Singh, P. Rawat, and J. M. Bonnin, "Cognitive radio for vehicular ad hoc networks (CR-VANETs): approaches and challenges," *EURASIP J. on Wireless Commun. and Net.*, vol. 2014, no. 1, pp. 1–22, Mar. 2014.

[48] M. Hasan, E. Hossain, and D. Niyato, "Random access for machine-to-machine communication in LTE-advanced networks: Issues and approaches," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 86–93, Jun. 2013.

[49] Y. Zhang, R. Yu, M. Nekovee, Y. Liu, S. Xie, and S. Gjessing, "Cognitive machine-to-machine communications: Visions and potentials for the smart grid," *IEEE Net.*, vol. 26, no. 3, pp. 6–13, May 2012.

[50] K. R. Jung, A. Park, and S. Lee, "Machine-type-communication (MTC) device grouping algorithm for congestion avoidance of MTC oriented LTE network," in *Proc. Int. Conf. on Security-Enriched Urban Computing and Smart Grid*, Springer, Heidelberg, Germany, pp. 167–178, Sep. 2010.

[51] S. H. Wang, H. J. Su, H. Y. Hsieh, S. P. Yeh, and M. Ho, "Random access design for clustered wireless machine to machine networks," in *Proc. Int. Black Sea Conf. on Commun. and Networking*, Batumi, Georgia, pp. 107–111, Jul. 2013.

[52] M. Darabi, B. Maham, W. Saad, A. Mehbodniya, and F. Adachi, "Joint machine-type device selection and power allocation for buffer-aided cognitive M2M communication," in *Proc. IEEE Int. Symp. on Personal, Indoor, and Mobile Radio Commun.*, Hong Kong, China, pp. 2288–2292, Sep. 2015.

[53] W. Ejaz, and M. Ibnkahla, "Machine-to-machine communications in cognitive cellular systems," in *Proc. IEEE Int. Conf. on Ubiquitous Wireless Broadband*, Montreal, QC, Canada, pp. 1–5, Oct. 2015.

[54] J. Yin, Y. Chen, G. Sang, B. Liao, and X. Wang, "QoE-oriented rate control and resource allocation for cognitive M2M communication in spectrum-sharing OFDM networks," *IEEE Access*, vol. 7, pp. 43318–43330, Apr. 2019.

[55] D. T. Hoang, and D. Niyato, "Performance analysis of cognitive machine-to-machine communications," in *Proc. IEEE Int. Conf. on Commun. Systems*, Singapore, pp. 245–249, Nov. 2012.

[56] H. K. Lee, D. M. Kim, Y. Hwang, S. M. Yu, and S. L. Kim, "Feasibility of cognitive machine-to-machine communication using cellular bands," *IEEE Wireless Commun.*, vol. 20, no. 2, pp. 97–103, Feb. 2013.

[57] M. A. Abdullah, Z. Abdullah, G. Chen, J. Tang, and J. Chambers, "Performance analysis of cognitive clustered M2M random networks with joint user and machine device selection," *IEEE Access*, vol. 7, pp. 83515–83525, Jun. 2019.

[58] G. Chen and J. A. Chambers, "Exact outage probability analysis for cooperative AF relay network with relay selection in presence of inter-cell interference," *Elect. Lett.*, vol. 48, no. 21, pp. 1346–1347, Oct. 2012.

[59] N. Alhussien and T. A. Gulliver, "Optimal resource allocation in cellular networks with H2H/M2M coexistence," *IEEE Trans. Vehic. Techol.*, vol. 69, no. 11, pp. 12951–12962, Nov. 2020.

[60] F. Calabrese, "Scheduling and link adaptation for uplink SC-FDMA systems," Ph.D. dissertation, Aalborg University, Aalborg, Denmark, Apr. 2009.

[61] R. Giuliano and F. Mazzenga, "Exponential effective SINR approximations for OFDM/OFDMA-based cellular system planning," *IEEE Trans. Wireless Commun.*, vol. 8, no. 9, pp. 4434–4439, Sep. 2009.

[62] H. H. Kha, H. D. Tuan, and H. H. Nguyen, "Fast global optimal power allocation in wireless networks by local DC programming," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 510–515, Feb. 2012.

[63] I. Stojmenovic, "Localized network layer protocols in wireless sensor networks based on optimizing cost over progress ratio," *IEEE Net.*, vol. 20, no. 1, pp. 21–27, Jan, 2006.

[64] A. Gallais, J. Carle, D. Simplot-Ryl, and I. Stojmenovic, "Localized sensor area coverage with low communication overhead," *IEEE Trans. Mobile Computing*, vol. 7, no. 5, pp. 661–672, Mar. 2008.

[65] K. Zheng, F. Hu, W. Wang, W. Xiang, and M. Dohler, "Radio resource allocation in LTE-advanced cellular networks with M2M communications," *IEEE Commun. Mag.*, vol. 50, no. 7, pp. 184–192, Jul. 2012.

[66] D. Astely, E. Dahlman, G. Fodor, S. Parkvall, and J. Sachs, " LTE release 12 and beyond," *IEEE Commun. Mag.*, vol. 51, no. 7, pp. 154–160, Jul. 2013.

[67] K. Doppler, M. Rinne, C. Wijting, C. B. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 42–49, Dec. 2009.

[68] Z. Zhou, M. Dong, K. Ota, J. Wu, and T. Sato, "Energy efficiency and spectral efficiency tradeoff in device-to-device (D2D) communications," *IEEE Wireless Commun. Let.*, vol. 3, no. 5, pp. 485–488, Oct. 2014.

[69] R. Yin, C. Zhong, G. Yu, Z. Zhang, K. K. Wong, and X. Chen, "Joint spectrum and power allocation for D2D communications underlaying cellular networks," *IEEE Tran. on Vehic. Technol.*, vol. 65, no. 4, pp. 2182–2195, Apr. 2015.

[70] L. Wei, R. Q. Hu, Y. Qian, and G. Wu, "Enable device-to-device communications underlaying cellular networks: Challenges and research aspects," *IEEE Commun. Mag.*, vol. 52, no. 6, pp. 90–96, Jun. 2014.

[71] T. D. Hoang, L. B. Le, and T. Le-Ngoc, "Energy-efficient resource allocation for D2D communications in cellular networks," *IEEE Trans. on Vehic. Technol.*, vol. 65, no. 9, pp. 6972–6986, Sep. 2015.

[72] J. Hu, W. Heng, X. Li, and J. Wu, "Energy-efficient resource reuse scheme for D2D communications underlaying cellular networks," *IEEE Commun. Lett.*, vol. 21, no. 9, pp. 2097–2100, Jun.2017.

[73] A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 68–73, Oct. 2004.

[74] Q. Zhang, J. Jia, J. Zhang, "Cooperative relay to improve diversity in cognitive radio networks," *IEEE Commun. Mag.*, vol. 47, no. 2, pp. 111–117, Feb. 2009.

[75] T. Nadkar, V. Thumar, G. Shenoy, A. Mehta, U. B. Desai, and S. N. Merchant, "A cross-layer framework for symbiotic relaying in cognitive radio networks," in *Proc. IEEE Int. Symp. on Dynamic Spectrum Access Networks*, Aachen, Germany, pp. 498–509, May 2011.

[76] Z. Zhou, C. Zhang, J. Wang, B. Gu, S. Mumtaz, J. Rodriguez, and X. Zhao, "Energy-efficient resource allocation for energy harvesting-based cognitive machine-to-machine communications," *IEEE Trans. Cognitive Commun. and Net.*, vol. 5, no. 3, pp. 595–607, Jun. 2019.

[77] A. Aijaz and A. H. Aghvami, "Cognitive machine-to-machine communications for Internet-of-Things: A protocol stack perspective," *IEEE Internet Things J.*, vol. 2, no. 2, pp. 103–112, Apr. 2015.

[78] X. Zhang, X. Zhang, L. Han, and R. Xing, "Utilization-oriented spectrum allocation in an underlay cognitive radio network," *IEEE Access*, vol. 6, pp. 12905–12912, Feb. 2018.

[79] Y. Cao, N. Zhao, F. R. Yu, M. Jin, Y. Chen, J. Tang, and V. C. Leung, "Optimization or alignment: Secure primary transmission assisted by secondary networks," *IEEE J. Select. Areas Commun.*, vol. 36, no. 4, pp. 905–917, Apr. 2018.

[80] B. Gu, C. Zhang, H. Wang, Y. Yao, and X. Tan, "Power control for cognitive M2M communications underlaying cellular with fairness concerns," *IEEE Access*, vol. 7, pp. 80789–80799, Apr. 2019.

[81] H. Liao, X. Chen, Z. Zhou, N. Liu, and B. Ai, "Licensed and unlicensed spectrum management for cognitive M2M: A context-aware learning approach," *IEEE Trans. Cognitive Commun. Networking*, vol. 6, no. 3, pp. 915–925, Sep. 2020.

[82] P. Apkarian, and H. D. Tuan, "Concave programming in control theory," *J. Global Optim.*, vol. 15, no. 4, pp. 343–370, Dec. 1999.

[83] M. El Tanab, and W. Hamouda, "Resource allocation for underlay cognitive radio networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 1249–1276, 2nd Qtr. 2017.

[84] A. Fehske, G. Fettweis, J. Malmodin, and G. Biczok, "The global footprint of mobile communications: The ecological and economic perspective," *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 55–62, Aug. 2011.

[85] K. Davaslioglu, and E. Ayanoglu, "Quantifying potential energy efficiency gain in green cellular wireless networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 2065–2091, 4th Qtr. 2014.

[86] Cisco annual internet report (2018-2023) white paper, Mar. 2020. Accessed: Jul. 1, 2020. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.pdf.

[87] M. Hasan, E. Hossain, and D. Niyato, "Random access for machine-to-machine communication in LTE-advanced networks: Issues and approaches," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 86–93, Jun. 2013.

[88] A. M. Maia, D. Vieira, M. F. de Castro, and Y. Ghamri-Doudane, "A mechanism for uplink packet scheduler in LTE network in the context of machine-to-machine communication," in *Proc. IEEE Global Commun. Conf.*, Austin, TX, pp. 2776–2782, Dec. 2014.

[89] A. Zappone, B. Matthiesen, and E. A. Jorswieck, "Energy efficiency in MIMO underlay and overlay device-to-device communications and cognitive radio systems," *IEEE Trans. Signal Process.*, vol. 65, no. 4, pp. 1026–1041, Feb. 2017.

[90] N. Xia, H.-H. Chen, and C.-S. Yang, "Radio resource management in machine-to-machine communications–A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 791–828, 1st Qtr. 2018.

[91] D. Feng, C. Jiang, G. Lim, L. J. Cimini, G. Feng, and G. Y. Li, "A survey of energy-efficient wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 167–178, 1st Qtr. 2013.

[92] Y. Zhang, R. Yu, M. Nekovee, Y. Liu, S. Xie, and S. Gjessing, "Cognitive machine-to-machine communications: Visions and potentials for the smart grid," *IEEE Network*, vol. 26, no. 3, pp. 6–13, May-Jun. 2012.

[93] J. Marinho, and E. Monteiro, "Cognitive radio: Survey on communication protocols, spectrum decision issues, and future research directions," *Wireless Networks*, vol. 18, no. 2, pp. 147–164, Feb. 2012.

[94] X. Zhang, X. Zhang, L. Han, and R. Xing, "Utilization-oriented spectrum allocation in an underlay cognitive radio network," *IEEE Access*, vol. 6, pp. 12905–12912, Feb. 2018.

[95] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[96] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 79–89, May 2007.

[97] I. F. Akyildiz, W.-Y. Lee, and K. R. Chowdhury, "CRAHNs: Cognitive radio ad hoc networks," *Ad Hoc Networks*, vol. 7, no. 5, pp. 810–836, Jul. 2009.

[98] K. llanko, M. Naeem, A. Anpalagan, and D. Androutsos, "Energy-efficient frequency and power allocation for cognitive radios in television systems," *IEEE Sys. J.*, vol. 10, no. 1, pp. 313-324, Mar. 2016.

[99] S. Wang, W. Shi, and C. Wang, "Energy-efficient resource management in OFDM-based cognitive radio networks under channel uncertainty," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3092–3102, Sep. 2015.

[100] A. Alabbasi, Z. Rezki, and B. Shihada, "Energy efficient resource allocation for cognitive radios: A generalized sensing analysis," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2455–2469, May 2015.

[101] D. Grace, J. Chen, T. Jiang, and P. D. Mitchell, "Using cognitive radio to deliver 'green' communications," in *Proc. Int. Conf. on Cognitive Radio Oriented Wireless Networks and Commun.*, Hanover, Germany, Jun. 2009.

[102] C. Yang, J. Li, M. Sheng, and Q. Liu, "Green heterogeneous networks: A cognitive radio idea," *IET Commun.*, vol. 6, no. 13, pp. 1952–1959, Sep. 2012.

[103] A. J. Luah, and C. K. Tan, "A Nash-based power control game for green communications via cognitive radio networks," in *Proc. IEEE Conf. on Sustainable Utilization and Development in Engineering and Technol.*, Kuala Lumpur, Malaysia. pp. 164–169, Oct. 2012.

[104] S. Dawaliby, A. Bradai, Y. Pousset, and C. Chatellier, "Joint energy and QoS-aware memetic-based scheduling for M2M communications in LTE-M," *IEEE Trans. Emerging Topics Computational Intelligence*, vol. 3, no. 3, pp. 217–229, Jun. 2019.

[105] R. Chai, C. Liu, and Q. Chen, "Energy efficiency optimization-based joint resource allocation and clustering algorithm for M2M communication systems," *IEEE Access*, vol. 7, pp. 168507–168519, Nov. 2019.

[106] Z. Zhou, J. Feng, Y. Jia, S. Mumtaz, K. M. S. Huq, J. Rodriguez, and D. Zhang, "Energy-efficient game-theoretical random access for M2M communications in overlapped cellular networks," *Computer Networks*, vol. 129, pp. 493–501, Dec. 2017.

[107] W. Ejaz and M. Ibnkahla, "Machine-to-machine communications in cognitive cellular systems," in *Proc. IEEE Int. Conf. on Ubiquitous Wireless Broadband*, Montreal, QC, Canada, Oct. 2015.

[108] D. Malak, H. S. Dhillon, and J. G. Andrews, "Optimizing data aggregation for uplink machine-to-machine communication networks," *IEEE Trans. Commun.*, vol. 64, no. 3, pp. 1274–1290, Mar. 2016.

[109] J. Rinne, J. Keskinen, P. R. Berger, D. Lupo, and M. Valkama, "Feasibility and fundamental limits of energy-harvesting based M2M communications," *Int. J. Wireless Inform. Networks*, vol. 24, no. 3, pp. 291–299, Sep. 2017.

[110] Z. Zhou, C. Zhang, J. Wang, B. Gu, S. Mumtaz, J. Rodriguez, and X. Zhao, "Energy-efficient resource allocation for energy harvesting-based cognitive machine-to-machine communications," *IEEE Trans. Cognitive Commun. Net.*, vol. 5, no. 3, pp. 595–607, Jun. 2019.

[111] Z. Yang, W. Xu, Y. Pan, C. Pan, and M. Chen, "Energy efficient resource allocation in machine-to-machine communications with multiple access and energy harvesting for IoT," *IEEE Internet Things J.*, vol. 5, no. 1, pp. 229–245, Feb. 2018.

[112] C. Kai, H. Li, L. Xu, Y. Li, and T. Jiang, "Joint subcarrier assignment with power allocation for sum rate maximization of D2D communications in wireless cellular networks," *IEEE Trans. Veh. Techol.*, vol. 68, no. 5, pp. 4748–4759, May 2019.

[113] K. Yang, S. Martin, C. Xing, J. Wu, and R. Fan, "Energy-efficient power control for device-to-device communications," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3208–3220, Dec. 2016.

[114] C. Isheden and G. P. Fettweis, "Energy-efficient multi-carrier link adaptation with sum rate-dependent circuit power," in *Proc. IEEE Global Telecommun. Conf.*, Miami, FL., Dec. 2010.

[115] W. Dinkelbach, "On nonlinear fractional programming," *Management Science*, vol. 13, no. 7, pp. 492–498, Mar. 1967.

[116] S. A. Grandhi, R. Vijayan, and D. J. Goodman, "Distributed power control in cellular radio systems," *IEEE Trans. Commun.*, vol. 42, no. 2/3/4, pp. 226–228, Feb./Mar./Apr. 1994.

[117] D. H. Ho and T. A. Gulliver, "Uplink power allocation for throughput and energy efficiency over Nakagami-m fading channels," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 1, pp. 190–202, Mar. 2021.

[118] R. L. Batista, C. F. M. e Silva, T. F. Maciel, J. M. B. da Silva, and F. R. Cavalcanti, "Joint opportunistic scheduling of cellular and device-to-device communications," *J. Commun. Info. Sys.*, vol. 32, no. 1, pp. 62–73, Aug. 2017.

[119] K. T. Phan, T. Le-Ngoc, S. A. Vorobyov, and C. Tellambura, "Power allocation in wireless relay networks: A geometric programming-based approach," in *IEEE Global Telecommun. Conf.*, New Orleans, LA, USA, pp. 1–5, Dec. 2008.

[120] C. Kai, L. Xu, J. Zhang, and M. Peng, "Joint uplink and downlink resource allocation for D2D communication underlying cellular networks," in *Proc. IEEE Int. Conf. on Wireless Commun. and Signal Process.*, Hangzhou, China, Oct. 2018.