Multimedia Internet of Things: Game Theoretic Approaches for Distributed Resource Allocation

Nicole M. Sawyer

October 2019

A thesis submitted for the degree of
Doctor of Philosophy of
The Australian National University
and CSIRO Data61

Australian National University

Research School of Engineering
College of Engineering and Computer Science
The Australian National University and CSIRO Data61

Declaration

The contents of this thesis are the results of original research and have not been submitted for a higher degree to any other university or institution.

Much of the work in this thesis has been published or has been submitted for publication as journal papers or conference proceedings.

The work in this thesis has been performed while supervised by Dr. David Smith (Data61 CSIRO and The Australian National University), Prof. Xiangyun Zhou (The Australian National University) and Prof. Leif Hanlen (The Australian National University), and also jointly collaborated with Prof. Walid Saad (Virginia Tech) and Dr. Mehdi Naderi Soorki (Isfahan University of Technology). The substantial majority of this work was my own.

Nicole M. Sawyer
Research School of Engineering,
The Australian National University,
Data61 CSIRO,
Canberra, ACT 2601,
Australia.

Acknowledgements

The work presented in this thesis would not have been possible without the support of a number of individuals and organizations, who are gratefully acknowledged below:

- Firstly, I would like to express my sincere gratitude to my primary supervisor Dr. David Smith. I am grateful for all the continuous support, opportunities, advice, guidance, and motivation throughout my PhD, it has made this journey enjoyable and unforgettable. Thank you for the freedom and flexibility to explore a range of research areas, and being patient during numerous in-depth discussions over Skype.

- I would also like to thank the rest of my supervisory panel: Prof. Xiangyun (Sean) Zhou and Prof. Leif Hanlen. They have provided me great support, advice, and insightful comments, while also asking the hard questions, to help me improve the quality of my thesis.

- I would like to thank Prof. Walid Saad for kindly welcoming me to visit his research group at Virginia Tech for two and a half months. This research visit was an invaluable experience and one of the highlights of my Ph.D. journey, which has lead to many interesting discussions and collaborations. Thank you Prof. Saad for all the comments and feedback during our collaborations. The collaborations with Virginia Tech, have strengthened my research skills. I would also like to thank the rest of the research group at Virginia Tech, for their hospitality during my visit.

- I would like to thank Dr. Mehdi Naderi Soorki for providing supervision and guidance during my research visit at Virginia Tech, as well as the continued supervision and guidance once I returned to ANU. Thank you for your time, support, insight, and patience, during our collaborative research.

- I would also like to thank everyone in the communications, signal processing, and acoustics research groups at the Research School of Engineering, ANU. Thank you to everyone, for making this a friendly and supportive environment to study at ANU. Special thanks to Prof. Salman Durrani, for his continued support, advice, guidance, and patience, during my tutoring and course convening. Thank you to my colleagues, both past and present, Dr. Nan Yang, Prof. Parastoo Sadeghi, Prof. Rod Kennedy, Dr. Alice Bates, Dr. Yeif Huang, Dr. Jing Guo, Dr. Prasanga Samarasinghe, Samaneh Movassaghi, Jihui (Aimee) Zhang, Noman Akbar, Samiya Shimly, Khurram Shahzad, Usama Elahi, Yuting Fang, and Katrina Zhou.

- I wish to thank my loving and supportive partner, Josh, for everything you have done for me during my studies. Thank you for all your encouragement and advice. I would also like to thank you for putting a smile on my face everyday, making me laugh, and making sure I took time to relax from my studies. Special thanks to Josh's parents and sister, for all their support and encouragement.

- Lastly, and most importantly, heartfelt thanks go to my parents, for their support, unconditional love, and encouragement, throughout my many years of study. You are always there for me. My brother, Oliver, thank you for your encouragement and always looking out for me. Many thanks to the rest of my family for all their support as well. Finally, special thanks go to Peggy Sue (my pet cat), for the many needed distractions and cuddles during my PhD.

Abstract

Multimedia Internet of Things (IoT) is an emerging paradigm enabling devices to form an interconnected network of all types of communication. Multimedia IoT is integral to future wireless networks, but there are many challenges that must be addressed, including: network modelling; resource management; massive scale; interference management; clustering; and user quality-of-experience.

This thesis focuses on three key enabling technologies of multimedia IoT: (i) device-to-device (D2D) communications; (ii) machine-to-machine (M2M) communications; and (iii) unmanned aerial vehicles (UAVs) as low flying base stations (BSs). Using game theory, we develop novel solutions for decision making processes required to address the key challenges and technologies of multimedia IoT, where interactions between users competing for resources are modelled in a fully distributed and autonomous manner.

In the first half of this thesis, we study game theoretic approaches for resource allocation in underlaid D2D communications. First, we propose a flexible application-driven resource allocation scheme, to enhance and improve D2D user quality-of-experience and reliability. We propose a multiple objective Stackelberg game using a non-scalarised approach to enable flexible resource allocation, by coordinating and reducing the effects of intra-cell interference, while ensuring D2D user quality-of-experience. We demonstrate that best social welfare is guaranteed across all D2D users.

Next, to further reduce intra-cell interference in underlaid D2D networks, we jointly optimise D2D mode selection, resource block allocation, and interference management. In existing D2D mode selection techniques, the BS assists D2D pairs in selecting a transmission mode, requiring complete channel state information (CSI), i.e., a network-assisted scenario. We propose both a non-cooperative game and a coalitional game to solve this joint optimisation problem in a network-assisted scenario. Next, we investigate a user-assisted approach as a viable D2D mode selec-

tion solution, where the BS has partial CSI. Hence, we extend the coalitional game to consider the user-assisted approach, while also coordinating intra-cell interference, scheduling D2D pairs, and allocating resource blocks efficiently. Extensive simulations validate the effectiveness of the proposed coalitional game in a user-assisted scenario.

In the second half of this thesis, we explore two emerging topics in M2M communications: (i) correlation-aware clustering in dense M2M networks; and (ii) UAVs as low flying BSs. We first propose a clustering algorithm to cluster a massive number of machine-type devices based on data correlation, which reduces the number of redundant bits being sent to the BS. To solve this, we propose an evolutionary game, and derive a novel utility function that captures the average machine-type device transmission power per cluster, using stochastic geometry. The proposed algorithm converges to a stable cluster formation, which is robust to stochastic changes in the M2M network environment.

Finally, we investigate an energy-aware scheduling scheme for mission critical M2M communications, utilising UAVs as low flying BSs. To solve this, we combine Lyapunov optimisation and a one-to-many matching game, where UAVs schedule themselves to collect sensed data from aggregators, to satisfy the ultra-reliable and low-latency constraint, whereas aggregators aim to decrease their energy consumption. Two-sided stable matching and stable power allocation is demonstrated for the proposed algorithm.

List of Publications

The work in this thesis has been published or has been submitted for publication as journal/conference papers. These papers are:

Journal papers:

J1 **N. Sawyer**, and D. B. Smith, "A Nash Stable Cross-Layer Coalitional Game for Resource Utilization in Device-to-Device Communications," *IEEE Trans. on Vehicular Technology*, vol. 67, no. 9, pp. 8608–8622, Sept. 2018.

J2 **N. Sawyer**, and D. B. Smith, "Flexible Resource Allocation in Device-to-Device Communications using Stackelberg Game Theory," *IEEE Trans. on Communications*, vol. 67, no. 1, pp. 653–667, Jan. 2019.

J3 **N. Sawyer**, M. Naderi Soorki, W. Saad, D. B. Smith, and N. Ding, "Evolutionary Games for Correlation-Aware Clustering in Massive Machine-to-Machine Networks," *IEEE Trans. on Communications*, May 2019.

J4 **N. Sawyer**, M. Naderi Soorki, W. Saad, N. Ding, D. B. Smith, M. H. Manshaei, and M. Mozaffari, "Energy-Aware Scheduling for Ultra-Reliable Internet of Things Communications using Unmanned Aerial Vehicles," (under preparation to be submitted to *IEEE Internet of Things Journal*).

Conference papers:

C1 **N. Sawyer**, and D. B. Smith, "Pareto-Efficient Cross-Layer Repeated Game for Device-to-Device (D2D) Communications," in *Proc. IEEE Int. Conf. on Communications (ICC)*, Kuala Lumpur, Malaysia, pp. 1–6, May 2016.

C2 **N. Sawyer**, and D. B. Smith, "A Nash Stable Cross-Layer Coalition Formation Game for Device-to-Device Communications," in *Proc. IEEE Int. Conf. on Communications (ICC)*, Paris, France, pp. 1-6, May 2017.

C3 **N. Sawyer**, M. Naderi Soorki, W. Saad, and D. B. Smith, "Evolutionary Coalitional Game for Correlation-Aware Clustering in Machine-to-Machine Communications," in *Proc. Global Communications Conf. (GLOBECOM)*, Singapore, pp. 1-6, Dec. 2017.

List of Abbreviations

3GPP	3rd Generation Partnership Project
5G	fifth generation
AWGN	additive while Gaussian noise
BS	base station
CTD	cellular-type device
CSI	channel state information
D2D	device-to-device
dB	decibel
dBm	decibel-miliwatts
ESS	evolutionary stable state
GHz	Gigahertz
IoT	Internet of Things
LTE	Long-Term Evoluion
M2M	machine-to-machine
MTD	machine-type device
OFDMA	orthogonal frequency-division multiple access
PDR	packet delivery ratio
PPP	Poisson point process
QAM	quadrature amplitude modulation
QoE	quality-of-experience
QoS	quality-of-service
QPSK	quadrature phase shift keying
SINR	signal to interference plus noise ratio
UAV	unmanned aerial vehicle
URLLC	ultra-reliable low latency communications

List of Notations

The following mathematical notations are consistent throughout the entire thesis:

a	scalar variable
$\boldsymbol{a}$	vector variable
$\mathbf{A}$	matrix variable
$\mathcal{A}$	set variable
$\|\|\boldsymbol{a}\|\|$	Euclidean norm of vector $\boldsymbol{a}$
$\|a\|$	absolute value of scalar a
$\|\mathcal{A}\|$	size of set $\mathcal{A}$
$\log_n(\cdot)$	logarithm of base n
$\mathbf{E}(\cdot)$	mathematical expectation operator
$\Gamma(\cdot)$	Gamma function
$(\cdot)!$	factorial
$\max(\cdot)$	maximum operator
$\emptyset$	empty set
$\setminus$	set exclusion operator
$\lfloor\cdot\rfloor$	floor operator
$\cup$	union operator

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

A predominant influence in the evolution of wireless cellular networks is the exponential growth in the number of wireless mobile devices, high mobile data traffic, and the need for better quality-of-experience (QoE). In particular, the evolution of wireless cellular networks has been a major driver for the development of multimedia Internet of Things (IoT). Multimedia IoT is an emerging paradigm, that encompasses both cellular-type devices (CTDs) and machine-type devices (MTDs) to form a fully interconnected network, which will provide unlimited access to a range of information and sharing of data [3–6]. In fact, wireless sensor networks are a core part of multimedia IoT paradigm [6].

According to Cisco, there were 18 billion global mobile devices and connections in the year 2017, which is expected to increase to 28.5 billion by 2022, i.e., a compound annual growth rate of 10% [7]. Out of the 28.5 billion global mobile devices and connections in 2022, 14.6 billion will be from MTDs (i.e., sensing, smart metering, actuators, wearable devices, etc.) and the rest will be from CTDS (i.e., smart phones or personal mobile-devices) [7]. In fact, mobile video traffic will account for 82 % of the total mobile data traffic by 2022 [7]. Hence, supporting the massive number of CTDs and MTDs in multimedia IoT is a key design challenge.

In traditional cellular networks, a BS is located at the center of the network and provides coverage to all users within the coverage area. Consider a simple scenario, as outlined in Fig. 1.1, where we have cellular user A that wishes to communicate with cellular user B within the coverage area of the same BS. Cellular user A will transmit its message to the BS via an uplink, where the BS will then forward (relay) the message to cellular user B via the downlink. Hence, all

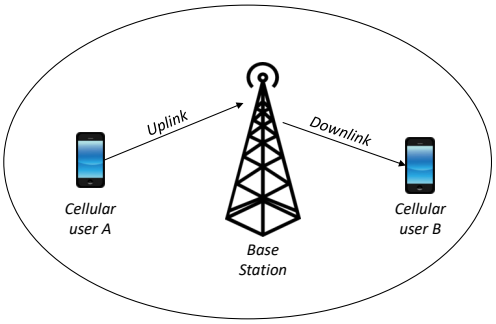


Figure 1.1: Illustration of traditional cellular communications.

cellular users relay their messages via the BS, and thus, the BS has control over all cellular communications within its coverage area. However, due to the number of CTDs exponentially increasing and the addition of the increasing number of MTDs within the cellular network, this will cause an increase in the traffic and demand on the BS and cellular resources [2, 8, 9]. In addition, individual mobile users are now demanding higher data rates for increased social networking applications, gaming, augmented virtual reality, and video streaming [2, 10, 11]. These factors for CTDs, along side the periodic/non-periodic data traffic of MTDs, will continue to increase the data traffic within cellular networks. Hence, this will lead to performance degradation of the network and the BS, which, in turn, will cause increased latency periods and outages, and reduced communication reliability and QoS for all users [2]. Thus, traditional cellular networks have become inadequate to support the massive number of devices, as well as the requirements and demands of these devices. To improve the current cellular networks, new paradigms such as multimedia IoT in next generation wireless networks are needed.

The next generation wireless network, 5G, is expected to provide: (i) enhanced mobile broadband, (ii) massive machine-type communication, and (iii) Ultra-Reliable Low Latency Communications (URLLC) [10, 11]. In fact, 5G systems will aim to provide increased capacity, higher data rate, increased coverage area, low end-to-end latency, high reliability, and support a massive number of devices [7,8,12]. To meet these criteria, it is expected that the future 5G network will provide the following enhancements over the current 4G network [7, 8, 10, 11, 13]:

- Increase date rate up to 1000$\times$ from 4G, in order to support applications such as high-definition video streaming and augmented virtual reality. Moreover,

cell-edge devices will receive an increased data rate of at least 100 Mbps.

- Ultra-low end-to-end latency of less than 1 ms, compared to current 4G roundtrip latencies of 15 ms, as well as ultra-high reliability guaranteeing no less than $1 - 10^{-5}$.

- Improving energy efficiency by decreasing the energy per bit and the cost per bit for each data link by 100×.

- Support a massive number of devices (1 million MTDs/km^2) densely deployed within the network.

Multimedia IoT will support a diverse range of devices within the network, i.e., CTDs, MTDs and UAVs, in order to revolutionise and improve our everyday lives [4, 5]. In particular, multimedia IoT will encompass all forms of media, ranging from high bandwidth applications for enhanced mobile broadband, to telemetry data sensing for massive machine-type communications, and ultra reliable and low latency for mission critical communications. Hence, in order to achieve these 5G requirements and assist in the evolution of the cellular network, new wireless technologies are being considered [8,14]. The range of new wireless technologies include: millimeter wave; massive multiple-input multiple-output (MIMO); heterogeneous networks; D2D communications; M2M communications; and UAVs [2, 15]. From these new wireless technologies, in order to investigate a wide range of media type applications, this thesis will focus on three key enabling technologies of multimedia IoT: (i) D2D communications; (ii) M2M communications; and (iii) UAVs as low flying BSs in M2M communications. In particular here, investigations into D2D communications capture the area of enhanced mobile broadband, as it focuses on increased data rate, spectral efficiency, user experience, and energy efficiency. Here, M2M communications are investigated for supporting a massive number of MTDs in the cellular network that are densely deployed, as well as for utilising UAVs as flying BSs to assist in guaranteeing URLLC constraints in order to provide low latency and high reliability for mission critical applications. Fig. 1.2 outlines the three main 5G use cases for multimedia IoT, as well as application examples.

D2D communications enables users who are within close proximity to transmit directly between one another, without relaying through a BS. Thus, D2D communications is an integral part in the current and future cellular network, as it alleviates network congestion, traffic, and demand on the BS. There are several

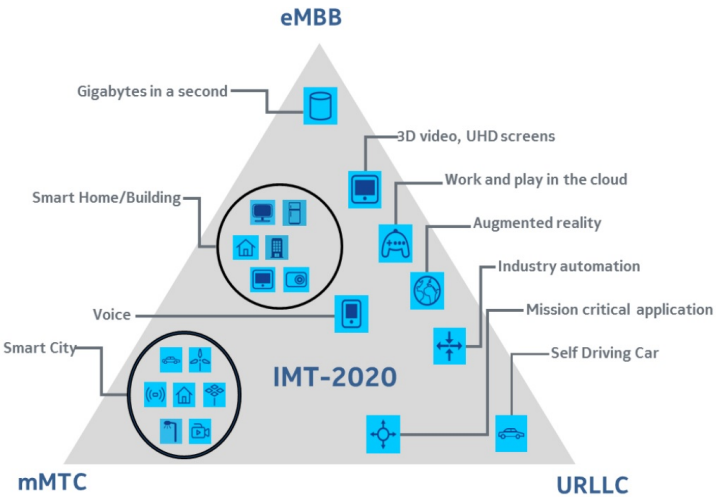


Figure 1.2: Three main 5G use cases for multimedia IoT and examples of applications [1].

advantages to both users and the network when adopting D2D communications, such as increased spectral efficiency, QoS, reliability, energy efficiency, throughput, system capacity, and scalability, as well as reduced latency [2, 9, 16]. D2D communications can be implemented in either in-band (licensed spectrum) or out-band (unlicensed spectrum). Moreover, enabling direct communication between CTDs in multimedia IoT, will cause an increase in battery lifetime and larger coverage areas.

Like D2D communications, M2M communications is also an integral part in the current and future cellular network, as it enables advanced network applications such as smart home technologies, healthcare, drone systems, and surveillance [17–22]. In fact, 3GPP conducted a study into further D2D enhancements [23], which proposed a user-to-network relay for IoT and wearables (i.e., MTDs) utilising D2D communications. This study aimed at utilizing the sidelink (i.e., the direct link) to increase energy-efficiency and communication reliability for all users in the network, as well as reducing power consumption and guaranteeing QoS for a range of different traffic types [23, 24]. An MTD can be a sensor, actuator, or smart meter whose typical role is to sense or measure an environment, and transmit the collected data to BSs. Thus, MTDs generally have little or no mobility and typically transmit small amounts of data [18]. Moreover, MTDs enable real-time monitoring and control of any physical environment, without direct human involvement, thus making processes more efficient and improving human welfare [25, 26].

Meanwhile, UAVs deployed in cellular networks can provide reliable and cost

effective solutions for practical scenarios, due to their high mobility, agility, and flexibility [27,28]. In fact, UAVs can be used as aerial BSs, which can deliver reliable communications, as well as increased network capacity and coverage area [29, 30]. UAV BSs can be deployed to several scenarios, ranging from highly densely populated areas to regions without coverage or poor connectivity. Furthermore, UAVs in multimedia IoT will play a key role in assisting M2M communications. For instance, an MTD may not have enough energy to transmit their data to the BS, due to limited battery lifetime, power constraints, and large transmission distances. Thus, UAVs can be deployed within close proximity to MTDs, in order to assist in forwarding MTD data to the cellular BS.

Implementing multimedia IoT in future wireless networks, poses significant challenges to D2D communications, M2M communications, and UAVs as low flying BSs, ranging from network modelling to resource management, scalability, interference management, user QoE, energy-efficiency, latency, and clustering. The challenges for these three key enabling technologies of multimedia IoT are discussed in the next subsection.

1.1.1 Research Challenges in Multimedia IoT

1.1.1.1 D2D communications

D2D communications are classified as cellular oriented IoT networks, which focuses on a range of media type applications, including high bandwidth applications. The introduction of D2D communications in the current and future cellular networks, can be implemented in either in-band or out-band. In the case of in-band, D2D users will utilise licensed spectrum, whereas in the case of out-band, D2D users will utilise unlicensed spectrum. In particular, in-band D2D communications (licensed spectrum) will guarantee higher QoS over out-band D2D communications (unlicensed spectrum), as out-band is susceptible to uncontrolled interference [16]. In this thesis, we focuses on in-band D2D communications. In particular, in-band D2D communications poses three main challenges, which are: D2D mode selection, resource allocation and interference management, and quality-of-experience. In the following we will give an overview of these main challenges.

Mode Selection: D2D mode selection enables D2D pairs utilising licensed spec-

trum, to select a transmission mode to operate in, i.e., D2D pairs need to determine whether they should transmit via a direct link or a cellular link [16]. Fig. 1.3 outlines an illustrative example of the three transmission modes for D2D communications. Enabling D2D mode selection, allows D2D pairs to operate more efficiently, which leads to improved system capacity and throughput. Hence, the set of possible transmission modes for D2D communications are [16]:

- *Cellular Mode*: Enables D2D pairs to transmit via traditional cellular communications, i.e., relaying messages through a BS (no direct communication link). This mode is typically used if the distance between the D2D pair transmitter and receiver is large, or if the D2D pair is located within close proximity to a BS.

- *Dedicated Mode*: D2D pairs transmit directly between themselves. The BS allocates these D2D pairs orthogonal resources, i.e., dedicated spectrum (overlay). Thus, D2D pairs in dedicated mode will not receive or cause any interference within the cellular network, however, they will need to wait till resources become available.

- *Reuse Mode*: D2D pairs also transmit directly between themselves. The BS allocates these D2D pairs non-orthogonal resources, i.e., D2D pairs share existing cellular resources (underlay). Hence, D2D pairs in reuse mode will cause interference to cellular users and other D2D pairs sharing the same resource/s.

In order for a D2D pair to determine a suitable transmission mode to operate in, the BS typically assists the D2D pairs in making such a decision, which is referred to as network-assisted D2D communications. Network-assisted D2D communications assumes that the BS must have global knowledge of the network, such as complete CSI and the location of all users [16, 31–33]. In order for the BS to acquire global knowledge of the network, this will cause signalling overhead and network complexity to significantly increase, especially with the number of users in the network is exponentially increasing. Furthermore, in practical scenarios, network-assisted D2D communications may not be ideal, as complete CSI cannot always be achieved for all users at the time.

Therefore, the main challenge is how to determine suitable criteria for D2D pairs to select an appropriate transmission mode, in a distributed manner. In particular,

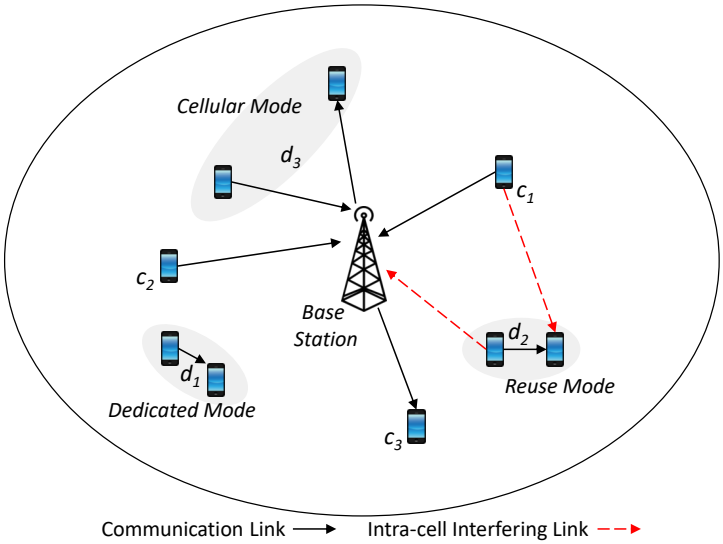


Figure 1.3: Illustrative scenario of all transmission modes for D2D communications. In this example, D2D pair d_1 is transmitting in dedicated mode, D2D pair d_2 is transmitting in reuse mode and causing intra-cell interference, and D2D pair d_3 is transmitting in cellular mode.

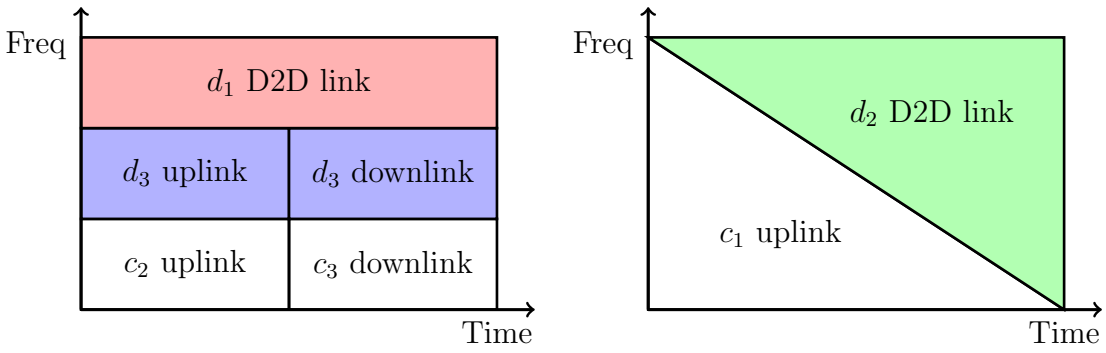


Figure 1.4: Resource block allocation for the example scenario shown in Fig. 1.3.

it is desirable to design a D2D mode selection approach, where the BS has partial CSI of the network, i.e., the BS will have little involvement in the mode selection decision. Moreover, D2D mode selection also relates to the resource allocation and interference management challenge in D2D communications, which focuses on, how resources are allocated to D2D users, as well as how interference can be handled within the network, given a particular mode selection.

Resource Allocation and Interference Management: Resource allocation and interference management is essential in in-band D2D communications. Fig. 1.4 outlines an illustrative example of the resource block allocation for all transmission

modes in D2D communications, based on the scenario in Fig. 1.3. For reuse mode, D2D pairs can reuse/share the uplink or downlink resource block/s of cellular user/s, which causes severe interference between users in the network. Hence, the interference may significantly degrade cellular user performance and QoS. In the cellular network, interference generated is characterised as intra-cell interference and inter-cell interference. In the case of intra-cell interference, this is generated between devices in a cell using the same time and frequency resources. In contrast, inter-cell interference is generated when users from different cells in the network use the same time and frequency resources.

In uplink resource sharing, D2D pairs will cause intra-cell interference to the cellular user's uplink, as well as all other D2D pairs sharing the same uplink resources. On the other hand, in downlink resource sharing, D2D pairs will cause intra-cell interference to the cellular user's downlink and the BS will cause interference to the D2D pair receiver. In particular, for uplink resource sharing, the BS can handle the extra interference that is caused by D2D communications, without requiring additional hardware [16, 34, 35]. Whereas, for downlink resource sharing, the D2D pair receiver may require extra hardware to handle the large amount of interference from the BS [16, 34, 35]. However, in either case, interference management is needed in order to ensure users operate more efficiently and their performance is guaranteed.

Meanwhile, for dedicated and cellular modes, the BS will allocate orthogonal resources. D2D pairs operating in cellular mode are guaranteed orthogonal resources, as they are essentially operating as traditional cellular users. In contrast, coordination is required between D2D pairs and the BS, when allocating orthogonal resources to D2D pairs in dedicated mode. This is because, the BS needs to assess if there are any free/idle resource blocks available to be allocated to D2D pairs [36].

Therefore, there is a need to adopt distributed resource allocation and interference management schemes, in order to coordinate and reduce the interference between users in the network, as well as allocate orthogonal resources efficiently.

Quality-of-Experience: There is a wide range of potential applications for D2D communications in multimedia IoT, such as, messaging, content sharing (such as videos and photos), local voice calling, gaming (single or multi-player), data shar-

ing, advertising, uploading/downloading or streaming content, and emergency services (local security and safety communications between neighboring police, firemen, and ambulance officers) [2,37,38]. These different applications in multimedia IoT focus on different forms of media, where each application has different QoE requirements. Thus, depending on the D2D application, particular factors such as, reliability, priority, throughput, or transmit power, can be enhanced in order to increase each D2D users QoE. For example, if low data rate was provided to a D2D user whose application is uploading content, then this would have a negative effect on user satisfaction, when compared to providing high data rate. Hence, guaranteeing D2D user QoE ensures end-user satisfaction, and it is an essential measure of communication quality, alongside QoS, in current and future cellular networks [6,39–41]. Therefore, the main challenge is, how to ensure that all D2D user's guarantee QoE, for a range of different applications, in a distributed manner, as well as how to determine the QoE requirements for multimedia IoT applications.

1.1.1.2 M2M communications

M2M communications is an important component of the emerging IoT system, and in typical applications MTDs capture telemetry data from sensing an environment and transmit that data to the BS. These are very different forms of media than those of cellular IoT. Moreover, M2M communications can be categorised into two groups, massive M2M communications and URLLC [42, 43]. Massive M2M communications focuses on increasing battery lifetime, data correlation, scalability, and data rate [44]. On the other hand, URLLC focuses on mission critical M2M communications, which focuses on decreasing end-to-end latency and increasing communication reliability for mission critical M2M applications [42].

Currently, the number of available resource blocks in cellular networks have been designed to suit the needs of CTDs. Due to the massive number of MTDs, the number of available resource blocks for each MTD is limited, and as a result, not all MTDs will be allocated orthogonal resources for transmission to the BS. If MTDs are unable to send their sensed data to the BS, this may lead to missing information for a particular environment. Moreover, due to the dense deployment of MTDs, MTDs within close proximity will gather correlated data due to sensing the same environment [44,45]. For example, sensors deployed to measure temperature in the same room, will measure very similar readings, which is highly correlated. Hence,

multiple MTDs will send a large number of redundant bits (same information) to the BS.

Recently, the idea of clustering MTDs into smaller groups has emerged as a promising technique to reduce the traffic load on the cellular BS and improve spatial reuse and energy efficiency, while reducing interference in the network [18, 26, 46], and [47]. In M2M communications there are two types of clustering, machine-centric and data-centric. Machine-centric clustering approaches, cluster MTDs in order to maximise MTD data rate and the number of supported MTDs, while minimising energy consumption. Such a machine-centric clustering approach does not take into account the individual data/information of each MTD. On the other hand, data-centric clustering approaches can be used to improve the data quality sent to the BS. Therefore the main challenge is how to design a distributed correlated-aware clustering approach for a massive number of MTDs.

1.1.1.3 UAVs as Low Flying BSs in mission critical M2M communications

The use of unmanned aerial vehicles (UAVs) as flying BSs have been used to provide reliable, cost-effective, and energy-efficient uplink communications for MTDs deployed in areas which experience intermittent or poor covereage, or even no access to cellular infrastructure [20, 28, 48–51]. In particular, the aerial nature of UAVs, such as their dynamic mobility, flexibility, and adaptive altitude, enables UAVs to move towards MTDs to collect their data [27]. In fact, due to the high altitude of UAVs as flying BSs, this enables them to establish a line-of-sight (LoS) communication link between MTDs [30, 52]. Thus, this can lead to reduced shadowing and signal blocking effects, while also establishing a reliable communication link with lower transmission power needed for sending the sensed data. However, despite the many advantages of utilising UAVs as flying BSs, there are a number of challenges that need to be considered that range from optimal three-dimensional deployment of UAVS to path planning (trajectory optimisation), flight time, dwelling time, and resource management.

In particular, mission critical M2M communications aim to guarantee ultra reliable and low latency communications. There is a range of different approaches to guarantee URLLC in mission critical applications, such as, short packet transmissions, shorter transmission time interval (TTI), caching, network slicing, or queue

latency [42, 53–59]. Hence, for mission critical applications, relaying on average requirements (such as, average rate or average delay) is no longer adequate [42].

Therefore, the main challenge is, how to ensure UAVs as flying BSs will satisfy the URLLC constraint for mission critical M2M communications, as well as how to determine the required number of UAVs to service all MTDs or the appropriate dwelling time for each UAV, in a distributed and autonomous manner.

1.2 Thesis Objective

The outlined research challenges for multimedia IoT leads to the following hypothesis and key research questions that are answered by this thesis:

In multimedia IoT — in order to obtain optimal reliability, scalability, energy-efficiency, and quality-of-experience — resource allocation can be performed in a distributed and autonomous manner. The design of such optimal resource allocation schemes is the aim of this research.

In this context, we develop novel solutions for distributed decision making processes using game theoretic tools. Game theory is a powerful mathematical tool that models the interactions between users competing for resources in a fully distributed and autonomous manner enabling us to *answer the following* in this thesis:

Q1. How can user satisfaction be enhanced in future wireless networks using game theory?

Q2. What are the effects of prioritising communications while enhancing user satisfaction?

Q3. How can a D2D user determine which transmission mode to operate in, assuming full channel state information?

Q4. What is the benefits in using cooperative game theory over non-cooperative game theory, to determine D2D mode selection?

Q5. How does partial CSI affect D2D users in selecting a transmission mode?

Q6. Can MTDs be clustered based on data correlation? Does this enable improved energy-efficiency?

Q7. What are the effects of considering a massive number of MTDs for fully distributed correlation aware clustering?

Q8. Can UAVs satisfy the ultra reliable low-latency communication constraints as aerial BSs, in an IoT network?

The key contributions of this thesis are outlined in four main technical chapters (Chapters 3 - 6). For the remainder of this chapter, we first provide background information on the key performance metrics and game theoretic approaches. Then, a detailed overview of the thesis contributions and an outline of the thesis layout. Chapter 2 provides a comprehensive literature review of the relevant work on multimedia IoT, which is divided into the three key areas we are investigating, i.e., D2D communications, M2M communications, and UAVs as low flying BSs. In Chapter 3, we propose a flexible resource allocation scheme for D2D communications, to enhance and improve D2D user QoE and reliability while coordinating and reducing the affects of interference. Then in Chapter 4, we jointly optimise D2D mode selection, resource block allocation, and power control, while investigating the effects of non-cooperative and cooperative game theoretic approaches for D2D communications. In particular, we also investigate the effects of D2D mode selection for complete and partial CSI. Moreover, in Chapter 5, we investigate MTD clustering based on data correlation for a massive number of MTDs, where the stable cluster formation is robust to stochastic changes in the M2M environment. In the final technical chapter (Chapter 6), we investigate an energy-aware scheduling scheme that guarantees ultra-reliability for MTDs, where UAVs are utilised as low flying BSs.

1.3 Performance Metrics and Analytical Tools

In this section, we provide background information of the techniques used in this thesis, which largely makes the thesis self-contained.

1.3.1 Key Performance Metrics

In the cellular network, fundamental performance metrics such as SINR, achievable spectral efficiency (channel rate), packet delivery ratio, and an ultra reliability constraint, are used to evaluate the performance of all users within the cellular network. In the following, we outline the four key performance metrics that are employed in this thesis, to analyse the individual user performance, as well as the network performance.

- **Signal-to-Interference-plus-Noise Ratio (SINR):** Is defined as a ratio between the signal power of a user, to noise power plus any interfering signals in the network. Typically SINR is used to measure the quality of a signal. Thus, to ensure interference within the network will not degrade any user's performance (QoS), we place a threshold on received SINR for all users in the network, $\gamma_i \geq \gamma_{\text{th}}$, where γ_i is user i's SINR.

- **Achievable spectral efficiency (channel rate):** Indicates the user's experience for a particular application, such as for a voice call or content sharing. In general, the achievable spectral efficiency, r, of a user assumes Gaussian signalling, and is given by:

$$r = \log_2 (1 + \gamma) . \tag{1.1}$$

  Achievable spectral efficiency and SINR are dependent on one another, where threshold channel rate r_{th} is a function of threshold SINR γ_{th}.

- **Packet Delivery Ratio (PDR):** Measures the communication success in terms of the number of packets successfully received at the receiver out of the toal number of transmitted packets. In particular, PDR assumes a practical modulation scheme and is defined as a compressed exponential function of inverse SINR [60], as follows:

$$pdr_i = \exp\left(-\left(\frac{1}{\gamma_i a_c}\right)^{b_c}\right), \tag{1.2}$$

  where a_c and b_c are constants that depend on packet size, type of modulation, and coding scheme. For simplicity, the PDR can be reexpressed as a function

of SINR, as follows:

$$pdr_i = \exp\left(a\gamma_i^b\right), \tag{1.3}$$

where $a = -(1/a_c)^{b_c}$, and $b = -b_c$.

To ensure acceptable PDR is achieved for all users we seek to meet or exceed a target PDR pdr_{th}, such that $pdr_i \geq pdr_{\text{th}}$, where target PDR is defined in terms of threshold SINR γ_{th} as, $pdr_{\text{th}} = \exp\left(a\gamma_{\text{th}}^b\right)$. PDR can also be used to prioritise communications, by assigning particular users with higher target PDR than others, which means that higher target PDR entails higher communication priority and greater reliability.

- **Ultra Reliability Constraint:** Is a reliability measure for URLLC, where the notion of risk is leveraged (risk is synonymous with losing urgent messages), i.e., the ultra reliability constraint aims to minimise the *risk of losing messages*. To do this, we use the entropic risk measure $\ln(\mathbb{E}\{e^{\delta Q_t}\})/\delta$ with a risk sensitivity parameter $\delta > 0$ as our reliability metric, where Q_t is the maximum queue length at time slot t (i.e., the worst-case queuing delay) [42]. By imposing a finite threshold κ on the the entropic risk measure, the network will be *ultra reliable*, if the following requirement is met:

$$\limsup_{t\to\infty} \frac{\ln(\mathbb{E}\{e^{\delta Q_t}\})}{\delta} \leq \kappa. \tag{1.4}$$

Hence, the ultra reliability constraint, ensures that the risk of the maximum queue length is finite and does not approach infinity, which in turn will reduce delay.

1.3.2 Game Theory

In this section, we review game theoretic approaches employed in this thesis. This thesis deals with both static deterministic games (i.e., non-cooperative power control game, Stackelberg game, matching game, and coalitional game) and dynamic games (i.e., evolutionary game), with imperfect and complete information. Game theory was developed by John Neuman and Oskar Morgenstern in 1944 [61]. Then, John Nash developed an important solution concept in game theory, which is called the *Nash Equilibrium*. Game theory has been widely used in many different fields, ranging from economics to biology, and wireless communications.

Game theory is a decision making process and models the interactions between users competing for resources in a fully distributed manner. It is a powerful mathematical tool which is useful for solving the aforementioned research questions, rather than using other classical optimisation techniques. In general, most classical optimisation techniques provide centralized solutions, where often a centralized solution can be NP-hard and complex, due to the combinatorial nature of such problems [62]. By contrast, game theory has a distributed decision making nature, that can enable faster and realistic convergence to feasible solutions. In order to determine the efficiency of a game compared to the classical optimization approaches, price of anarchy can be used to measure the selfishness of the users in a game [61].

Within any type of game there are three main elements: a set of players; a set of actions (strategy set); and payoffs (payoffs can also be known as utility functions). The set of players are assumed to be rational, where each player selects an action from the set strategy set. In this thesis, since we have considered games with imperfect and complete information, which means that all players will select an action simultaneously (i.e., at the same time), but will also know the action set and payoffs of all players [63, 64]. Furthermore, since players are rational, this means that they aim to maximise their payoff [61, 63, 65, 66]. In particular, a utility function measures the performance of a player, and is a function of the player's selected action. Hence, during a game, the players can dynamically update their selected action, in order to achieve a better payoff. Note that, in the static deterministic games, the dynamics of a repeated game are analysed, i.e., the best response, which is different to determining the equilibrium of a dynamic game.

Games can be classified into two types: non-cooperative or cooperative. In non-cooperative game theory, no communication exists between players, where each player is selfish and chooses its action independently to improve its own payoff [61, 65, 67]. On the other hand, in cooperative game theory, players cooperate with each other and form groups known as coalitions. Within each coalition, players cooperate and share information, in order to maximise the coalition value [61, 65, 67]. Moreover, a game can be static or repeated, where a static game is a single-shot game which means that there is only one iteration, and players will only have one move/interaction. Whereas, a repeated game will have many iterations, and the players will interact more than once. During each iteration, the strategy set in a game can be either pure or mixed. In the case of pure strategies,

players will select one action per iteration and play it [63, 66, 68]. However, in the case of mixed strategies, each player will assign a set of probabilities for each strategy in the strategy set [63, 66, 68].

Non-cooperative Power Control Game

A non-cooperative power control game is a popular tool in wireless communications for addressing the interference management problem. The aim of this game is to minimise transmit power for all users in the network, which also reduces interference, and increases power efficiency and user battery lifetime. A non-cooperative power control game is defined as $G_{\text{NPC}} = [\mathcal{N}, \mathcal{A}, u_i(\cdot)]$. In this game, $\mathcal{N}$ is the set of players, $\mathcal{A}$ is the finite set of strategies for all players, and $u_i(\cdot)$ is the utility function for player i that best describes the player's performance. The set of players in the power control game is the set users in the wireless network, and transmit power is the action set which is bounded by a minimum and maximum transmit power. The players have imperfect information, which means that all players select their action simultaneously. The non-cooperative power control game maximises each player's utility function, with respect to all other players in the game, in order to solve for optimal transmit power.

In a non-cooperative game, there are two important solution concepts which are used to analyse the performance of a game, which are: Nash Equilibrium and Pareto optimality [63]. The Nash equilibrium is a stable point in the game, where no player will have incentive to unilaterally change their strategy without reducing their payoff [61, 66, 68]. In this thesis, we have considered the concept of Nash Equilibrium, as it provides strong assurance for the best response across all players in the game. Moreover, the Nash Equilibrium for G_{NPC}, is defined as:

Definition 1.1 The strategy profile $\mathcal{A}^* = \{a_1, a_2, \ldots, a_N\}$, is a Nash Equilibrium for each stage in the game, if for every player $i \in \mathcal{N}$, there exists:

$$u_i(a_i^*, \boldsymbol{a}_{-i}^*) \geq u_i(a_i, \boldsymbol{a}_{-i}^*), \forall a_i \in \mathcal{A}_i. \tag{1.5}$$

Moreover, a strategy profile $\mathcal{A}$ is said to be Pareto optimal, if an outcome of player $i \in \mathcal{N}$ cannot be made better off, without making any other player's outcome worse off. Thus, Pareto optimality is defined as:

Definition 1.2 Strategy profile $\mathcal{A}$ is Pareto optimal if $\forall i \in \mathcal{N}$, $u_i(\mathcal{A}) \geq u_i(\mathcal{A}')$,

and there exists some $j \in \mathcal{N}$ for which $u_j(\mathcal{A}) \geq u_j(\mathcal{A}')$, where $\mathcal{A}'$ is another strategy profile and $\mathcal{A}' \neq \mathcal{A}$.

Stackelberg Game

A Stackelberg game consists of a hierarchical structure with a leader and follower. In fact, Stackelberg games can be made up of either multiple individual leader-follower pairs, or a single leader with multiple followers (or vice versa). In a Stackelberg game, the leader moves first, by setting a price that is charged to the follower/s. Next, the follower/s move, and react to the charged price by updating their strategy to maximise their individual utility. In the case where there are multiple followers, the followers compete with each other for resources. Thus, depending on how the follower/s react to the charges price, this will, in turn, affect the leader's strategy. Furthermore, a Stackelberg game considers one utility function assigned to the leader/s, which may or may not be different to the one utility function assigned to the follower/s.

Meanwhile, a Stackelberg game can be classed as a bi-level multi-objective optimisation problem (MOP), which are a subset of multi-level MOPs. Multi-level (and bi-level) MOPs consist of a leader level and multiple follower levels, where several objective (utility) functions are available to followers [69, 70]. To solve a multi-level MOP, typically scalarisation is used, which combines all the separate objective functions at each follower level into one objective (utility) function with multiple objectives. Thus, the problem now considers only one utility function for the leader and one utility function for the followers. Consider the bi-level MOP example from [69, Example 2]. This scalarisation approach can be generalized further by considering multi-level Stackelberg , where each level consists of multiple, different, utility functions, and using the approach in [69] or [70], the multiple, different, utility functions can be reduced to one utility at each follower level.

In Stackelberg game theory, the Stackelberg equilibrium is a popular solution concept, which is a stable point in the game where the leader/s and follower/s have no incentive to increase their payoffs without changing their strategy unilaterally, given all other player's strategies. The Stackelberg equilibrium is a hierarchal equilibrium, which means that, once the leader achieves a Nash equilibrium, the followers will then achieve a Nash Equilibrium thereafter.

Evolutionary Game

Evolutionary game theory extends from the traditional non-cooperative game, by considering a population as a set of players which can be either finite or infinite. Hence, evolutionary game theory was developed by biologists, to model and predict the interactions between different specie populations over time and in the presence of conflicting or cooperative objectives [61, 71]. An evolutionary game is defined as $G_{\mathrm{E}} = [\mathcal{N}, \mathcal{A}, \boldsymbol{x}, u(\cdot)]$. In this game, $\mathcal{N}$ is the population of players that is assumed to be large, $\mathcal{A}$ is the finite set of strategies for all players (pure strategy set), $\boldsymbol{x}$ is the population state, and $u(\cdot)$ is the utility function. The population state vector $\boldsymbol{x} \in \mathbb{R}^N$, where $\sum_{i \in \mathcal{A}} x_i = 1$, captures the percentage of the population choosing strategy type $i \in \mathcal{A}$. Each element x_i of $\boldsymbol{x}$ represents the average percentage of the population selecting strategy i. Moreover, in an evolutionary game, the evolutionary process is characterized by replicator dynamics. Replicator dynamics is used to model the evolution of a population selecting a particular strategy type. Depending on the problem that is being modeled, there are a range of replicator dynamics that can be used to capture the evolution of the population considered, such as imitative dynamics [61]. Over time, the population will update heir reference in $\boldsymbol{x}$ and will become more certain about what strategy type they would prefer [71]. For instance, in continuous time replicator dynamics, the rate of the population select strategy type i is proportional to the difference between the fitness of strategy type i and the average expected fitness of the population [61]. The fitness of a strategy type is defined as the average payoff (utility) of that type, and is a function of the population state $\boldsymbol{x}$. A popular replicator dynamics to model the evolution of strategy type i, is given by:

$$\frac{dx_i(t)}{dt} = \dot{x}_i(t), \tag{1.6}$$

$$\dot{x}_i(t) = x_i(t)\left(\overline{u}_i(\boldsymbol{x}, t) - U(\boldsymbol{x}, t)\right), \tag{1.7}$$

where $\overline{u}_i(\boldsymbol{x}, t)$ is the fitness of strategy type i, and $U(\boldsymbol{x}, t)$ is the average expected fitness of the population. The fitness of strategy type i, is defined as:

$$\overline{u}_i(\boldsymbol{x}, t) = \sum_{j \in \mathcal{A}} u_{i,j}(t) x_j(t), \tag{1.8}$$

where $u_{i,j}(t) = u_i(t)$ if $u_i \geq u_j$, or $u_{i,j}(t) = u_j(t)$ if $u_j > u_i$; and $x_j(t)$ is the population state of strategy type $j \in \mathcal{A}$. Furthermore, the average expected fitness

of the population, is given by:

$$U(\boldsymbol{x}, t) = \sum_{i \in \mathcal{A}} \overline{u}_i(\boldsymbol{x}, t) x_i(t). \tag{1.9}$$

In evolutionary game theory, another important solution concept is an evolutionary stable strategy (ESS). The ESS is a stable state in the evolutionary game, which is robust to a small portion of the population changing their strategy type (i.e., mutation) [61, 71]. In the context of biology, a mutation would occur if there was an invasion in the population. Moreover, the concept of ESS is stronger than the Nash equilibrium and is more robust.

Coalitional Game

Coalitional game theory, enables players to self-organise into cooperative groups (coalitions), in a fully distributed and autonomous manner [67]. A coalitional game is defined as $G_{\mathrm{C}} = [\mathcal{N}, v, \mathcal{F}]$. In this game, $\mathcal{N}$ is the finite set of players, v is the coalition value, and $\mathcal{F}$ is the set of coalitions (or the coalition structure). The set of coalitions $\mathcal{F}$ can be either a disjoint set or overlapping. In this thesis, we will focus on disjoint coalitions, where a coalition $\mathcal{F}_j \in \mathcal{F}$ is defined as a subset of $\mathcal{N}$, such that $\bigcup_{\mathcal{F}_j \in \mathcal{F}} \mathcal{F}_j = \mathcal{N}$, where $\forall \mathcal{F}_j, \mathcal{F}_{j'} \in \mathcal{F}$ is $\mathcal{F}_j \neq \mathcal{F}_{j'}$ then $\mathcal{F}_j \cap \mathcal{F}_{j'} = \emptyset$.

Coalitional games can either have a transferable utility function of non-transferable utility function. For transferable utility games, the coalition value v is a real number, defined by, $v : 2^{\mathcal{N}} \rightarrow \mathbb{R}$, which is divided among all players within the coalition [67]. The division of the coalition value depends on the problem that is being solved. For instance, if all players in a coalition contribute equally, then the coalition value can be equally shared amongst everyone. Whereas, if some players in the coalition contributed more, then they may receive a higher share of the coalition value compared to the rest of the players in the coalition. For non-transferable utility games, the coalition value v is a set of vectors capturing the individual player utilities within the coalition [67]. Thus, each player will have a utility function which is dependant on the actions of all other players in the coalition. The coalition value $v(\mathcal{F}_j)$ is given by, $v(\mathcal{F}_j) = \{\boldsymbol{u}(\mathcal{F}_j) \in \mathbb{R}^{|\mathcal{F}_j|} | u_i(\mathcal{F}_j), \forall i \in \mathcal{F}_j\}$, where $u(\mathcal{F}_j)$ is a nonempty vector of space $\mathbb{R}^{|\mathcal{F}_j|}$, and $u_i(\mathcal{F}_j)$ is the utility of user i in coalition $\mathcal{F}_j$ and is an element of $\boldsymbol{u}(\mathcal{F}_j)$.

In a coalitional game, players can either leave a coalition and join (form) a

new coalition, or stay in their current coalition. In order for a player to make such a decision, each player generates a preference-order over all coalitions. A preference-order is an ordered list of coalition preferences that each player prefers being a member of, a concept presented in [67]. A preference-order is a defined as a complete, reflexive and transitive binary relation over the set $\{\mathcal{F}_j \subseteq \mathcal{N} : i \in \mathcal{F}_j\}$ [72]. Preference-order is denoted by $\succ_i$, for any player $i \in \mathcal{N}$. Hence, for player $i \in \mathcal{N}$ to determine which coalition they prefer being a member of, they must calculate their utility in both coalition, as follows:

$$\mathcal{F}_j \succ_i \mathcal{F}'_j \Leftrightarrow u_i(\mathcal{F}_j) > u_i(\mathcal{F}_{j'}), \tag{1.10}$$

where $\mathcal{F}_j, \mathcal{F}_{j'} \in \mathcal{F}$. Equation (1.10) states that player i, strictly prefers being a member of coalition $\mathcal{F}_j$ over $\mathcal{F}_{j'}$, as player i obtains a better utility function (payoff) in coalition $\mathcal{F}_j$.

Nash stability is an important solution concept in coalitional game theory, which is an efficiency measure for coalitional games and is different to that of Nash Equilibrium, which is a best response condition for non-cooperative games. Thus, Nash stability is the strongest notion of stability for coalition partition, which ensures that no player has incentive to change to change their coalition formation [72, 73].

Matching Game

Matching games models the interactions between two disjoint sets of players. It is a powerful mathematical tool that is decentralized [74]. In fact, a matching game can model a one-to-one matching, a one-to-many matching, and a many-to-many matching. In the one-to-one matching game, one player from one set is matching to one player in the other set. On the other hand, in the one-to-many matching game, one player from one set is matched with a subset of players from the other set. In particular, matching games for wireless communications enables interactions between heterogenous devices, i.e., interactions between different devices in the network, such as CTDs and MTDs [75].

A matching game is defined as $G_{\mathrm{M}} = [\mathcal{N}, \mathcal{K}, \succ_n, \succ_k]$. In this game, $\mathcal{N}$ and $\mathcal{K}$ are disjoint sets of players, where each set of players are selfish and rational. $\succ$ is denoted as a preference relation, which is similar to the preference order defined in coalitional games, i.e., players from each set will have ranked preferences over players in the other set, and vice versa. Hence, in matching theory, there are many

other types of games, i.e., matching with transfer or matching with externalities.

In general, two-sided matching games have relied on the deferred acceptance algorithm. The deferred acceptance algorithm is an iterative process, where players in one set make proposals to the players in the other set. Players in this game, will make their matching decision based on their individual preferences and maximising their utility/payoff. Moreover, the deferred acceptance algorithm is guaranteed to converge to a stable matching [75].

1.4 Thesis Contributions and Organisation

The main focus of this thesis is on distributed resource allocation in multimedia IoT, and using game theory to develop novel solutions for decision making processes required to address the key challenges and area of multimedia IoT. In particular, the three key enabling technologies of multimedia IoT that we will consider are: D2D communications; M2M communications; and enabling UAVs as low flying BSs. The main body of the thesis consists of four technical chapters, as outlined in Fig. 1.5. The first two technical chapters of this thesis investigates interference management, resource allocation, and end-user satisfaction challenges that arise in D2D communications. The last two technical chapters of this thesis investigates the challenges relating to M2M communications, i.e., data correlation for a massive number of MTDs, and enabling UAVs as low flying BSs in mission critical IoT communications.

A detailed summary of the major contributions of this thesis are as follows:

Chapter 3 - Flexible D2D Application-Driven Resource Allocation

In Chapter 3, we address the fundamental problem of flexibility for application-driven distributed resource allocation, to provide significantly improved D2D user QoE in cellular networks for underlaid D2D communications. In particular, we formulate this problem as a multiple objective Stackelberg game, in order to optimise throughput, energy consumption, and efficient resource block allocation across all users in the network, while providing greater reliability and user satisfaction guarantees.

The main contributions of this chapter are:

- A flexible resource allocation scheme for D2D communications using a Stack-

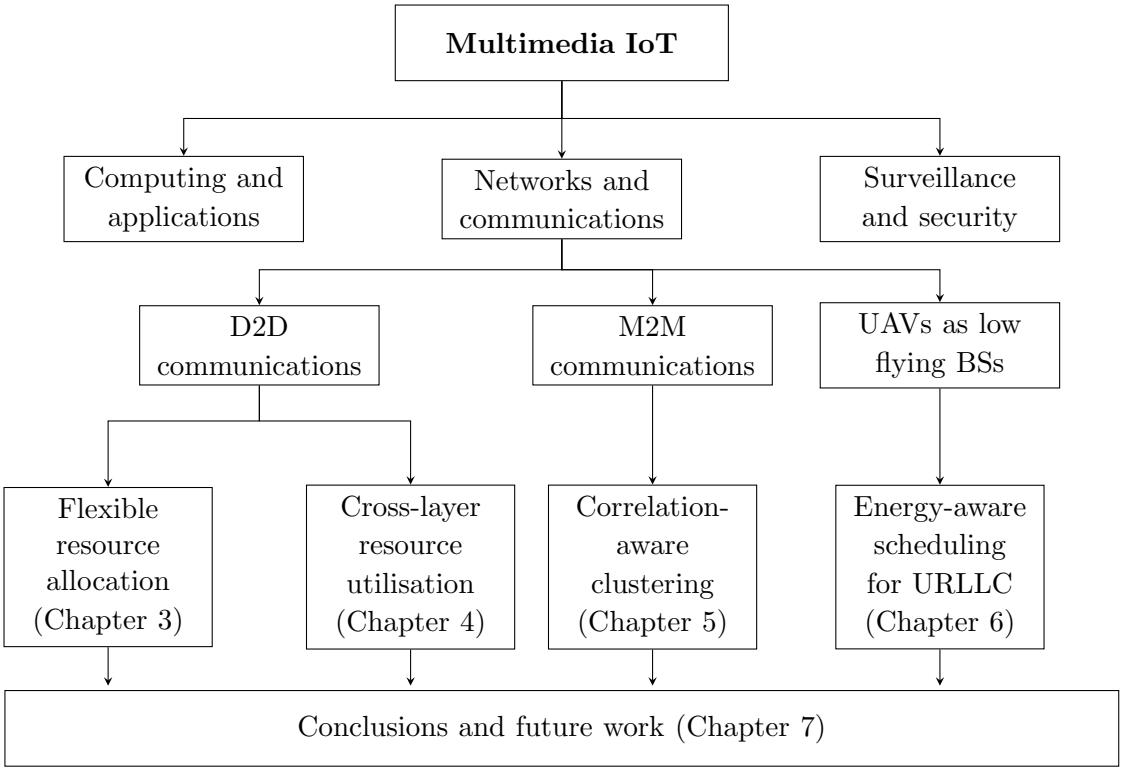


Figure 1.5: The thesis outline.

elberg game is proposed. We consider a single leader and multiple followers, where the BS is the leader and the D2D pairs are the followers. The BS charges the D2D pairs a fee for network (BS) satisfaction/certainty with respect to all users (both cellular and D2D) in the network, and a fee for reusing a given cellular user's resource block, to coordinate and reduce the intra-cell interference. Whereas, the D2D pairs react to the leader's charging price, and compete to find optimal transmit power and resource block allocation, while guaranteeing QoS and QoE.

- To enhance D2D user QoE and provide greater flexibility, D2D users are categorised into one of three application classes, based on their practical application or service: (i) *casual class*; (ii) *interactive class*; and (iii) *streaming class*. Each class formulates a criterion set and is mapped to a different utility function. Hence, we propose a crucial innovation, where the several, different, utility functions available to followers, are solved using a non-scalarised solution, as a scalarised approach, to this multi-criteria optimisation problem, is generally infeasible in real-time D2D cellular communications.

- We demonstrate beneficial effects of prioritising communications, along with the implicit prioritisation from different utility functions, by setting different PDR targets.

- A distributed algorithm based on a dynamic Stackelberg game with multi-criteria decision making is proposed. The proposed algorithm is shown to converge to a unique Stackelberg equilibrium across all users, which is a sub-game perfect equilibrium for each game stage. Moreover, simulation results show that the proposed approach reduces transmit power effectively and increases throughput, while also guaranteeing best social welfare and satisfaction across all D2D users.

The results in this chapter have been presented in [76], which is listed again for ease of reference:
[76] **N. Sawyer**, and D. B. Smith, "Flexible Resource Allocation in Device-to-Device Communications using Stackelberg Game Theory," *IEEE Trans. on Communications*, vol. 67, no. 1, pp. 653–667, Jan. 2019.

Chapter 4 - D2D Mode Selection and Resource Allocation in D2D Networks

In Chapter 4, we jointly optimise mode selection, resource allocation, and interference management for D2D communications, in a distributed manner. We investigate a non-cooperative game and a coalitional game to solve the network-assisted joint optimisation problem. In the non-cooperative game, the BS will assist D2D pairs in selecting a transmission mode (i.e., either cellular mode or reuse mode), while also allocating resources to cellular and D2D users. However, in the coalitional game, we extend the mode selection to consider all three transmission modes (i.e., either cellular mode, dedicated more, or reuse mode). Network-assisted D2D communications requires global network information, and can cause signalling overhead and network complexity to increase due to the large number of users in the network. Thus, we also investigate user-assisted D2D communications as a viable solution, in which the BS has partial knowledge of the network and partial CSI. Thus, the D2D pairs must determine their transmission mode without the assistance of the BS. Hence, we extend the coalitional game to consider the user-assisted approach.

The main contributions of this chapter are:

- We propose a non-cooperative cross-layer repeated game, which combines a non-cooperative power control game and a two-armed bandit game, for network-assisted D2D communications. The non-cooperative power control game aims to minimise transmit power for each user, which leads to a reduction in the intra-cell interference. Whereas, the two-armed bandit games aims to select the best transmission mode for D2D pairs.

- A distributed algorithm based on the non-cooperative cross-layer repeated game is proposed. The proposed non-cooperative algorithm is shown to converge to a unique Nash equilibrium, that is Parteo-optimal.

- On the other hand, we proposed a distributed cross-layer coalitional (CLC) game-theoretic approach, to cluster D2D pairs based on resource block allocation and D2D mode selection, in a fully distributed and autonomous manner. This approach provides insight into practical scenarios when considering all possible D2D transmission modes.

- To further reduce interference in the network, the proposed CLC game-theoretic approach incorporates scheduling D2D pairs reusing cellular user resources along with power control. That is, we aim to find an optimal number of D2D pairs that can reuse a cellular user resource block at a given time while ensuring user performance.

- Furthermore, we investigate the impacts of network-assisted and user assisted D2D communications, when clustering D2D pairs. The network-assisted and user-assisted scenarios are analysed in terms of minimising transmission power, user throughput, and spectral efficiency.

- A fully distributed clustering algorithm based on the CLC game is proposed to find a stable cluster formation of D2D pairs. The proposed clustering algorithm converges to a Nash stable coalition partition that is socially efficient and has a non-empty core.

- Extensive simulation results validate the effectiveness of the proposed non-cooperative cross-layer repeated game and the proposed CLC game. The performance of the proposed non-cooperative and coalitional games are compared in terms of total transmission power consumption in the network and total network throughput.

The results in this chapter have been presented in [77, 78] and [79], which are listed again for ease of reference:
[77] **N. Sawyer**, and D. B. Smith, "Pareto-Efficient Cross-Layer Repeated Game for Device-to-Device (D2D) Communications," in *Proc. IEEE Int. Conf. on Communications (ICC)*, Kuala Lumpur, Malaysia, pp. 1–6, May 2017.
[78] **N. Sawyer**, and D. B. Smith, "A Nash Stable Cross-Layer Coalition Formation Game for Device-to-Device Communications," in *Proc. IEEE Int. Conf. on Communications (ICC)*, Paris, France, pp. 1-6, May 2017.
[79] **N. Sawyer**, and D. B. Smith, "A Nash Stable Cross-Layer Coalitional Game for Resource Utilization in Device-to-Device Communications," *IEEE Trans. on Vehicular Technology*, vol. 67, no. 9, pp. 8608–8622, Sept. 2018.

Chapter 5 - Correlation-Aware Clustering in M2M Networks

Chapter 5 focuses on the self-organising, correlation-aware clustering for a dense network of MTDs deployed over a cellular network. In dense M2M networks, MTDs are typically located within close proximity and will gather correlated data, and, thus, clustering MTDs based on data correlation will lead to a decrease in the number of redundant bits transmitted to the base station. We formulate the correlation-aware clustering problem for a massive number of MTDs, as an evolutionary game.

The main contributions of this chapter are:

- We propose a novel distributed correlation-aware clustering scheme for a massive number of MTDs. The proposed correlation-aware clustering scheme is formulated using an evolutionary game, in which MTDs are clustered based on data correlation, in order to decrease MTD transmission power by reducing the number of redundant bits. We investigate the impact on MTD cluster formation, for different MTD densities, MTD data correlation, and the worst case inter-cluster interference.

- To define the utility function for the proposed evolutionary game, we first derive a closed-form upper bound expression for inter-cluster interference. Then, based on this interference analysis, we introduce a novel utility function that captures the average MTD transmission power per cluster, while mitigating the preference between larger or smaller cluster size according to the cost of signalling overhead.

- We introduce a new distributed algorithm based on an evolutionary game to find the stable cluster formation specifically for a massive number of MTDs. The robustness of the proposed distributed algorithm is analysed in terms of the maximum number of MTDs that can change their cluster formation. We derive the maximum portion of MTDs that can deviate from the ESS, while still maintaining a stable cluster formation.

- The accuracy of the stochastic geometry analysis as well as the effectiveness of the game-theoretic approach are corroborated by extensive simulations. In particular, simulation results show that, as the network density and data correlation increase, the proposed distributed algorithm determines the number of MTDs within each cluster, based on network density and cluster radius. Moreover, the simulation results verify the tradeoff between cluster size and transmission power per MTD per cluster, given the network density and correlation constant.

The results in this chapter have been presented in [80] and [81], which are listed again for ease of reference:
[80] **N. Sawyer**, M. Naderi Soorki, W. Saad, and D. B. Smith, "Evolutionary Coalitional Game for Correlation-Aware Clustering in Machine-to-Machine Communications," in *Proc. Global Communications Conf. (GLOBECOM)*, Singapore, pp. 1-6, Dec. 2017.
[81] **N. Sawyer**, M. Naderi Soorki, W. Saad, D. B. Smith, and N. Ding, "Evolutionary Games for Correlation-Aware Clustering in Massive Machine-to-Machine Networks," *IEEE Trans. on Communications*, May 2019.

Chapter 6 - Utilising UAVs in Mission Critical M2M Communications

Chapter 6 focuses on a novel framework that addresses energy-aware scheduling for mission critical M2M communications, utilising UAVs as flying BSs. In particular, we consider a multimedia IoT network in which MTDs are clustered and aggregators are employed as cluster heads to collect the sensed data from MTDs. In the considered IoT network, UAVs must optimise their schedule in order to collect the data from each aggregator in a distributed manner, without requiring a-priori knowledge of the probabilities associated with the event-driven arrival requests from MTDs. Hence, we formulate the ultra reliable energy-aware scheduling problem by combining both Lyapunov optimisation and matching theory.

The main contributions of this chapter are:

- The use of Lyapunov optimisation is appropriate as it does not require a-priori knowledge of the event-driven traffic of MTDs. Using the Lyapunov optimisation technique, the problem is transformed into a linear weighted function, which ensures the maximum queue length and the queue length variance across all aggregators in the network are bounded.

- Based on the Lyapunov optimisation of our problem, matching theory is used to formulate and model the energy-aware scheduling problem as a one-to-many matching game. In the formulated game, each UAV will determine a subset of aggregators that it will serve and sufficient dwelling time in order to satisfy the URLLC constraint, whereas, the aggregators will minimise their energy consumption when transmitting their data to a UAV.

- A fully distributed algorithm based on a one-to-many matching is proposed to find the two-sided stable matching. Simulation results show that, the maximum queue length mean and variance, across all aggregators, are bounded by finite threshold values, which satisfy the URLLC constraints for mission critical M2M communications. The results show that, as the transmission probability of packet arrival and maximum queue length thresholds increase, the minimum number of required UAVs to service cluster heads (aggregators) in the network decreases. Moreover, a trade-off is observed between reducing aggregator transmission power and increasing network throughput, when varying the tunable parameter from the Lyapunov optimisation.

The results in this chapter have been presented in [82], which is listed again for ease of reference:
[82] **N. Sawyer**, M. Naderi Soorki, W. Saad, N. Ding, D. B. Smith, M. H. Manshaei, and M. Mozaffari, "Energy-Aware Scheduling for Ultra-Reliable Internet of Things Communications using Unmanned Aerial Vehicles," (under preparation to be submitted to *IEEE Internet of Things Journal*).

Finally, Chapter 7 draws conclusions from this thesis and provides suggestions for future research directions arising from the work presented in the technical chapters (Chapters 3-6).

Chapter 2

Literature Review

In this chapter, we review the relevant literature covering the three key areas of multimedia IoT addressed in this thesis. Thus, we first review existing work on resource management and end-user satisfaction in underlaid D2D communications. Then, we discuss relevant work on data correlation and machine-type device (MTD) clustering in M2M networks, and finally, we analyse existing work on ultra-reliability low latency communications (URLLC) in M2M communications, in the context of UAVs being utilised as low flying BSs to service MTDs. Moreover, in each section of this chapter, we outline the limitations of existing work and state our thesis contribution for each particular area.

2.1 Resource Management and End-User Satisfaction in D2D Communications

The introduction of D2D communications in current and future cellular networks, enables users who are within close proximity to communicate directly with one another. D2D mode selection enables D2D users to determine which transmission mode to operate in, that is, D2D pairs need to determine when they should transmit directly or when they should transmit via the cellular link. Determining the best transmission mode for a particular D2D pair, is an important decision, which will ensure D2D pairs operate more efficiently, and thus, increase system capacity, spectral efficiency, and network efficiency [83]. The three transmission modes are: cellular mode, dedicated mode, and reuse mode. For D2D pairs in cellular mode or dedicated mode, the fundamental research challenge that has been investigated is resource allocation. On the other hand, for D2D pairs in reuse mode which shares/reuses cellular spectrum, the fundamental research challenge is resource al-

location and interference management. Hence, there is a need to adopt distributed mode selection, resource allocation and interference management schemes, in order to coordinate and reduce the intra-cell interference between users in the network, while also ensuring D2D users operate more efficiently, guaranteeing end-user satisfaction, and are allocated resource blocks fairly.

2.1.1 D2D Transmission Mode Selection

For D2D pairs to determine a suitable transmission mode to operate in, they base their decision on two factors, i.e., the measured performance gain they will receive and information about the network. Typically, the BS will assist D2D users in determining a particular transmission mode to operate in, which is referred to as network-assisted D2D communications [16]. Thus, the BS will help D2D users by providing additional network information, as well as coordinating the D2D pairs and cellular users within the network to reduce severe intra-cell interference. The type of network information that the BS will provide to the D2D users will include, the physical distance between the D2D pair transmitter and receiver, and the complete CSI, as well as the interference that will be received. Moreover, mode selection can also be based on the D2D pair's performance, i.e., the D2D pair will prefer to operate in a particular transmission mode if its throughput is increased or its energy consumption is reduced.

Within literature, there is a wide range of existing D2D mode selection techniques, such as in [16, 33, 84–88] and [89]. Distance dependant mode selection is one of the simplest mode selection approaches, where the distance between the D2D pair transmitter and receiver must be within a threshold distance to transmit directly to each other, otherwise the D2D pair will operate like traditional cellular users. The work in [84] proposes a mode selection approach based on a threshold distance in order to minimise transmission power. The proposed mode selection approach in [84], found that the distance threshold was inversely proportional to the BSs density, as well as being strictly increasing with the path loss exponent. Meanwhile, the work in [86] considers three factors for mode selection which are, the distance between the D2D pair transmitter and receiver, link quality, and a bias factor. Thus, the proposed approach in [86] enables D2D pairs to only transmit directly between one another if their channel link quality is at least the same quality as the cellular user uplink. On the other hand, in [88] a mode selection

technique based on D2D and cellular link quality, and interference was proposed. In particular, the interference considered in the mode selection approach in [88] also factors in the intra-cell and inter-cell interference. In contrast, the work in [89] considers a mode selection approach based on optimising the sum rate of D2D communications, and takes into account the cellular, dedicated, and reuse modes.

As outlined in Section 1.3.2, game theory is a powerful mathematical tool that enables decision making and models the interactions between users. Game theoretic techniques have been considered for dynamic mode selection, such as in [77,90–95], and [96], to model the interactions between D2D pairs wishing to change their transmission mode. In the aforementioned works, the mode selection problem has been formation using a two-armed bandit game, a coalitional game, or an evolutionary game. The work in [90] proposed a two-armed Levy bandit game for D2D mode selection, where mode selection is based on maximising an expected reward, which is a function of the strategy the D2D pair chooses and its utility. On the other hand, the work in [93] proposes a coalitional game for energy-efficient mode selection, where D2D pairs in the same transmission mode can cooperate and share information. Thus, in [93], the energy-efficient mode selection approach focuses on minimising transmission power while guaranteeing rate requirements. Meanwhile, the work in [96] proposed an evolutionary game for D2D mode selection, where replicator dynamics are used to model the preferences of D2D pair's transmission mode based on the utility function. Therefore, game theoretic approaches enable D2D pairs to dynamically change their transmission mode, only if their payoff is increased.

In fact, existing works, such as in [91–95, 97–99] consider a distributed mode selection scheme, however, D2D pairs have only been considered to operate in either reuse mode or cellular mode. In addition, these works only consider one D2D pair to reuse only one or more cellular user's resource block. However, in practice, these approaches are not ideal, as spectral efficiency and network performance cannot be maximised due to limiting the number of D2D pairs reusing cellular user resources. A limited number of works, such as [36, 93], and [100] have considered D2D pairs operating in either reuse mode, cellular mode, or dedicated mode. In particular, when considering dedicated mode, coordination is required between D2D pairs and the BS, as the BS needs to assess if there are any free resource blocks available to be allocated to D2D pairs [36]. This coordination can cause signalling overhead to

increase, however, D2D pairs operating in dedicated mode will have ultra-reliable communications, which, in turn, will improve overall network performance and throughput.

In most of the works described thus far implement network-assisted D2D communications, where the BS assists D2D users in determining which transmission mode to operate in. Thus, in network-assisted D2D communications, it is assumed that the BS must have global knowledge of the network, such as, complete CSI and the location of all users [16,31–33]. In order for the BS to acquire global knowledge of the network, this will cause signalling overhead and network complexity to significantly increase, especially with the increasing number of users in the network. Furthermore, in practical scenarios, network-assisted D2D communications is not ideal, as complete CSI cannot always be achieved for all users all the time.

Hence, as the network begins to evolve towards multimedia IoT, users will start to become more self-sufficient and intelligent. Thus, the BS won't need to provide information to users all the time, i.e., the BS will not be required to have full knowledge of the network. The work in [36] considers a D2D cognitive communication system with assistance-free D2D mode selection, i.e., assuming partial CSI, that analyses the network performance and and the impact of direct transmission between D2D users on the cellular network. Whereas [100] jointly considers mode selection, user scheduling, and rate adaptation, with partial CSI, to ensure user fairness and reduced interference in the network. Even though some of the proposed works, such as, [36,101], and [100], consider limited/partial CSI, they do not jointly consider mode selection, resource allocation, and interference management for D2D communications in a distributed and autonomous manner.

2.1.2 Interference Management and Resource Allocation

Once the transmission mode is determined for a D2D pair, interference management and resource allocation need to be addressed, in order to guarantee each user's QoS within the network. D2D pairs in reuse mode will cause harmful interference to both cellular and D2D users, due to reusing a cellular user's uplink resource block/s. Thus, D2D pairs in reuse mode need to be allocated a suitable cellular user's uplink resource block/s in order to reduce the impact of the intra-cell interference. One approach to do this, is to enable a D2D pair in reuse mode to reuse a cellular user's resource block who is located far away within the network. Thus the interference

the cellular user will receive will be reduced (not mitigated). On the other hand, D2D pairs in cellular mode or dedicated mode, will not cause interference in the network, however will need to be allocated orthogonal resource blocks. Within literature many works have been proposed to address the resource allocation and interference management problem in underlaid D2D communications, as in [34, 35, 65, 77, 83, 86, 87, 89, 91–99, 101–107]. Hence, addressing the resource allocation and interference management problem, will lead to reduced intra-cell interference, and increased spectral efficiency and energy efficiency.

In [87], a joint mode selection and resource allocation scheme was proposed, that optimises the tradeoff between maximising sum rate and increasing the number of supported D2D users in the network. The work in [35] investigated the problem of energy efficient resource allocation in D2D communications, while taking into account transmission power and QoS constraints, as well as mutual preferences of cellular and D2D users. In [102], a cognitive D2D communication system was considered, in order to improve resource allocation and spectral efficiency. Additionally, the work in [103] considers an energy efficient power allocation scheme for clustered D2D pairs, to reduce interference between users, minimise energy consumption, and maximise sum rate.

2.1.2.1 Game Theoretical Approaches

Game theory has been widely used to address resource allocation and interference management in wireless communications, as in [34, 35, 65, 77, 83, 89, 91–99, 101, 104, 107–121]. Recall, from Section 1.3.2, game theory is a decision making process and models the interactions between users competing for resources in a fully distributed and autonomous manner. A number of works have considered non-cooperative power control games, as in [83], [113] and [114], auction-based games, as in [34], or Stackelberg games, as in [97, 104, 108–112, 115–119], and [120], to maximise data rate and reduce intra-cell interference in the network. In particular, power control game theory is a popular approach, which is used to minimise transmission power of users in the network, which in turn, leads to a reduction in intra-cell interference. In [77], a non-cooperative mode selection and power control game was proposed, to reduce intra-cell interference, increase communication reliability, and assist D2D pairs selecting optimal transmission mode. On the other hand, in [34], the authors proposed a joint channel and power allocation iterative combinatorial auction game,

that optimises the tradeoff between transmission rate and battery lifetime for each user.

Modelling the resource allocation and interference management problem in D2D communications as a Stackelberg game, allows the base station (BS), cellular users, and D2D pairs to coordinate the intra-cell interference within the network, as in [97, 104, 108–112, 115–120]. A number of works such as in [109], [110], and [111] have considered distributed power control using a Stackelberg game framework. The goal of these works is to reduce the interference in the network and solve for optimal transmission power. The work in [104] and [97], aims to coordinate and efficiently manage the intra-cell interference using Stackelberg game theory, such that network throughput is maximised while guaranteeing cellular user performance. Additionally, the work in [117] jointly considers a heuristic channel assignment based on the adaptive interference restricted region, and distributed power control scheme using a Stackelberg game. On the other hand, the work in [119] proposes a distributed resource allocation Stackelberg game, where the D2D pairs (followers) jointly learn the resource and power allocation using an uncoupled stochastic learning algorithm, while coordinating the interference in the network. In [118], a distributed power control and interference management scheme is proposed, where novel uniform and differentiated interference pricing is proposed in the Stackelberg game. Furthermore, the work in [120] proposes a discount interference pricing scheme for D2D pairs deciding to cooperate with nearby cellular users. Meanwhile, the work in [116] considers cluster-oriented D2D communications, where power allocation is formulated as a Stackelberg game, and bandwidth allocation and link selection are then jointly considered, in order to increase network data rate.

Moreover, a Stackelberg game can be classed as a bi-level multi-objective optimisation problem (MOP), which are a subset of multi-level MOPs. Multi-level (and bi-level) MOPs consist of a leader level and multiple follower levels, where several objective (utility) functions are available to followers [69, 70]. Up-till now scalarisation has generally been used to solve the multi-level MOP by combining all the separate objective functions into one objective (utility) function with multiple objectives, at each follower level. The work in [69] investigates a multi-level MOP with multiple followers, where several objective functions available to followers are combined into one function by linear scalarisation and nonlinear scalarisation. However, the linear or nonlinear scalarisation approach in a MOP, puts a restric-

tion on the amount of diversity in design of multiple objective functions. Therefore, each objective function cannot be chosen arbitrarily, if scalarisation is used. On the other hand, a limited number of works have considered non-scalarisation methods to solve the multi-level MOP, with several, different, objective functions available to followers. The work in [122] investigates a non-scalarisation method to solve a multi-level multi-objective decision making problem with multiple followers. However, the non-scalarisation method considered in [122], does not allow followers with similar objectives to optimise their objective functions competitively and non-cooperatively. Thus, the non-scalarisation approach represents a more practical and tractable scenario for D2D communications.

In particular, existing work on current Stackelberg games for D2D communications [97, 104, 108–112, 115–120] consider one utility function assigned to the leader/s, which may or may not be different to the one utility function assigned to the follower/s. Thus, such Stackelberg games for D2D communications have been fundamentally constrained, due to only limiting followers to optimising one particular utility function across all D2D users, which does not necessarily guarantee all D2D users' QoE. If this constraint was removed, as in multi-level Stackelberg games [69], then different utility functions suitable to particular service requirements (e.g., energy, or throughput) could be made available to better enable D2D user applications. Yet, designing individual utility functions, for each possible D2D application, would result in a large number of utility functions. Thus, categorising D2D pair applications into classes based on similar requirements, as done in [123] and [124], reduces the number of multiple, different, utility functions available to followers (D2D users), in an effective manner. Although [123] and [124] consider defining different classes for D2D user satisfaction, they do not jointly consider QoE, QoS, resource allocation, and interference management, as well as signalling overhead, for D2D communications in a fully distributed and autonomous manner. Hence, the advantages of providing followers with several utility functions in a Stackelberg game are: 1) network problems such as dynamic resource allocation and interference management can improve D2D users' experience by properly accounting for the desired service each user; 2) utility functions can be tailored for specific groups of practical applications, where, for example, the desired outcomes from a utility for content/data sharing would be different to a utility for gaming; and 3) cellular networks will have a more flexible communication infrastructure,

where D2D user satisfaction/QoE is better accounted for and enhanced.

Meanwhile, clustering D2D pairs based on cooperative game theoretic techniques, has recently emerged as a promising technique to improve spatial reuse, energy efficiency, and cooperation between users in a distributed and autonomous manner [35, 91–96, 99, 101]. The goal of these works is to increase energy efficiency, improve fairness, and maximise channel rate, while cooperatively sharing the spectrum in order to reduce intra-cell interference and guarantee cellular user requirements. The work in [94] and [95] proposed a distributed resource sharing scheme, in order to maximise channel rate and improve network performance. In [96], a clustered cognitive D2D communication system was proposed, where an evolutionary game was used to assist D2D users to select a transmission mode, in order to increase spectral efficiency. Moreover, in [101], a joint subchannel and power allocation scheme is considered, based on a cooperative Nash bargaining game, that optimises the tradeoff between maximising sum rate and fairness.

2.1.2.2 Scheduling Based Approaches

D2D pair scheduling has attracted considerable interest in order to reduce intra-cell interference in the network, as studied in [106, 125–131]. Two popular scheduling schemes have been proposed for distributed D2D pair scheduling, which are flash link quality (FlashLinQ) [125] and information-theoretic link quality (ITLinQ) [126]. These D2D pair scheduling schemes focus on scheduling D2D links within the network based on interference between users. Although D2D link scheduling will reduce the number of transmitting D2D pairs at a particular time, this will not only cause intra-cell interference to reduce, but also system capacity and network throughput. Hence, existing work on incorporating D2D pair scheduling as an interference management scheme along with power control in D2D communications remains limited. The work in [127, 129–131], jointly considers D2D link scheduling with power control, in order to reduce intra-cell interference, while also increasing the number of supported D2D pairs. However, most of these works [127, 129–131] do not consider other factors such as resource allocation, and mode selection.

2.1.3 End-User Satisfaction

There is a wide range of potential applications for D2D communications, such as, messaging, content sharing (such as videos and photos), local voice calling, gaming

(single or multi-player), data sharing, advertising, uploading/downloading/streaming content, and emergency services (local security and safety communications between neighboring police, firemen, and ambulance officers) [2, 37, 38]. Depending on the D2D application, particular factors such as, reliability, priority, throughput, or transmit power, can be enhanced in order to increase each D2D user's QoE. Guaranteeing D2D user QoE ensures end-user satisfaction, and it is becoming an essential quality of measure, alongside QoS, in current and future cellular networks [39–41].

Meanwhile, the idea of guaranteeing QoE for D2D pairs under different applications has recently emerged as a promising technique to ensure D2D users are satisfied, as outlined in [6, 39–41, 123, 124, 132–138]. In current and future cellular networks, guaranteeing QoE is becoming an essential quality of measure, alongside QoS [6, 39–41]. D2D user satisfaction can be expressed as a D2D user's perceived benefit, meaning that D2D user satisfaction is the measure of benefit that a particular user feels [6, 39, 123, 124, 133]. For example, if low data rate was provided to a D2D user whose application is uploading content, then this would have a negative effect on user satisfaction, when compared to providing high data rate. Thus, D2D user satisfaction can be measured in the following ways: mean opinion score [39, 124]; a satisfaction function [40, 41]; guaranteeing QoS [132]; or ensuring high data rate is provided [123]. Additionally, user-in-the-loop (UIL) is another approach that aims to aid user's experience by reducing traffic and delay within the network, by controlling/influencing user behaviour/requirements, as studied in [133–138]. Furthermore, the work in [123] and [124] characterizes D2D user satisfaction (QoE) into different classes, based on achievable data rate and media service delivery, respectively, and defines specific QoE functions for each class. The work in [123] studies the possibility of joint resource allocation, in order to maximise the average system satisfaction degree. However, [123] relies on an artificial fish swarm approach, which is not a practical approach for large-scale D2D communications, as it may lead to increased signalling overhead and slow convergence speed [139]. Whereas, [124] proposes a distributed media service delivery and resource allocation scheme for D2D communications, to address the massive media content dissemination problem in cellular networks. Even though some of the proposed works, such as, [123] and [124], consider defining different classes for D2D user satisfaction, they do not jointly consider QoE, QoS, resource allocation, and interference management, as well as signalling overhead, for D2D communications

in a fully distributed and autonomous manner.

2.1.4 Summary

Therefore, this thesis aims to investigate resource allocation, interference management, and end-user satisfaction for D2D communications, using a multiple objective (equivalently multi-criteria) Stackelberg game, without any scalarisation across the different objectives. In addition, this thesis also investigates the joint optimisation of mode selection and resource allocation using non-cooperative and cooperative game theory, in order to reduce the intra-cell interference and improve energy-efficiency, where a mode selection scheme scheme is considered with partial CSI.

2.2 Data Correlation in M2M Communications

Recently, the idea of clustering MTDs into smaller groups has emerged as a promising technique to reduce the traffic load on the cellular BS and improve spatial reuse and energy efficiency, while reducing interference in the network, as studied in [18, 26, 46], and [47]. Fig. 2.1 and 2.2 outline an illustrative example of an M2M network topology, without clustering and with clustering. In particular, a clustered M2M network, as outlined in Fig. 2.2, this will effectively reduce the number of MTDs transmitting to the BS. Existing clustering techniques for M2M communications [26, 46, 47, 80, 140–148] have focused on clustering MTDs based on resource allocation, location, load on the random access channel (RACH), and data correlation. Clustering has been considered in literature, as an effective approach to alleviate the potential massive congestion caused by MTDs, as done in [140–142] and [143]. These aforementioned works aim to maximise the number of MTDs that attempt to simultaneously access the BS, while minimising network congestion, the load on the RACH and signalling overhead [141, 142]. In [140], an energy-efficient cluster formation (load adaptive multiple access scheme) and cluster head selection scheme was proposed, to maximise network lifetime in a massive M2M network. The work in [142], investigated the problem of random access contention between cooperative groups of MTDs that coordinate their random access channel, while taking into account energy consumption and time varying queue length. However, the algorithm presented in [142] cannot cope with a massive number of MTDs, as its complexity will grow significantly. On the other hand, clustering techniques based

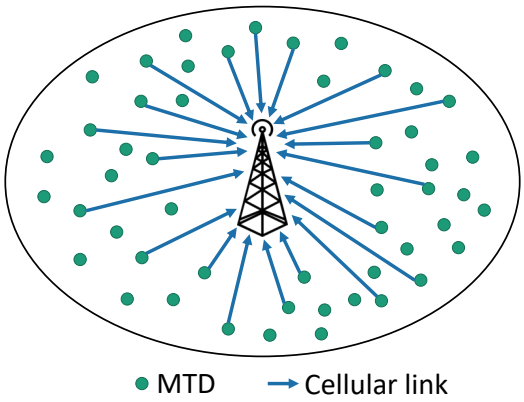


Figure 2.1: M2M network topology, without clustering

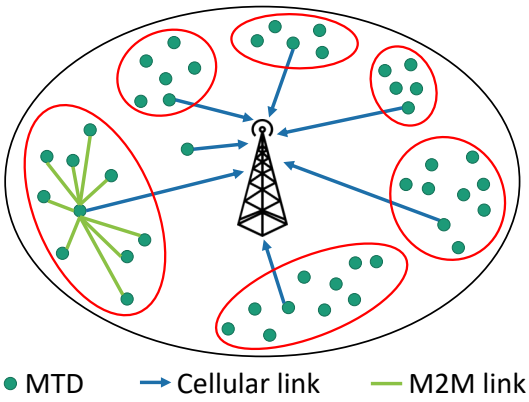


Figure 2.2: M2M network topology, with clustering

on the QoS requirements and locations of MTDs, are proposed in [26], [144–147] and [148], in order to maximise the number of supported MTDs. A cluster prioritisation scheme for massive access management is studied in [26], where MTDs are clustered based on QoS requirements. The work in [144] proposes a self-organised cluster formation mechanism in which MTDs form clusters with neighboring MTDs. In addition, a number of works such as in [145–147, 149], and [150] have also considered joint clustering and resource allocation. The goal of these works is to maximise MTD data rate, allocate resource blocks efficiently, reduce interference to the cellular network, and optimise the battery lifetime of MTDs. The work in [149] proposes a distributed resource (time) allocation scheme, to address the diverse QoS requirements in an IoT network, while taking into account data rate of CTDs and energy consumption of MTDs. The aforementioned works [26, 140–150] consider machine-centric clustering approaches, that cluster MTDs in order to maximise MTD data rate and number of supported MTDs, while minimising energy consumption. Such a machine-centric clustering approach does not take into account the individual data/information of each MTD. Additionally, in [151], the authors proposed a clustering approach that uses a coalitional game to optimise the tradeoff between sum-rate gains and power costs.

Meanwhile, the dense deployment of MTDs in M2M networks, will enable MTDs within close proximity to gather correlated data, thus often sending the same information (redundant bits) to the BS (e.g., see [44,47] and [152]). Hence, a *data-centric clustering approach* can be used to improve the data quality sent to the BS. Existing work on clustering MTDs with respect to data correlation remains limited.

Primarily, the works in [46] and [47] have studied the possibility of MTD clustering based on location and correlation, however, these works rely on centralized approaches that are not practical for large-scale M2M networks. Such centralized clustering approaches can lead to significant signalling overhead as they require gathering of global information, such as location and data correlation factors, for a large number of MTDs. Indeed, in practice, centralized clustering approaches are not robust to the dynamic changes in the MTD networking environment, such as the joining of new MTDs, MTD loss of battery, or rapid fluctuations in the sensing environment. Thus, there is a need to introduce new distributed correlation-aware clustering approaches.

To develop such distributed solutions for cooperation in wireless network, it is customary to resort to tools from game theory [61,67]. Particularly, in [80,142,149], and [153], game theory has been used for distributed cluster formation in M2M networks. The work in [153] proposed a distributed correlation-aware cell association algorithm, that maximises the information sent to the BS while maximising the number of assigned IoT devices to every BS. The game considered in [153] is a two-sided matching game, however the proposed solution cannot cope with dynamic changes in the M2M network environment. Indeed, the existing works in [26, 46, 47, 80, 140–148], and [153] consider only clustering a small, finite number of MTDs. However, this is not the case in practical IoT scenarios as the number of MTDs within the network is massive, which causes substantial interference and impacts the way in which correlation-aware clustering must be performed. Meanwhile, the aforementioned works are also not robust to the dynamics of a large-scale M2M network that results from factors such as the arrival or departure of new MTDs, or the deactivation of MTDs (e.g., due to battery loss, or rapid fluctuations in the sensing environment). In particular, in [80], we have developed an evolutionary coalitional game for correlation-aware clustering, for a finite number of MTDs. However, our proposed distributed algorithm in [80] cannot cope with a massive number of MTDs, as its complexity will grow significantly. Furthermore, [80] also relies on a simplistic utility function, that does not capture the real-world deployment of MTDs.

Therefore, this thesis aims to investigate a distributed correlation-aware clustering approach for a finite and massive number of MTDs, while ensuring low signalling overhead and robustness for small stochastic changes in the M2M environment.

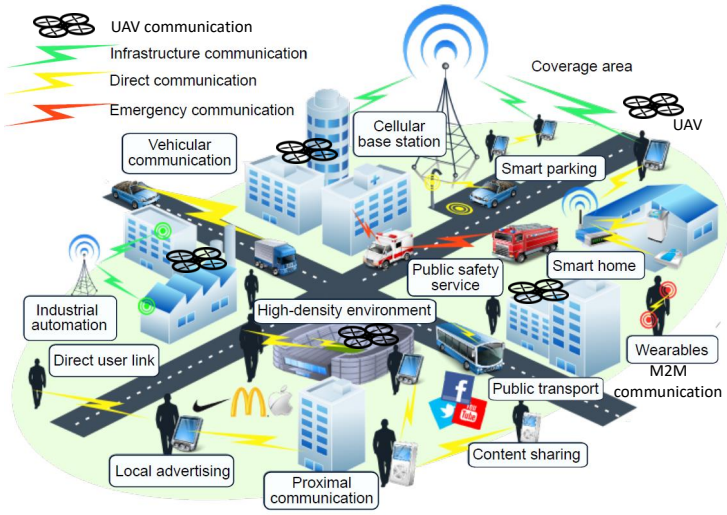


Figure 2.3: Use cases of multimedia IoT communications, including UAVs. This figure is a modification of Fig. 1 in [2].

2.3 Enabling UAVs as Flying BSs for Mission Critical M2M Communications

Recently, the use of UAVs as flying BSs and mission critical M2M communications have attached considerable attention in multimedia IoT. Fig. 2.3 outlines an illustrative example of multimedia IoT use cases, with UAV BSs. However, within literature these two hot topics, UAV BSs and mission critical M2M communications, have generally been considered separately.

In some multimedia IoT applications/scenarios, MTDs may be deployed in areas which experience intermittent or poor coverage from cellular BSs, or MTDs may be deployed in environments with no cellular BS infrastructure (e.g. desert). Moreover, in some cases, MTDs might be unable to transmit their data due to the large distances to the cellular BS, limited battery lifetime of MTDs, and MTD power constraints. Therefore, the aforementioned scenarios for existing cellular infrastructure, will lead to increased latency and sensed data to be potentially lost. Therefore, the use of UAVs as flying BS will help over come these problems. The work in [52] studied the optimal altitude for low flying UAVs, in order to provide maximum coverage area per UAV across all ground users. In [154], an efficient deployment of UAVs as flying BSs under different network densities was proposed, to provide maximum user coverage, while also minimising the number of deployed

UAVs. Moreover, the work in [155], investigated the optimal deployment of UAVs as aerial BSs, while also jointly optimising radio resource allocation and user association, in order to minimise power consumption of IoT devices for uplink UAV coverage. On the other hand, optimal path planning (trajectory) approaches for UAVs are proposed in [156, 157], and [158], in order to maximise throughput of ground users and minimise energy consumption of UAVs. The work in [156] proposes a joint optimisation for user scheduling, power control, and UAV trajectories, in order to maximise throughput and fairness amongst all users. In [157], a relay-based UAV system was considered, where the UAV relay trajectory and source/relay transmit power allocation are jointly optimised to maximise the throughput.

A number of works such as in [49, 159], and [51] have also considered UAVs as flying BSs to assist with data collection. The goal of these works is to deploy UAVs to assist sensors on the ground that are unable to send their sensed data to a cellular BS, where the UAV collects data from each sensor. The work in [49] considers a clustered sensor network, where UAVs are employed to collect data from each cluster head and recharge the cluster head. While, in [159], a priority-based data access scheme was proposed, which utilises UAVs as relays to collect the sensed data from the wireless sensor network, in order to reduce energy consumption, increase throughput, and suppress redundant sensors. However, the ground users in the aforementioned works [49, 51, 52, 154–159] do not consider ultra-reliability and low latency constraints, especially when there are not enough UAVs to cover a massive number of MTDs deployed in a vast geographical area.

Meanwhile, mission critical M2M communications aims to guarantee ultra reliable and low latency communications. Within literature, a range of different approaches have been considered, in order to guarantee URLLC in mission critical applications, such as, short packet transmissions, shorter transmission time interval (TTI), caching, network slicing, or queue latency [42, 53–59]. The work in [53] investigates short packet communications and the resources required for transmission, which leads to minimising latency and ensuring ultra reliability. In [54], a nonadaptive and adaptive retransmission schemes are proposed for short packet communications. On the other hand, the work in [55] proposed a uplink contention based access scheme to handle the arrival of URLLC packets, using diversity transmission.

Another approach to ensuring URLLC for MTDs, is to focus on queuing latency,

and in particular ensuring that the queue length does not exponentially increase to infinity which will cause the network to become unstable, this has been investigated in [57] and [58]. However, to determine the queue length in practice is challenging, as the arrival of packets are dynamic and non-deterministic, which means that the probability model of the event-driven traffic of MTDs is unknown. Moreover, in the aforementioned works [42, 53–59], the presence of UAVs were also not considered. In fact, none of the prior studies in [42, 49, 51–59, 154–159] considered the problem of jointly optimising UAV scheduling and guaranteeing URLLC for M2M communications on the ground.

Despite the abundance of prior works on UAV BSs, there exists only a limited number of works that address the problem of URLLC with UAV BSs. Primarily, the works in [160] and [161] have studied the possibility of utilising UAVs to support mission critical communications. In [160], guaranteeing quality-of-service (QoS) for URLLC links from ground users to UAVs is studied, in order to find optimal UAV altitude and minimise the total required bandwidth for URLLC. On the other hand, the work on [161] proposed utilising UAVs as floating relays in a heterogenous cellular network, to serve low-priority MTD traffic and URLLC traffic by investigating frequency reuse, interference, backhaul resource allocation, and coverage. However, these works rely on a given transmission probability of packets, as well as using QoS to measure the reliability an latency which is not necessarily guaranteed for URLLC all the time. Therefore, a fully distributed and autonomous energy-aware scheduling approach for mission critical M2M communications utilising UAVs as flying BSs is needed, to ensure ultra reliability, low latency, and network stability.

Therefore, this thesis aims to investigate a fully distributed and autonomous energy-aware scheduling approach for mission critical M2M communications utilising UAVs as flying BSs, while ensuring ultra reliability, low latency, and network stability.

2.4 Summary

In this chapter, we have provided a comprehensive literature review of the key research challenges for multimedia IoT, with particular focus on: D2D communications, M2M communications, and UAVs as low flying BSs. However, there

are still open issues that need to be addressed in this area. Hence, we outline a high-level summary on the limitations of existing work in these areas, as follows:

- QoE is an emerging measure of quality for future cellular communications. QoE guarantees end-user satisfaction, which can be expressed as a benefit that a particular user feels.

- Existing multiple objective Stackelberg games, typically use scalarisation to solve such games, which combines all separate objective functions into one objective (utility) function. Thus, the objective functions cannot be chosen arbitrarily, and the scalarised solution will cause real-time information exchange between users (as well as affecting the non-cooperative nature of the game).

- There is a need to adopt a distributed mode selection, resource allocation, and interference management schemes for D2D communications, in order to coordinate and reduce intra-cell interference, enable D2D pairs to operate more efficiently, and improve network throughput and performance. D2D pair scheduling has attracted considerable interest in order to reduce intra-cell interference in the network. However, most existing work on D2D pair scheduling do not jointly consider other factors such as resource allocation or mode selection. Moreover, most of the existing works implement network-assisted D2D communications, where the BS assists D2D users in determining which transmission mode to operate in, i.e., the BS has global knowledge of the network (complete CSI and knowledge all user locations). In order for the BS to acquire global knowledge of the network, this will cause signalling overhead and network complexity to significantly increase, especially with the increasing number of users in the network. In practical scenarios, network-assisted D2D communications is not always ideal, as complete CSI cannot always be achieved for all user all the time.

- Clustering MTDs based on data correlation, is used to improve the data quality sent to the BS. Existing works consider only cluster a small, finite number of MTDs. However, this is not the case in practical IoT scenarios as the number of MTDs within the network is massive, which causes substantial interference and impacts the way in which correlation-aware cluster must

be performed. Meanwhile, the existing works are also not robust to small stochastic changes in the M2M environment.

- Mission critical M2M communications aims to guarantee ultra reliable and low latency communications. In some instances, MTDs may be deployed in remote geographical regions with poor cellular coverage or no coverage at all. UAVs as flying BSs have been considered, in order to assist MTDs in these situations by forwarding their data to the nearest BS. However, there exists only a limited number of works that address guaranteeing URLLC for ground users, with UAVs BSs.

Chapter 3

Flexible D2D Application-Driven Resource Allocation

In underlaid D2D communications — where D2D pairs are operating in reuse mode — a distributed interference management and resource allocation scheme can enable interference to be coordinated and mitigated between cellular and D2D users within the network, and therefore improve D2D user QoE, reliability, and energy-efficiency. This chapter addresses the fundamental problem of flexibility for application-driven distributed resource allocation, using a dynamic multiple-objective (equivalently multi-criteria) Stackelberg game. In this game, we consider a single leader (BS) and multiple followers (D2D pairs), where the leader reduces interference within the network by charging a price to followers. The followers react to the charged price, and compete to find optimal transmit power and resource block allocation, while guaranteeing QoS and QoE. To enhance D2D user QoE, D2D users are categorised into one of three application classes, where each class formulates a criterion set. In contrast to previous studies on interference management using multiple-objective Stackelberg game theory in D2D communications, we consider arbitrarily different utility functions for each application class that best describes the D2D pair's practical application or service and meets the application class criterion.

We propose a crucial innovation where the several different utility functions, available to followers, are solved using a non-scalarised approach, which will reduce real-time information exchange between D2D pairs, while also maintaining the non-cooperative and selfish behaviour between them. To solve this novel Stackelberg game, a distributed algorithm is proposed, and it is shown to converge to a unique Stackelberg equilibrium across all users, which is a sub-game perfect equilibrium

for each game stage. Finally, our simulation results show that transmit power is reduced and data rate is increased for all D2D pairs in each application class, while the BS ensures best social welfare is guaranteed across all D2D users. Moreover, we demonstrate beneficial effects of prioritising communications, along with the implicit prioritisation from different utility functions, by setting different PDR targets.

This chapter is organised as follows. The system model and problem formulation for the application-driven resource allocation for underlaid D2D communications are presented in Section 3.1. In Section 3.2, a novel Stackelberg game is proposed, which considers multiple, different, utility functions available to followers. Within Section 3.3, the relationship between the leader and followers is analysed. Section 3.4 outlines the proposed algorithm and Section 3.5 analyses the convergence of the proposed distributed algorithm. Simulation results are illustrated in Section 3.6. Finally, Section 3.7 concludes the chapter.

3.1 System Model and Problem Formulation

Consider a single BS serving C cellular users and D D2D pairs, where D2D communication underlays cellular communications. In the network, each cellular user j from the set $\mathcal{C}$ (i.e. the user accessing the BS) and each D2D pair i from the set $\mathcal{D}$, are independently and uniformly distributed throughout the network, as illustrated in Fig. 3.1. The number of D2D pairs is considered to be much larger than the number of cellular users in the network, $D > C$ [162, 163]. We assume that all D2D pairs considered in the network reuse the cellular user's uplink resource block, which means that for each D2D pair the distance $\xi_{x,y}$ between transmitter x and receiver y is less than or equal to a threshold distance $\overline{\xi}$, $\xi_{x,y} \leq \overline{\xi}$. Moreover, the BS and all transmitters/receivers are equipped with omni-directional antennas, as device orientation is not pertinent to our model and use-cases.

In the network, the BS assigns K orthogonal resource blocks from the set $\mathcal{K}$, to cellular users. Each cellular user j is allocated one resource block for uplink transmission, based on an orthogonal frequency-division multiple access (OFDMA) system. We assume that the number of utilised resource blocks in the network is equal to the number of cellular users in the network, $K = C$. On the other hand, the BS allocates non-orthogonal spectrum to D2D pairs, that is, allocating D2D pairs

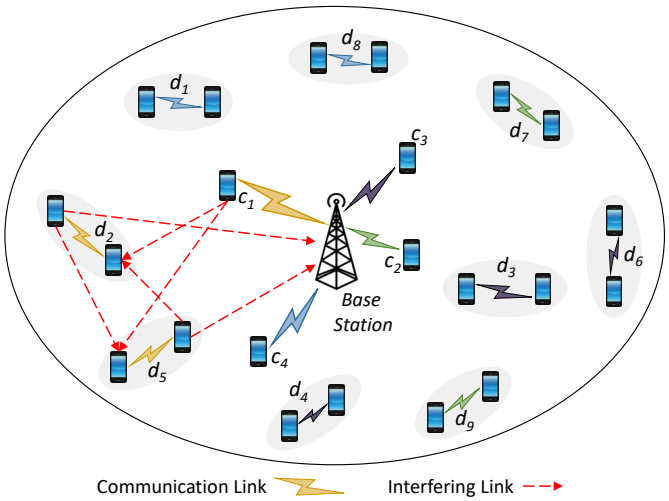


Figure 3.1: Illustration of the D2D communication system topology for reuse mode only. The intra-cell interference for cellular user c_1 is outlined in red. Note that, if the colour of the communication link for either a cellular user and/or a D2D pair is the same, this indicates that the BS has allocated the same resource block for transmission.

to a cellular user resource block. Thus, paired D2D transmitters reusing a cellular user's resource block will cause the BS to suffer interference, while paired D2D receivers reusing the cellular user's resource block will suffer interference from the cellular user and all other non-paired D2D transmitters sharing the same resource block (intra-cell interference). (And we assume that inter-cell interference is dealt with separately).

Fig. 3.1 is an illustrative example of the D2D communication system topology, where four cellular users are transmitting to the BS (uplink), and nine D2D pairs are operating in reuse mode. The colours of the transmission links indicates which D2D pairs are reusing which cellular user's uplink resource block. The intra-cell interference generated due to resource block sharing is outlined for cellular user c_1 in Fig. 3.1. The example scenario considered in Fig. 3.1 is quite simple and can be extended to a more complex scenario with many more users and BSs.

The received signal-to-interference-plus-noise ratio (SINR) of D2D pair i reusing the uplink resources of cellular user j, is:

$$\gamma_{di} = \frac{p_{di}|g_{di,di}|^2}{p_{cj}|g_{cj,di}|^2 + \sum_{\substack{m\in\mathcal{M}\\ m\neq i}} p_{dm}|g_{dm,di}|^2 + N_0}. \tag{3.1}$$

The received uplink SINR of cellular user j who is sharing its uplink resource with

a set $\mathcal{M} \subset \mathcal{D}$ of M D2D pairs, is:

$$\gamma_{cj} = \frac{p_{cj}|g_{cj,BS}|^2}{\sum_{m\in\mathcal{M}} p_{dm}|g_{dm,BS}|^2 + N_0}, \tag{3.2}$$

where p_{di} is the transmit power of D2D pair i; p_{cj} is the transmit power of cellular user j; a set $\mathcal{M} \subset \mathcal{D}$ of M D2D pairs reusing the same cellular user's resource block; p_{dm} is the interfering transmit power of D2D pair $d_m \in \mathcal{M}$; and N_0 is the additive white Gaussian noise (AWGN) power. The channel gain $g_{a,b}$ between transmitter a and receiver b, considers a slowly time-selective flat Rayleigh fading channel model (small-scale fading), as follows:

$$g_{x,y} = A_{PL}A_{SSF}\left(\frac{\xi_0}{\xi_{x,y}}\right)^{\frac{\alpha}{2}}, \tag{3.3}$$

where A_{PL} is the free space path loss channel attenuation; A_{SSF} is the small-scale fading channel attenuation with fixed normalised Doppler spread; ξ_0 is the reference distance between the transmitter and receiver; and α is the path loss exponent. Furthermore, to ensure interference within the network will not degrade any user's performance (QoS), we place a threshold on received SINR for both cellular users $\gamma_{cj} \geq \overline{\gamma}_c \; \forall \; j \in \mathcal{C}$, and D2D pairs $\gamma_{di} \geq \overline{\gamma}_d \; \forall \; i \in \mathcal{D}$.

Performance Metric: We consider two metrics in this chapter, for evaluating the performance of cellular users and D2D pairs: (i) *achievable spectral efficiency* (1.1), and (ii) *PDR* (1.3). To ensure acceptable PDR is achieved for all D2D pairs we seek to meet or exceed a target PDR, such that $pdr_{di} \geq pdr_{\text{th}}$, where pdr_{th} is target PDR and is defined in terms of threshold SINR $\overline{\gamma}_d$ as, $pdr_{\text{th}} = \exp(a(\overline{\gamma}_d)^b)$. Ideally, we want users to have very good perceived benefit throughout the cell, which includes achieving PDR above target PDR, that is, guaranteeing QoS and maintaining desired communication reliability. In fact, PDR can also be used to also prioritise communications, by assigning particular users a higher target PDR than others, which means that higher target PDR entails higher communication priority and greater reliability.

Typical modulation schemes employed in current cellular networks (LTE-A) and future 5G networks and beyond, include, QPSK, 16-QAM, and 64-QAM, which will be coded to assist with error correction. The error correction coding techniques considered are, low-density parity-check (LDPC) codes or turbo codes for

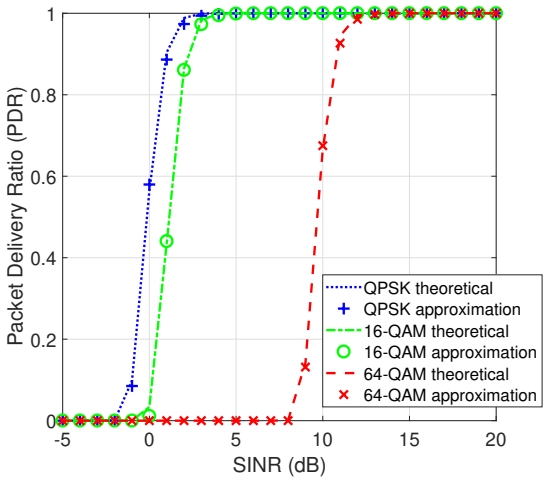


Figure 3.2: SINR versus PDR simulation and compressed exponential approximation (1.2), for the three modulation types and a packet size of 1500 bytes, as outlined in Table 3.1.

Table 3.1: Estimated parameters, a_c and b_c, from (1.2)

Modulation	**Coding Gain**	a_c	b_c	a	b
QPSK	8 dB [165]	1.097	6.552	-0.545	-6.552
16-QAM	11 dB [165]	0.816	7.378	-4.483	-7.378
64-QAM	7 dB [166]	0.114	7.127	-5.27×10^{6}	-7.127

example [164], which introduces a coding gain over uncoded modulation schemes. Fig. 3.2 compares the theoretical PDR results to the PDR approximation, and Table 3.1 outlines the coding gain and approximated values for a_c and b_c from (1.2), for all three modulations, with the same packet length of 1500 bytes.

Given this system model, our goal is to address resource allocation and interference management in D2D communications, while considering application-based QoE for all D2D pairs. Since the BS allocates uplink cellular resource blocks for D2D pairs to reuse, severe intra-cell interference is caused, which is illustrated in the system model. If this scenario is modelled as a non-cooperative power control game, the BS and D2D pairs will behave selfishly and may choose a strategy that will maximise their individual performance, without considering the effects on other users (cellular users, BS, and D2D pairs) in the network [108, 112]. Thus, the BS may not have any incentive to share the cellular users' uplink resource blocks with D2D pairs, as the BS (and the cellular users' uplink) will suffer severe intra-cell interference, as well as for other D2D pairs sharing the same resource blocks.

Furthermore, to enhance D2D pair's communication experience, D2D users are

Table 3.2: Practical application examples of D2D pairs for each class

Casual class	Short messaging service (SMS), local voice call, and advertising
Interactive class	Video calling (video conference), content sharing (such as, video and photo), and data sharing
Streaming class	Gaming, uploading/downloaded/streaming content, and emergency services[2]

categorised into three application-based classes[1]: *casual class*, *interactive class*, and *streaming class*. Table 3.2 outlines a small subset of practical examples of applications serviced by each class. Based on the practical applications, the casual application class will seek to lower transmit power requirements while also guaranteeing QoS, whereas, the interactive application class will seek to lower transmit power requirements while also increasing reliability requirements. While, the streaming application class will seek to maximise the achievable spectral efficiency requirements, such that content can be fully uploaded/downloaded/streamed in a reasonable timeframe, as well as to increase priority and reliability, and to reduce latency. To measure these requirements for the three application classes, SINR (3.1), achievable spectral efficiency (1.1), and PDR (1.3), will be used.

Next, we propose a fully distributed and autonomous flexible resource allocation framework based on *Stackelberg game theory* [65, 104, 108, 112, 115] for D2D communications. The proposed Stackelberg game considers a multi-criteria optimisation problem, with several different utility functions available to the followers (D2D pairs). We define three utility functions for the followers, that is, one for each application class, such that the utility functions meet a criterion set for each applicaiton/service to enhance D2D user QoE. Whereas, the objective of each utility function considered for the followers in the proposed Stackelberg game, is to optimise transmit power of the D2D pairs. Moreover, the several utility functions considered are arbitrarily different, and thus, a non-scalarised approach is proposed, which avoids real-time information exchange between D2D pairs.

[1]D2D pairs reusing the same cellular user's resource block, will cause intra-cell interference to other D2D pairs, even if they have been categorised in different application classes.

[2]Local security and safety communications between neighbouring police, firemen, and ambulance officers.

3.2 Flexible Resource Allocation Stackelberg Game

Here, we propose a dynamic flexible resource allocation Stackelberg game with several utility functions available to followers. The proposed Stackelberg game aims to coordinate intra-cell interference and allocate resources within the network efficiently, while enhancing D2D user QoE. We consider a single leader (BS) and multiple followers (D2D pairs) in our proposed game. The BS owns the resource blocks and charges the D2D pairs a reuse fee for reusing a cellular user's uplink resource block. The BS also charges D2D pairs a satisfaction/certainty fee, with respect to all users in the network, where BS satisfaction is defined as follows:

Definition 3.1 *BS satisfaction* is measured according to the BS's knowledge and certainty of the network and its perceived benefit of all D2D pairs and cellular users.

Our proposed Stackelberg game follows a hierarchical game structure, where the leader first sets a charging price to the followers, which is a function of BS satisfaction and reuse fees. The followers react to the leader's charging price by choosing optimal transmit power. The proposed Stackelberg game is defined as follows:

Definition 3.2 The *dynamic Stackelberg game with multiple utility functions available to followers* is defined by $G_S = (\mathcal{D}, \boldsymbol{x}, \boldsymbol{p}, u_L, u_F)$, where $\mathcal{D}$ is the finite set of rational D2D pairs; $\boldsymbol{x}$ is the leader's finite pure strategy set; $\boldsymbol{p}$ is the follower's finite pure strategy set; u_L is the leader's (BS) utility function; and u_F is the follower's (D2D pair) utility function.

In our proposed Stackelberg game, the leader's finite strategy set $\boldsymbol{x} = (0, 1]$ represents the BS's satisfaction fee across all users in the network. The finite strategy set for each D2D pair $i \in \mathcal{D}$ (follower) $\boldsymbol{p} = [p_{\min}, p_{\max}]$ denotes transmit power. Thus, the strategy space for all D2D pairs in each iteration of the game is $\mathcal{P} = \{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_D\}$. The leader's utility function, u_L, best describes the performance of the BS in order to improve its satisfaction with respect to all users in the network. The three separate utility functions available to each follower, $u_F \in \{u_{\mathrm{Ca}}, u_{\mathrm{Ia}}, u_{\mathrm{Sa}}\}$, map directly to a particular application class, which best describes the performance of each D2D pair's application or service, in order to improve QoE.

3.2.1 Leader Utility Function

We consider a single leader in the proposed Stackelberg game, where the utility of the leader $u_L(x, \alpha_{j,i}, p_{di})$ is a function of the BS's satisfaction fee for the network across all users, x, and the gain the BS earns from sharing cellular user resource blocks to D2D pairs, $\alpha_{j,i}$. The BS satisfaction fee x in the proposed leader's utility is set proportional to the throughput of all users in the network, D2D pair transmission power, and the total interfering cost of each D2D pair to the BS. Whereas, the reuse fee for D2D pair i reusing cellular user j's resource block $\alpha_{j,i}$ in the proposed leader's utility, is set proportional to the total interference the BS (and each cellular user) observes from each D2D pair on a particular resource block. Hence, the leader sets two fees, where the satisfaction fee is bounded by $x \in (0, 1]$ and the reuse fee is bounded by $\alpha_{j,i} > 0$. The BS has perfect information about the current state of the cell at each stage in the game, and measures the perceived benefit across all users in the network by charging the two fees. Note that, when the BS satisfaction fee is close to 0 this means the BS has little knowledge and certainty of the network, as well as little perceived benefit of the users in the network. Whereas a BS satisfaction fee of 1 means the BS has total satisfaction of the network (that is, very certain about the network) across all D2D users. The utility function $u_L(x, \alpha_{j,i}, p_{di})$ for the leader (BS) is given by:

$$u_L(x, \alpha_{j,i}, p_{di}) = x\beta \left(\overline{p} + \overline{I}_{BS}(\alpha_{j,i})\right) - \lambda \left(x\overline{r}\right)^2 - \left(\overline{I}_{BS}(\alpha_{j,i})\right)^2, \tag{3.4}$$

where $\overline{p} = \sum_{i \in \mathcal{D}} \frac{p_{di}}{D}$; $\overline{I}_{BS}(\alpha_{j,i}) = \sum_{j \in \mathcal{C}} \sum_{\substack{i \in \mathcal{M} \\ \mathcal{M} \subseteq \mathcal{D}}} \frac{\alpha_{j,i} p_{di} \widetilde{G}_{i,BS}}{D}$; $\overline{r} = \sum_{\substack{n \in \mathcal{N} \\ \mathcal{N} = \mathcal{C} \cup \mathcal{D}}} \frac{r_n}{N}$ and $r_n = \log_2(1 + \gamma_n)$; β is a positive scalar with respect to the game iteration, $\beta > 0$; a set $\mathcal{M} \subset \mathcal{D}$ of M D2D pairs reusing the same cellular user's resource block; $\mathcal{N}$ is the set of all cellular users and D2D pairs, $\mathcal{N} = \mathcal{C} \cup \mathcal{D}$; $\widetilde{G}_{i,BS} = \frac{|g_{i,BS}|^2}{|g_{i,i}|^2}$ is the normalised interfering channel gain from D2D pair i to the BS (and cellular user j); and λ is a tuning parameter, which determines how soon the BS's perception and satisfaction/certainty of the network converges to optimal satisfaction, $\lambda > 0$. The leader finds the best response satisfaction of the network with respect to all users, and the best response reuse fee, by maximising its utility function (3.4) with respect to the satisfaction and reuse fees, as follows:

$$\begin{aligned} &\underset{x,\alpha_{j,i}}{\text{maximise}} \quad u_L(x, \alpha_{j,i}, p_{di}) \\ &\text{subject to} \quad x \in (0,1];\ \alpha_{j,i} > 0, \quad \forall j \in \mathcal{C}, \forall i \in \mathcal{D}. \end{aligned} \tag{3.5}$$

3.2.2 Follower Utility Function

In the proposed Stackelberg game, multiple followers (D2D pairs) are considered, and the utility function of D2D pair i, $u_{F,i}$, depends on its application or service, where $u_{F,i}(x, \alpha_{j,i}, p_{di}) \in \{u_{\text{Ca}}, u_{\text{Ia}}, u_{\text{Sa}}\}$, that is, $u_{F,i}(x, \alpha_{j,i}, p_{di})$ maps directly to either the casual class utility $u_{\text{Ca},i}$, interactive class utility $u_{\text{Ia},i}$, or the streaming class utility $u_{\text{Sa},i}$. These utility functions are arbitrarily different, and cannot be scalarised and reduced into one single utility function. Initially, D2D pairs sort themselves into one application class that best describes the performance of the desired application or service. We assume that once D2D pairs have assigned themselves to a class, they cannot change classes during the finite stages of the proposed game. The utility function of follower i models the difference between D2D pair i's performance for a particular application, and the leader's charging price. The leader's charging price is a function of BS satisfaction fee x and reuse fee $\alpha_{j,i}$. In order to reduce intra-cell interference in the network, D2D pairs aim to find optimal transmit power. Hence, the individual finite action set for each follower i is transmit power, p_{di}, which is bounded by a minimum and maximum transmit power. The D2D pairs in the proposed game have imperfect information, which means that all D2D pairs select their action simultaneously. Each D2D pair determines their utility with respect to all other D2D pairs sharing the same resource block.

The casual class utility is a power/SINR balancing function, as defined in [167], with an additional cost charged by the leader. The criterion set for the casual class (examples outlined in Table 3.2) requires low transmission power, and acceptable achievable spectral efficiency, reliability, and QoS. The applications or service associated with this class are common/everyday functionalities of D2D users. The casual class utility $u_{\text{Ca},i}$ for D2D pair i is defined as:

$$u_{\text{Ca},i}(x, \alpha_{j,i}, p_{di}) = \left(\frac{\overline{\gamma}_d}{\gamma_{di}}\right) p_{di} - p_{di} x \delta (0.5 + \alpha_{j,i} \widetilde{G}_{i,BS}), \tag{3.6}$$

where δ is a positive constant. Thus, utility function (3.6) was chosen for the casual

application class, as it provides large emphasis on minimising transmit power, while also guaranteeing QoS for D2D pairs.

The interactive class utility is a trade-off between maximising SINR and minimising transmit power, as defined in [167], with an additional cost charged by the leader. The criterion set for the interactive class (examples outlined in Table 3.2) requires similar transmission power to the casual application class, however, it requires more reliability than the casual class, which is due to the transferring of data files between users, for example. Thus, this utility function will ensure good signal quality at the receiver without large latency periods. The interactive class utility $u_{\mathrm{Ia},i}$ for D2D pair i is defined as:

$$u_{\mathrm{Ia},i}(x, \alpha_{j,i}, p_{di}) = -qp_{di} - k\left(\overline{\gamma}_d - \gamma_{di}\right)^2 - p_{di}x\delta(0.5 + \alpha_{j,i}\widetilde{G}_{i,BS}), \tag{3.7}$$

where q and k are positive constants to ensure all parts of (3.7) have the same magnitude; and $\overline{\gamma} - \gamma_{di}$ is the SINR error, which is the difference between target SINR and received SINR of D2D pair i. If SINR error is less than 0, then the received SINR of D2D pair i is greater than target SINR, which means that QoS is guaranteed for D2D pair i. The utility function (3.7) was chosen for the interactive application class, as there is large emphasis on good signal quality in order to increase reliability, while also providing emphasis on reducing transmit power.

The streaming class utility is a trade-off between maximising achievable spectral efficiency and minimising transmit power, as defined in [78], with an additional cost charged by the leader. The criterion set for the streaming class (examples outlined in Table 3.2) allows much larger transmit power, and requires higher achievable spectral efficiency and reliability when compared to the other two classes, while also being able to adapt to changes in the network where QoS and reliability are guaranteed for each D2D pair. In particular, this utility function was chosen for the streaming application class, as it provides greater emphasis on achievable spectral efficiency, such that, content can be fully uploaded/downloaded/streamed in a reasonable timeframe, serious gamers won't experience any latency concerns, and emergency services will have higher priority and reliability. The streaming class utility $u_{\mathrm{Sa},i}$ for D2D pair i is defined as:

$$u_{\mathrm{Sa},i}(x, \alpha_{j,i}, p_{di}) = vr_{di} - wp_{di}^2 - p_{di}x\delta(0.5 + \alpha_{j,i}\widetilde{G}_{i,BS}), \tag{3.8}$$

where r_{di} is the channel of D2D pair i; and v and w are positive scale factors to ensure all parts of (3.8) have the same magnitude.

We observe that transmit power p_{di} of each D2D pair i belongs to a nonempty, convex, and compact subset of Euclidean space $\mathbb{R}^{|\mathcal{D}|}$. Each follower i finds the best response (optimal) transmit power, by maximising its utility function $u_F \in \{u_{\text{Ca}}, u_{\text{Ia}}, u_{\text{Sa}}\}$ (that is, the utility function that best maps to the follower's application or service) with respect to transmit power, as follows:

$$\begin{aligned} \underset{p_{di}}{\text{maximise}} \quad & u_{F,i}(x, \alpha_{j,i}, p_{di}) \; \forall \; i \; \in \; \mathcal{D} \\ \text{subject to} \quad & p_{\min} \leq p_{di} \leq p_{\max}. \end{aligned} \tag{3.9}$$

Conjecture 3.1 A non-scalarisation approach to multi-criteria optimisation – considering several, and arbitrarily different, utility functions for followers – can result in a best response outcome for all players when using a Stackelberg game.

3.2.3 Leader's Charging Price to the Followers

The aim of the leader's charging price, is to link the leader with each follower (and vice versa), in order to coordinate intra-cell interference and allow resource blocks to be utilised efficiently. Thus, the charging price the leader assigns to the followers must not be too high, otherwise the followers will choose high transmit power, which, in turn, will lead to severe intra-cell interference in the network. The proposed pricing strategy models the BS's satisfaction fee proportional to the follower's transmit power and interfering cost. The proposed pricing strategy fee θ_i charged by leader to each follower i, is:

$$\theta_i(x, \alpha_{j,i}, p_{di}) = p_{di} x \delta \left(0.5 + \alpha_{j,i} \widetilde{G}_{i,BS}\right), \tag{3.10}$$

where BS satisfaction fee x and reuse fee $\alpha_{j,i}$, are functions of follower transmit power p_{di}.

Initially, in the network, the BS will have little certainty and low satisfaction (and perceived benefit) of the network, which means that the leader's charging price assigned to the followers will not have a large impact on the follower utility functions. However, when the BS's perceived benefit from the network increases, this will lead to an increase in BS/network satisfaction, as the BS is becoming

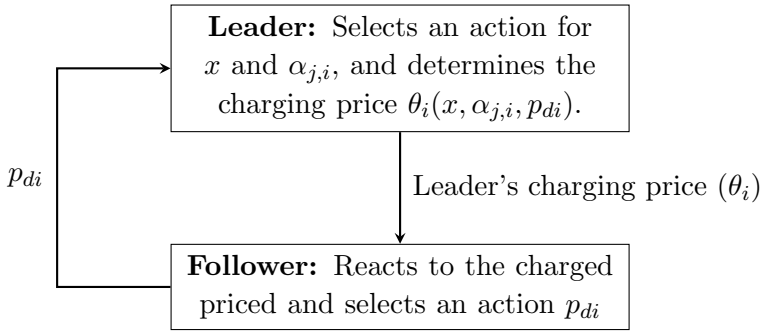


Figure 3.3: Flow chart of the proposed flexible resource allocation Stackelberg game.

more certain/satisfied about the users in the network, the network conditions, and the network layout. Thus, there is more coordination across users in the network. Additionally, when the BS satisfaction fee and reuse fee increase, the interfering cost assigned to the followers also increases. As the leader's charging price increases, this will cause the follower utility functions to further decrease and result in follower's reducing transmission power.

Fig. 3.3 outlines a flow chart of the proposed flexible resource allocation Stackelberg game. As we can see from Fig. 3.3, the leader first selects its actions and then sets a charging price to the followers, and the followers react to the leader's charging price by choosing optimal transmit power.

3.2.4 Resource Block Allocation

The allocation by the BS of cellular user uplink resource blocks for D2D pairs to reuse can lead to severe intra-cell interference. Power control is an effective tool to reduce such intra-cell interference, which also reduces transmit power. However, optimising transmit power by itself is not necessarily adequate, in order to reduce the impact of intra-cell interference. Thus, with the assistance of the BS, D2D pairs must find a suitable cellular user's uplink resource block to reuse, while ensuring cellular user performance does not degrade. Here, we propose a resource block allocation scheme along with the proposed flexible resource allocation Stackelberg game.

The resource block allocation scheme, takes the D2D pair optimal transmit power derived in the proposed Stackelberg game, that is, in (3.9), and determines its channel rate for each cellular user's uplink resource block in the network. Thus, in order for D2D pair i to determine which cellular user's uplink resource block to

reuse from the set $\mathcal{C}$, the following maximisation must be satisfied, as follows:

$$j^* = \max\left(\boldsymbol{\eta}_i\right), \tag{3.11}$$

where $j^* \in \mathcal{C}$ is the selected cellular user that D2D pair i obtains maximum channel rate, and $\boldsymbol{\eta}_i$ is a channel rate vector of D2D pair i for each cellular user $j \in \mathcal{C}$. Before D2D pair i switches to reuse cellular user j^*'s resource block, the QoS of cellular user j^* must be guaranteed, as follows $\gamma_{cj^*} \geq \overline{\gamma}_c$. If the QoS of cellular user j^* is not guaranteed, then D2D pair i will remain reusing the current cellular user's resource block, until the next time slot. Note that, when a D2D pair decides to switch the cellular user resource block they are reusing to get better channel rate, D2D pairs do not need to consider their effect on other D2D pairs. The D2D pairs only need to consider received interference if they start to reuse that resource block, as D2D pairs are selfish and only want to maximise their own payoff. Thus, allowing D2D pairs to choose which cellular user's resource block to reuse, will improve overall network performance, as well as further reducing D2D pair transmit power and intra-cell interference.

3.3 Analysis of the Proposed Stackelberg Game

Next, we analyse the leader's utility function and the three follower utility functions in the proposed Stackelberg game, using backward induction. Particularly, we investigate the trade-off between the leader's charging price and the followers transmit power, where we derive the best response BS satisfaction fee and best response reuse fee for the leader, and best response transmit power for each follower for the three application-based classes. Furthermore, we prove that there exists a unique Stackelberg equilibrium in the proposed game, across all users.

3.3.1 Follower Analysis

Given the satisfaction fee x and the reuse fee $\alpha_{j,i}$, each follower aims to maximise its utility by choosing optimal transmit power. Since we have considered three different utility functions for the set of followers, we will firstly analyse each type of utility function individually.

Proposition 3.1 Given x and $\alpha_{j,i}$, each utility function defined in (3.6), (3.7), and (3.8) for any D2D pair $i \in \mathcal{D}$, is strictly concave and continuous with respect to D2D pair i's transmit power.

Proof: See Appendix A.1. ■

From (3.9), the best response (optimal) transmit power p^*_{di} for follower i, can be solved by setting the first derivative of the three follower utility functions in (A.1), (A.2), and (A.3) to zero, given the BS's satisfaction fee x and reuse fee $\alpha_{j,i}$. However, instead of setting (A.1) to zero, the best response transmit power $p^*_{\mathrm{Ca},di}$ for the casual class is given by (3.6), that is, $p^*_{\mathrm{Ca},di}(x, \alpha_{j,i}) = u_{\mathrm{Ca},i}$, since the casual class utility is a linear and power balancing function, with an added cost. The best response transmit power $p^*_{\mathrm{Ca},di}$ for casual class utility is searched within $\{p_{\min}, p^*_{\mathrm{Ca},di}, p_{\max}\}$. Whereas, the best response transmit power $p^*_{\mathrm{Ia},di}$ for the interactive class, $\frac{\partial u_{\mathrm{Ia},i}(\cdot)}{\partial p_{di}} = 0$, is derived as follows:

$$p^*_{\mathrm{Ia},di}(x, \alpha_{j,i}) = \frac{I_{di} + N_0}{|g_{di,di}|^2} \left(\overline{\gamma}_d - \frac{(I_{di} + N_0)(q + \theta'_i)}{2k|g_{di,di}|^2} \right), \tag{3.12}$$

where the best response transmit power for the interactive class utility is searched within $\{p_{\min}, p^*_{\mathrm{Ia},di}, p_{\max}\}$. Furthermore, the best response transmit power $p^*_{\mathrm{Sa},di}$ for the streaming class, $\frac{\partial u_{\mathrm{Sa},i}(\cdot)}{\partial p_{di}} = 0$, is derived as follows:

$$\begin{aligned} p^*_{\mathrm{Sa},di}(x, \alpha_{j,i}) = &\left(\frac{(I_{di} + N_0)\left[w(I_{di} + N_0) - |g_{di,di}|^2 \theta'_i\right]}{4w(|g_{di,di}|^2)^2} + \frac{8vw + \ln(2)(\theta'_i)^2}{16w^2 \ln(2)} \right)^{\frac{1}{2}} \\ &- \frac{I + N_0}{2|g_{di,di}|^2} - \frac{\theta'_i}{4w}, \end{aligned} \tag{3.13}$$

where the best response transmit power for the streaming class is searched within $\{p_{\min}, p^*_{\mathrm{Sa},di}, p_{\max}\}$.

Thus, the best response transmit power $p^*_{\mathrm{Ca},di}$, $p^*_{\mathrm{Ia},di}$, and $p^*_{\mathrm{Sa},di}$ for the three application-based classes maximise their corresponding utility, and are monotonically decreasing with BS satisfaction fee x and reuse fee $\alpha_{j,i}$. As the leader becomes more certain and satisfied with the network, its satisfaction fee increases, which in turn, allows total interfering cost to be charged to the followers, and as a result, causing best response transmit power to decrease.

If the reuse fee $\alpha_{j,i}$ does not have an upper bound, the BS can infinitely increase this cost, causing the followers to decrease their transmission power to $p_{\min}$, and

defeating the purpose of the Stackelberg game [112, 115]. Instead, we set a lower and upper bound for the reuse fee $\alpha_{j,i}$, using the best response transmit power of the casual, interactive, and streaming application classes in (3.6), (3.12) and (3.13) respectively, to ensure the BS charges a reasonable fee. The reuse fee expressions for the three application classes, are given by:

$$\text{Casual class:} \quad \alpha_{j,\text{Ca}} = \frac{\overline{\gamma}_d \left(I_{di} + N_0\right)}{p_{di} x \delta \widetilde{G}_{i,BS} |g_{di,di}|^2} - \frac{1 + 0.5 x \delta}{x \delta \widetilde{G}_{i,BS}}; \tag{3.14}$$

$$\text{Interactive class:} \ \alpha_{j,\text{Ia}} = \frac{2k|g_{di,di}|^2}{x \delta \widetilde{G}_{i,BS} \left(I_{di} + N_0\right)} \left(\overline{\gamma}_d - p_{di} \left(\frac{|g_{di,di}|^2}{I_{di} + N_0}\right)\right) - \frac{q + 0.5 x \delta}{x \delta \widetilde{G}_{i,BS}}; \tag{3.15}$$

$$\text{Streaming class:} \ \alpha_{j,\text{Sa}} = \frac{v|g_{di,di}|^2}{\ln(2) x \delta \widetilde{G}_{i,BS} \left(p_{di}|g_{di,di}|^2 + I_{di} + N_0\right)} - \frac{2 w p_{di} + 0.5 x \delta}{x \delta \widetilde{G}_{i,BS}}; \tag{3.16}$$

where p_{di} is equal to $p_{\min}$ for the reuse fee upper bound for the three application classes, $\alpha_{j,\text{Ca}-\min}$, $\alpha_{j,\text{Ia}-\min}$, $\alpha_{j,\text{Sa}-\min}$; on the other hand, p_{di} is equal to $p_{\max}$ for the reuse fee lower bound for the three application classes, $\alpha_{j,\text{Ca}-\max}$, $\alpha_{j,\text{Ia}-\max}$, $\alpha_{j,\text{Sa}-\max}$. Thus, the reuse fee $\alpha_{j,i}$ lower and upper bounds are in terms of maximum and minimum transmit power respectively, and is given by, $\boldsymbol{\alpha}_{j,i} = [\alpha_{j,\min}, \alpha_{j,\max}]$, where $\alpha_{j,\min} = \max\{\alpha_{j,\text{Ca}-\min}, \alpha_{j,\text{Ia}-\min}, \alpha_{j,\text{Sa}-\min}\}$ and $\alpha_{j,\max} = \min\{\alpha_{j,\text{Ca}-\max}, \alpha_{j,\text{Ia}-\max}, \alpha_{j,\text{Sa}-\max}\}$, while $\alpha_{j,\min} > 0$ and $\alpha_{j,\max} > 0$, for cellular user $j \in \mathcal{C}$.

In the following, we will continue to analyse the leader's utility function and find the best response satisfaction fee and the best response reuse fee.

3.3.2 Leader Analysis

Recall that the Stackelberg game has a hierarchal structure, where the leader sets a fee and knows that the followers will react to this cost by searching for an optimal transmit power within their strategy set, $\boldsymbol{p}$. According to the best response transmit power expressions in (3.6), (3.12) and (3.13), as well as the leader's utility function in (3.4), it is evident that there exists a trade off between the leader's charging price and maximising the leader's payoff/utility.

Proposition 3.2 The utility function of the leader in (3.4) is strictly concave and continuous with respect to the BS satisfaction fee x and the reuse fee $\alpha_{j,i}$, given the best response transmit power of all the followers from the strategy set

$p_{\min} \leq p_{di}^{*} \leq p_{\max}$.

Proof: See Appendix A.2. ■

From (3.4), the best response BS satisfaction fee x^* and reuse fee $\alpha_{j,i}^{*}$ can be solved by setting the first derivative of the leader's utility function in (A.9) and (A.10) to zero, given the D2D pair's and cellular user's transmit power. The unique maximum solution (best response) for the BS satisfaction fee x^* is derived as follows:

$$x^{*}(\alpha_{j,i}, p_{di}) = \frac{\beta\left(\overline{p} + \overline{I}_{BS}(\alpha_{j,i}^{*})\right)}{2\lambda\overline{r}^{2}}, \tag{3.17}$$

where $\overline{I}_{BS}(\alpha_{j,i}^{*}) = \sum_{j\in\mathcal{C}} \sum_{\substack{i\in\mathcal{M} \\ \mathcal{M}\subseteq\mathcal{D}}} \frac{\alpha_{j,i}^{*} p_{di} \widetilde{G}_{i,BS}}{D}$.

The unique solution of the BS's satisfaction fee x^* is dependent upon the best response reuse fee $\alpha_{j,i}^{*}$ that the BS assigns to follower i for reusing the cellular user j's resource block, and follower transmit power p_{di}. Since the BS satisfaction fee is bounded by $(0, 1]$, the BS will search for the optimal satisfaction fee within $\{x_{\min}, x^{*}, x_{\max}\}$. Furthermore, the best response BS satisfaction fee of the network with respect to all users, guarantees best *social welfare* across all followers (D2D pairs). Social welfare for the proposed Stackelberg game is defined as follows:

Definition 3.3 *Social welfare* is the summation of all follower utility functions across all application classes.

Remark 3.1 The BS can determine the social welfare of the network by summing all the D2D users utility functions (D2D user satisfaction) [168, 169], where maximising social welfare is a socially optimal outcome for particular network users-of-interest [170].

The best response reuse fee $\alpha_{j,i}^{*}$ for cellular user j and D2D pair i, is derived as follows:

$$\alpha_{j,i}^{*}(x, p_{di}) = \frac{\beta x^{*}}{2 p_{di} \widetilde{G}_{i,BS}}. \tag{3.18}$$

The unique solution of the reuse fee $\alpha_{j,i}^{*}$ is dependent upon the BS satisfaction x^*, and the interfering power that each D2D pair i causes the BS (and cellular user j's uplink resource block). Since the reuse fee is bounded by $\alpha_{j,i} > 0$ and $[\alpha_{j,\min}, \alpha_{j,\max}]$, the BS will search for best response reuse fee within $\{\alpha_{j,\min}, \alpha_{j,i}^{*}, \alpha_{j,\max}\}$.

Therefore, by substituting optimal BS satisfaction fee x^* (3.17) and optimal reuse fee $\alpha^*_{j,i}$ (3.18) into the best response transmit power expressions for each application class, (3.6), (3.12) and (3.13), this will generate a maximal (optimal) transmit power response for each class.

3.4 Proposed Algorithm

To solve the dynamic flexible resource allocation Stackelberg game, with several utility functions available to followers, we propose a distributed algorithm. The distributed algorithm aims to find optimal BS satisfaction fee and reuse fee for the BS, while finding optimal transmission power for the followers. Algorithm 1 enables the BS to set a price to charge to the followers, for reusing a cellular user's resource block and to coordinate intra-cell interference between users, where the followers react to the charging price by selecting transmit power. During the proposed algorithm, it is assumed that the D2D pairs are established (and the number of D2D pairs in the cell is constant) over a finite time horizon of the game. If any additional D2D pairs are established in the network or any D2D pairs in the network are terminated, then the new set of D2D pairs will be considered in the next occasion when the game is played. Note that, the proposed Stackelberg game can be re-initiated any-time after convergence, with the new set of D2D pairs. Hence, Algorithm 1 outputs the best response BS satisfaction fee and reuse fee for the leader (BS), as well as the best response transmit power for all D2D pairs (followers).

Next, we prove that there exists a unique Stackelberg equilibrium for the proposed Stackelberg game with different, multiple, utility functions available to the followers.

3.5 Stackelberg Equilibrium Analysis

To solve the proposed Stackelberg game with multiple utility functions available to followers, we consider the concept of a *Stackelberg equilibrium*. The Stackelberg equilibrium is a stable point in the game, where the leader and followers have no incentive to increase their payoffs without changing their strategy unilaterally, given all other player's strategies. The Stackelberg equilibrium for our proposed

game is defined as:

Definition 3.4 The set of strategies $\boldsymbol{A}^* = [x^*, \boldsymbol{\alpha}^*, \boldsymbol{p}^*]$ for the BS (leader) and all D2D pairs (followers), $i \in \mathcal{D}$, there exists a Stackelberg equilibrium in the proposed Stackelberg game, if the following is satisfied:

1. The leader first achieves a Nash equilibrium:

$$u_L([x^*, \alpha^*_{j,i}, p^*_{di}], [x^*, \boldsymbol{\alpha_{j,-di}}^*, \boldsymbol{p_{-di}}^*]) \geq u_L([x, \alpha_{j,i}, p_{di}], [x^*, \boldsymbol{\alpha_{j,-di}}^*, \boldsymbol{p_{-di}}^*]); \tag{3.19}$$

2. The followers then achieve a Nash equilibrium thereafter:

$$u_{F,i}([x^*, \alpha^*_{j,i}, p^*_{di}], [x^*, \boldsymbol{\alpha_{j,-di}}^*, \boldsymbol{p_{-di}}^*]) \geq u_{F,i}([x, \alpha_{j,i}, p_{di}], [x^*, \boldsymbol{\alpha^*_{j,-di}}, \boldsymbol{p^*_{-di}}]); \tag{3.20}$$

for all $i \in \mathcal{D}$ and $j \in \mathcal{C}$; where $u_{F,i}(x, \alpha_{j,i}, p_{di}) \in \{u_{\mathrm{Ca}}, u_{\mathrm{Ia}}, u_{\mathrm{Sa}}\}$; p^*_{di} is an element of $\boldsymbol{p}^*$, which is the best response transmit power of D2D pair i; $\boldsymbol{p_{-di}}^*$ are elements of $\boldsymbol{p}^*$, which are the best response transmit powers for all other D2D pairs other than D2D pair i; and $\boldsymbol{\alpha_{j,-di}}^*$ is an element of $\boldsymbol{\alpha}^*$ and is the best response reuse fees for all other D2D pairs other than D2D pair i for cellular user j.

Remark 3.2 The Stackelberg equilibrium is a hierarchal equilibrium, that is, once the leader achieves a unique Nash equilibrium, the followers will then achieve a unique Nash equilibrium thereafter, resulting in a unique Stackelberg Equilibrium[3].

From our analysis of the leader and followers in the proposed Stackelberg game, we analysed a subgame of the proposed game. Within the subgame we solved the best response BS satisfaction fee and best response reuse fee for the leader, as well as the best response transmit power for the followers, that is, an equilibrium of the subgame. Subgame perfection of the Stackelberg game is defined as follows:

Definition 3.5 A Stackelberg equilibrium of the proposed game is said to be *subgame perfect* if and only if the strategy profile $\boldsymbol{A}^*$ results in a Nash equilibrium for

[3]In a typical single-leader multiple-follower Stackelberg game, the leader converges to a best response outcome, which causes the followers to converge to a Nash Equilibrium, such that the Nash (equivalently Stackelberg) Equilibrium is not converged to separately from the leaders best response.

Algorithm 1 Flexible Resource Allocation Stackelberg Game in D2D Communications

Input: Leader's (BS) finite strategy set: $\boldsymbol{x} = (0, 1]$; Follower's (D2D pair) finite strategy set: $\boldsymbol{p} = [p_{\min}, p_{\max}]$;
The BS randomly assigns each D2D pair $i \in \mathcal{D}$ to reuse cellular user $j \in \mathcal{C}$ resource block;
Each D2D pair i determines preferred application class: {casual (Ca), interactive (Ia), streaming (Sa)};
The BS initially selects a satisfaction fee $x \in \boldsymbol{x}$ and a reuse fee $\alpha_{j,i} \in [\alpha_{j,\min}, \alpha_{j,\max}]$ $\forall j \in \mathcal{C},\ \forall i \in \mathcal{D}$;
Each D2D pair i initially select transmit power $p_{di} \in \boldsymbol{p}$;
while $t \leq T$ **do**
    BS calculates the utility u_L, with strategy x and $\alpha_{j,i}$ $\forall j \in \mathcal{C},\ \forall i \in \mathcal{D}$, as in (3.4), given p_{di} $\forall i \in \mathcal{D}$;
    Update the BS's satisfaction fee $x(t)$, as in (3.17), and reuse fee $\alpha_{j,i}(t)$ $\forall j \in \mathcal{C},\ \forall i \in \mathcal{D}$, as in (3.18);
    for D2D pair $i \in 1 \dots D$ **do**
        Find suitable cellular user's uplink resource block to reuse, using (3.11);
        if D2D pair $i \in$ Ca **then**
            Calculate the utility $u_{\text{Ca},i}$ using (3.6), given the leader's charging price, θ_i as in (3.10);
            Update D2D pair i's transmit power $p_{di}(t+1) = u_{\text{Ca},i}$;
        end if
        if D2D pair $i \in$ Ia **then**
            Calculate the utility function, $u_{\text{Ia},i}$, using (3.7), given θ_i as in (3.10);
            Update D2D pairs i's transmit power $p_{di}(t+1)$ using (3.12);
        end if
        if D2D pair $i \in$ Sa **then**
            Calculate the utility function, $u_{\text{Sa},i}$, using (3.8), given θ_i as in (3.10);
            Update D2D pairs i's transmit power $p_{di}(t+1)$ using (3.13);
        end if
    end for
    $t = t + 1$;
end while
$p^*_{di} = p_{di}(T)$ $\forall i \in \mathcal{D}$; $\alpha^*_{j,i} = \alpha_{j,i}(T)$ $\forall j \in \mathcal{C},\ \forall i \in \mathcal{D}$; $x^* = x(T)$;
Output: Stackelberg equilibrium strategies $\boldsymbol{A^*} = [x^*, \boldsymbol{\alpha}^*, \boldsymbol{p}^*]$ for the BS (leader) and all D2D pairs (followers).

the leader and followers in every subgame of the proposed game, that is independent of the game history.

Based on [171] and our definition of a Stackelberg equilibrium in Definition 3.4, the proposed Stackelberg game using a non-scalarisation approach to solving the several different utility functions available to the followers, will result in a unique Stackelberg equilibrium. A unique Stackelberg equilibrium is the best response for the leader and all the followers that exists over all stages of the proposed game

(subgame perfect is a subgame that is independent of the game history).

Theorem 3.1 The Stackelberg game with different, multiple utility functions for followers, has a unique Stackelberg equilibrium.

Proof: See Appendix A.3. ■

Remark 3.3 When the game is operating at a Stackelberg Equilibrium, different applications across the three classes enable improvement of all D2D users QoE, throughput, and application requirements are met effectively.

Definition 3.6 The proposed dynamic Stackelberg game is Pareto-efficient when the leader's satisfaction is maximised, and all followers minimise transmit power and achieve desired PDR.

Proposition 3.3 When the leader's BS satisfaction fee converges to the unique maximum solution, this is an optimal outcome for the leader, which leads to the best response of the follower's transmit power and guaranteed best social welfare across all the followers, so the resultant outcome is Pareto-efficient.

Proof: A Pareto-efficient outcome is fully defined over all game stages, according to Algorithm 1. The leader first converges to an optimal BS satisfaction fee and reuse fee, then the followers converge to a Pareto-efficient outcome, such that the game achieves Pareto efficiency. When the BS satisfaction fee x is maximum the desired outcome in terms of network satisfaction is achieved. Hence, each D2D pair i will minimise its transmit power with respect to all other D2D pairs $-\boldsymbol{i}$ reusing the same resource block, until no D2D pair can further decrease its transmit power over the finite game, while desired PDR is achieved for each D2D pair. ■

3.6 Simulation Results

For our simulations, we consider a single BS located at the center of a circular area with a 500 m radius. D2D pairs and cellular users are distributed randomly throughout the network. We consider a dense D2D network, where all D2D pairs considered are within close proximity, that is, within the maximum D2D transmission range. All users encounter a slowly time-selective flat Rayleigh fading channel, with a normalised Doppler spread of 0.001, which for example, could correspond

Table 3.3: Simulation Parameters and Values

Parameter	**Value**
Cell radius, R	500 m
Maximum distance between D2D pairs	50 m
D2D pair reference distance, ξ_0	25 m
Number of D2D pairs per cell, D	48
Number of cellular users per cell, C	4
D2D free space path loss attenuation, A_{PL}	$10^{-3.32}$
D2D pair path loss exponent, φ	4
D2D user transmit power range	0-23 dBm
Cellular user SINR threshold, $\overline{\gamma}_c$	6 dB
AWGN power, N_0	-84 dBm

to a maximal Doppler frequency of 2 Hz with a game iteration occurring at 0.5 ms intervals. In particular, the interactive application class utility function (3.7) has the following constants $z = 0.005$ and $k = 1$, the streaming application class utility function (3.8) has the following constant $w = 2$, and the leader's charging price (3.10) has the following constant, $\delta = 10$. We assume each cellular user transmits at maximum transmission power (worst case scenario), 23 dBm. We consider a carrier frequency of 2 GHz. To ensure enough packets are successfully received at the D2D pair receiver, we introduce a minimum target PDR of 0.9 for all application classes. Turbo coding is applied to the modulated schemes used in simulations. The rest of the simulation parameters are outlined in Table 3.3.

Initially, an equal number of D2D pairs are randomly assigned to each cellular user, to reuse their resource block. In addition, an equal number of D2D pairs are assigned to each application class initially. At the beginning of the game, we assume that the BS has little satisfaction (certainty) and perceived benefit from the cell and the users within the network, and thus, at $t = 0$ the BS's satisfaction fee is set to $x = 0.001$.

The proposed Stackelberg algorithm is evaluated using Monte Carlo simulations with varying cellular users C and D2D pairs D. We compare our proposed Stackelberg algorithm to three different baseline algorithms, that is, a non-cooperative power control (NPC) algorithm using the utility function $u_{\text{NPC}}(p_{di}) = vr_{di} - wp_{di}^{q}$ from [[78], eq. (7)]; a distributed power control Stackelberg algorithm [109]; a joint spectrum and power allocation Stackelberg algorithm [104]; and a distributed resource allocation Stackelberg algorithm [119].

In Fig. 3.4, we show the BS's satisfaction fee for the network with respect to the

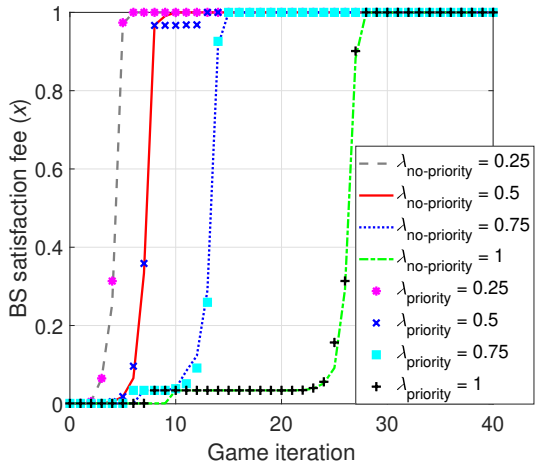


Figure 3.4: BS's satisfaction fee with and without priority, using coded 16-QAM.

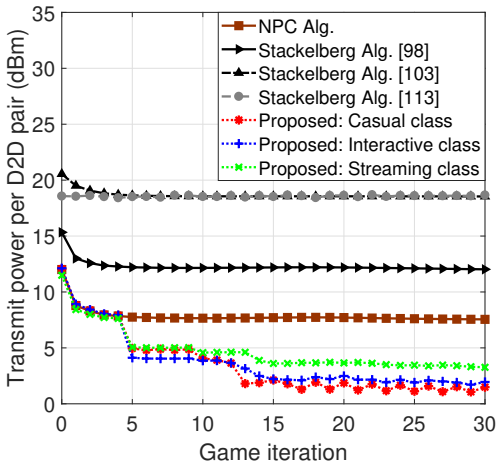


Figure 3.5: Transmit power per D2D pair, (no priority), using coded QPSK.

cellular and D2D users, for prioritised and non-prioritised D2D communications. We observe that the proposed Stackelberg game has fast convergence to an optimal BS satisfaction fee of $x = 1$. This means that the BS achieves full knowledge, certainty, and perceived benefit from the network across all users, especially D2D users. The BS's satisfaction fee evolves due to the leader's charging price charged to the followers, and the followers reacting to this price. Fig. 3.4 also shows the effect of the tuning parameter λ in the leader's utility function (3.4). As λ increases from 0.25 to 1, the game iteration at which the BS's satisfaction fee converges to an optimal outcome increases from 5 to 28, respectively. At $\lambda = 0.75$, this is a trade off point between transmit power and PDR. As λ increases from 0.25 to 0.75, transmit power decreases and PDR increases. However, once λ increases from 0.75 to 1, transmit power increases and PDR decreases. Thus, we consider $\lambda = 0.75$ for the rest of our simulations. Prioritising each application class, ensures particular communications will take priority over other communications. To prioritise communications, we set different minimum target PDR for each application class. Fig. 3.4 shows that there is no considerable effect on the leader's satisfaction fee, when prioritised or non-prioritised D2D communications is considered.

Fig. 3.5 shows the dynamics of average transmission power per D2D pair, without considering priority. We observe that once the BS satisfaction fee x and reuse fee $\alpha_{j,i}$ converge to optimal outcomes, the leader's charging price, defined in (3.10), also converges to an optimal outcome. Thus, D2D pairs (followers) further reduce their transmit power, as they are willing to allocate more from their utility in order

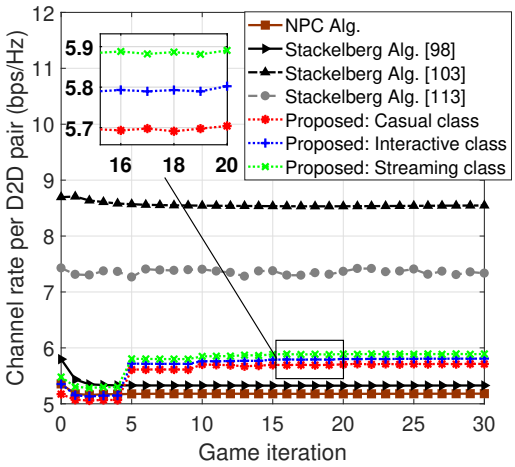


Figure 3.6: Channel rate per D2D pair, (no priority), using coded QPSK.

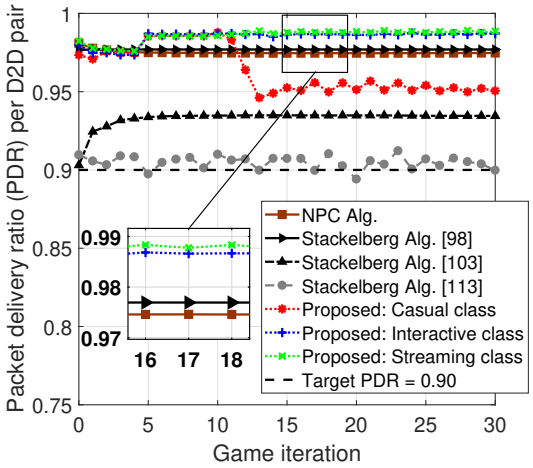


Figure 3.7: PDR per D2D pair, (no priority), using coded QPSK.

to achieve better QoE, as well as to reduce the interference to the BS and respective cellular user uplink resource block. From this figure, we can see that as the the D2D pairs switch to different cellular users' uplink resource blocks, D2D pair transmit power is reduced further on average, thus reducing intra-cell interference which leads to a power efficient network. Due to the effects of slow fading, transmit power does not converge to a single value in most cases. For each application class, average transmit power per D2D pair is decreased by around 91.5% for the casual class, 90.2% for the interactive class, and 89.3% for the streaming class, compared to the Stackelberg algorithm in [104]. Additionally, for each application class from the proposed algorithm, average transmit power per D2D pair is decreased by around 98.1% for the casual class, 97.8% for the interactive class, and 96.9% for the streaming class, compared to the Stackelberg algorithms in [109] and [119]. Overall, D2D pairs reduce their transmit power to a Pareto-efficient outcome, while guaranteeing best social welfare across all D2D pairs.

Fig. 3.6 shows the dynamics of average maximum achievable spectral efficiency (channel rate) per D2D pair, without considering priority. We observe that as the leader's charging price converges to an optimal outcome and D2D pairs are switching to different cellular user's uplink resource blocks, the achievable spectral efficiency also increases per D2D pair on average. Each application class from the proposed Stackelberg algorithm increases average achievable spectral efficiency per D2D pair by around 10.2% for the casual class, 12% for the interactive class, and 13.7% for the streaming class, compared to the NPC algorithm. As shown in

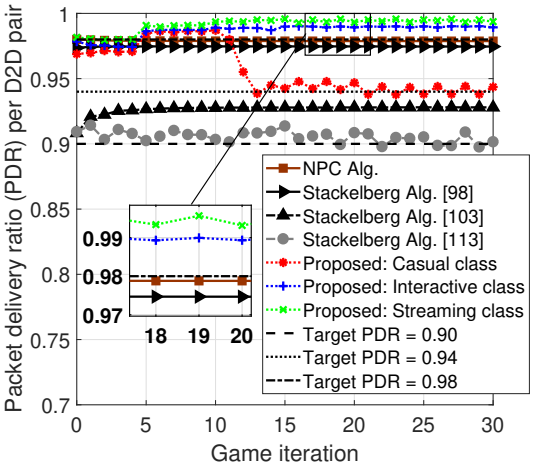


Figure 3.8: PDR per D2D pair, with priority, using coded QPSK.

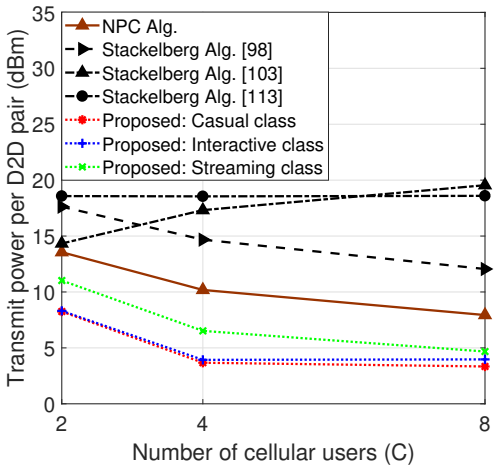


Figure 3.9: Transmit power per D2D pair under increasing number of cellular users when D=48, (no priority), using coded 16-QAM.

Fig. 3.6, the Stackelberg algorithms in [109] and [119] both maximise the average achievable spectral efficiency per D2D pair, compared to the proposed Stackelberg algorithm. However, considering the results in Fig. 3.5, the baseline Stackelberg algorithms in [109] and [119] yielded a larger transmit power, compared to the proposed distributed algorithm. Thus, each application class decreases average achievable spectral efficiency per D2D pair by around 19.3% for the casual class, 18% for the interactive class, and 16.8% for the streaming class, compared to the baseline Stackelberg algorithms in [109] and [119]. Moreover, due to the reduction in intra-cell interference and transmit power, as well as allowing D2D pairs to switch cellular user resource blocks, the proposed Stackelberg algorithm improves overall network performance.

In Fig. 3.7, we show the dynamics of average PDR per D2D pair, without considering priority. We observe that the proposed algorithm has different effects on PDR depending on the application class. In Fig. 3.7, D2D pairs in the proposed Stackelberg algorithm converge to a Pareto-efficient outcome, while maintaining their PDR above target PDR, $pdr_{tgt} = 0.9$. D2D pairs who selected the streaming and iterative application classes maximise PDR over the casual application class in the proposed Stackelberg algorithm, and all baseline algorithms considered. Hence, D2D pairs in the streaming and interactive application classes increase PDR per D2D pair by around 1.3% compared to the NPC baseline algorithm on average, and by around 5.2% compared to all Stackelberg baseline algorithms on average. Thus,

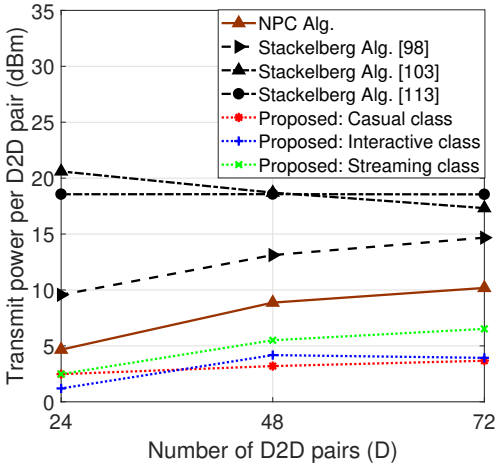


Figure 3.10: Transmit power per D2D pair under increasing number of D2D pairs when C=8, (no priority), using coded 16-QAM.

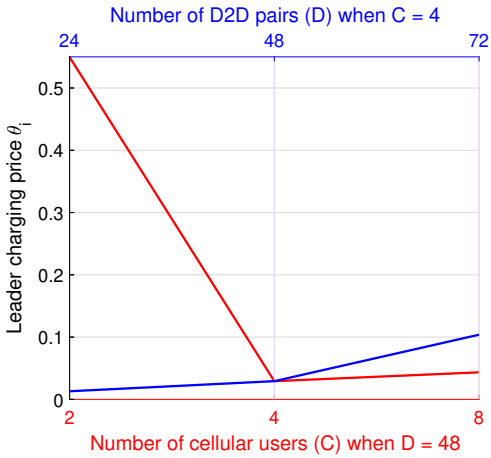


Figure 3.11: Leader's charging price assigned to the followers, under increasing number of cellular users and D2D pairs, (no priority), using coded QPSK.

the proposed streaming and interactive application classes have improved communication reliability and higher achievable spectral efficiency with little latency. However, the casual application class decreases PDR per D2D pair by around 2.4% compared to the NPC baseline algorithm, and increases PDR by around 1.3% compared to all baseline Stackelberg algorithms. The proposed casual application class maintains desired communication reliability for all users (that is, above minimum target PDR).

Fig. 3.8 shows the dynamics of average PDR per D2D pair, considering priority for each application class in the proposed Stackelberg algorithm. In particular, the casual application class is assigned a minimum target PDR of 0.9; interactive application class, minimum target PDR of 0.94; and streaming application class, minimum target PDR of 0.98. Note that, all baseline algorithms maintain a target PDR of 0.9. In Fig. 3.8, the outcome for the streaming application class, in our proposed Stackelberg algorithm, achieves its minimum target PDR of 0.98 and increases average PDR per D2D pair by around 6.3% compared to all baseline Stackelberg algorithms. As a result, D2D pairs who selected the streaming application class in the proposed Stackelberg algorithm will have further improved communication reliability, priority, higher achievable spectral efficiency, and reduced latency. For the other application classes, the proposed algorithm exceeds their minimum target PDRs. Moreover, from Fig. 3.8, the proposed algorithm provides more flexibility in terms of priority and reliability depending on the D2D

pair application, while also ensuring QoE, compared to the baseline algorithms.

Fig 3.9 shows average transmit power per D2D pair, under different number of cellular users in the network. We observe that as the number of cellular users increase (fixed number of D2D pairs) under the current network conditions, D2D pair transmit power decreases. This is due to the number of D2D pairs reusing a cellular user's resource block also decreasing, and thus, reducing intra-cell interference between users in the network. On average, the proposed Stackelberg algorithm decreases transmit power per D2D pair by around 69.8% for the casual class, 67.9% for the interactive class, and 56% for the streaming class, compared to all the baseline Stackelberg algorithms.

In Fig.3.10, we show average transmit power per D2D pair under different number of D2D pairs in the network. We observe that as the number of D2D pairs increases (fixed number of cellular users) under the current network conditions, D2D pair transmit power also increases. This is due to the number of D2D pairs reusing the cellular user's uplink resource block also increasing, and thus causing intra-cell interference to increase. Thus, the proposed Stackelberg algorithm decreases average transmit power per D2D pair by around 81.3% for the casual and interactive classes, and 70.9% for the streaming class, compared to all the baseline Stackelberg algorithms.

Moreover, in Fig. 3.11, we analyse the effects on the leader's charging price assigned to the followers, under an increasing number of cellular users and D2D pairs. We observe that, as the number of cellular users increase, the leader's charging price assigned to the followers reduces. This is due to the intra-cell interference reducing, and thus, causing D2D pair transmit power to also reduce, as observed in Fig, 3.9. On the other hand, as the number of D2D pairs increase, we observe that the leader's charging price assigned to the followers increases, and thus, intra-cell interference also increases due to increasing D2D pair transmit power, as outlined in Fig. 3.10.

3.7 Summary of Contributions

In this chapter, we have developed a novel dynamic multiple-objective Stackelberg game with arbitrarily different utility functions available to followers, to enable more efficient and flexible application-based resource allocation in D2D commu-

nications. Hence, in this chapter we answer questions Q1 and Q2, posed in Section 1.2.

In the proposed dynamic multiple-objective Stackelberg game, D2D pairs, following the BS (leader), were categorised into one of three application classes in order to enhance D2D user QoE and provide best service according to their communications requirements. Each application class was assigned an arbitrarily different utility function, and thus, it was not feasible to scalarise the three utility functions into a single combined utility function. Based on the flexible resource allocation Stackelberg game – with non-scalarised multi-criteria optimisation across all followers – we have proposed a distributed algorithm, in which D2D pairs aim to minimise transmit power, maximise data rate, and guarantee QoE. We have shown that the proposed distributed algorithm converges to a unique sub-game perfect Stackelberg Equilibrium. Simulation results showed that, on average, the proposed Stackelberg algorithm reduces transmit power per D2D pair, and increases PDR and achievable spectral efficiency per D2D pair, while guaranteeing best social welfare and QoE across all D2D followers, compared to the baseline algorithms.

Chapter 4

D2D Mode Selection and Resource Allocation in D2D Networks

Chapter 3 investigated the resource allocation and interference management problem for D2D pairs in reuse mode only, while also ensuring D2D end-user satisfaction for a range of different applications. However, in practice, D2D pairs not only transmit in reuse mode, they can also transmit in cellular mode or dedicated mode. Thus, D2D mode selection and resource allocation is another crucial area for investigation, due to intra-cell interference and the allocation of orthogonal and non-orthogonal resource blocks.

In this chapter, we jointly optimise D2D mode selection, resource allocation, and power control in a distributed and autonomous manner. Thus, this chapter solves this joint optimisation problem for network-assisted D2D communications, using two game theoretic approaches, i.e., a non-cooperative game and a coalitional game. Firstly, we propose a non-cooperative cross-layer repeated game, which combines a non-cooperative power control game and a two-armed bandit game, which enables D2D pairs to selfishly and independently select a transmission mode, while also minimising transmission power. The proposed non-cooperative game based algorithm is shown to converge to a unique Nash Equilibrium, that is Pareto-optimal. Secondly, we propose a cross-layer coalitional game, based on a leave-and-join approach, to cluster D2D pairs based on resource block allocation and D2D mode selection. Hence, in this game within each cluster, D2D pairs cooperate and share information in order to improve each cluster member's overall payoff. The proposed cross-layer coalitional game based algorithm, is shown to converge to a

Nash stable coalition partition that is socially efficient and has a non-empty core.

Moreover, in this chapter, we investigate user-assisted D2D communications as an alternative viable solution for D2D mode selection, when clustering D2D pairs. The user-assisted scenario requires the BS to only have partial knowledge of the network (and partial CSI), which can significantly reduce the signalling overhead and network complexity, compared to a network-assisted scenario (i.e., an ideal scenario).

This chapter is organised as follows. In Section 4.1, the system model for underlaid D2D communications is presented for D2D pairs in cellular mode, dedicated mode, and reuse mode. Section 4.2 presents the non-cooperative cross layer repeated game, for a network-assisted scenario. In this section, we outline the two-armed bandit game and the non-cooperative power control game, and we analyse the stability and convergence of the proposed non-cooperative game. On the other hand, Section 4.3 presents a dynamic cross-layer coalition formation game, for both a network-assisted and a user-assisted scenario. In this section, we outline the proposed coalitional game, and analyse the proposed game stability. In the proposed coalitional game, we design a more comprehensive interference management scheme, and also propose a dedicated mode resource allocation algorithm. Section 4.4 outlines the proposed distributed cross-layer coalitional algorithm. The stability and computational cost of the proposed coalitional algorithm is analysed in Section 4.5. Simulation results for the non-cooperative game and the coalitional game are illustrated and compared in Section 4.6. Moreover, in the simulation results, we also compare the two network scenarios in terms of network performance. Finally, Section 4.7 concludes the chapter.

4.1 System Model

Consider a single BS serving C cellular users and D D2D pairs, where D2D communication underlays the cellular network. In the network, each cellular user j from the set $\mathcal{C}$ (i.e. the user accessing the BS) and each D2D pair i from the set $\mathcal{D}$, are independently and uniformly distributed throughout the network, as illustrated in Fig. 4.1. The number of cellular users is considered to be less than the number of D2D pairs in the network [99], $C < D$. We focus on uplink resource allocation/sharing, as uplink resources are not as heavily utilised as downlink re-

sources [16, 35]. In addition, the BS can also handle some of the extra interference caused by D2D communications utilising the uplink resources [16,34,35]. D2D pairs are able to select one of three transmission modes to operate in, that is, (i) cellular mode; (ii) dedicated mode; or (iii) reuse mode. D2D pairs operating in cellular mode will communicate via the BS, whereas D2D pairs operating in dedicated mode or reuse mode will communicate directly between the transmitter and receiver. In order for D2D pairs to operate in dedicated mode or reuse mode, the distance $\zeta_{a,b}$ between transmitter a and receiver b must be less than or equal to a threshold distance $\overline{\zeta}$, $\zeta_{a,b} \leq \overline{\zeta}$. If this is not the case, $\zeta_{a,b} > \overline{\zeta}$, D2D pairs must operate in cellular mode (assuming enough resource blocks are available) until the transmitter or receiver moves closer. Moreover, the BS and all transmitters/receivers are equipped with omni-directional antennas, as device orientation is not pertinent to our model and use-cases.

In the network, the BS assigns R orthogonal resource blocks from the set $\mathcal{R}$, to cellular users and D2D pairs. Each cellular user j is allocated at most one resource block for uplink transmission from the set $\mathcal{R}$, where $C < R$. If a D2D pair chooses to operate in cellular mode, the BS will assign an orthogonal resource block from the set $\mathcal{R}$ to the D2D pair. Whereas, if a D2D pair chooses to operate in dedicated mode, then the D2D pair must wait to see if there are any remaining resource blocks that are not being used, before being allocated an orthogonal resource block. Otherwise, the D2D pair in dedicated mode, may have to wait till a resource block become available or choose to operate in reuse mode. Hence, D2D pairs in cellular mode and dedicated mode do not receive or cause any interference within the network, as they are each assigned an orthogonal resource block. On the other hand, the BS allocates D2D pairs operating in reuse mode non-orthogonal resource blocks, that is, D2D pairs in reuse mode will reuse/share part of a cellular user's uplink resources. We assume that each D2D pair in reuse mode can only share part of once cellular user's resource block at a time, while more than one D2D pair in reuse mode can share the same cellular user's resources. Therefore, D2D pairs operating in reuse mode will cause the BS to suffer interference, while paired D2D receivers will suffer interference from the cellular user and all other non-paired D2D transmitters sharing the same resource block (intra-cell interference). The interference generated from other communication cells within the network (inter-cell interference) is out of the scope of this paper, and it is assumed to be managed

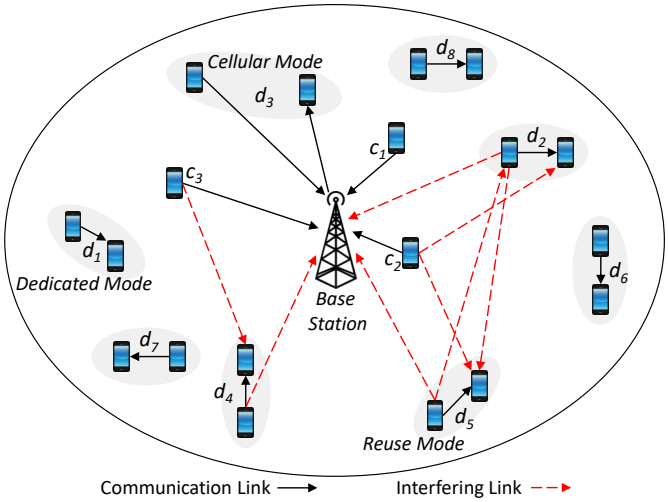


Figure 4.1: D2D communication system topology for all three transmission modes. The intra-cell interference for cellular users c_2 and c_3 is outlined in red. Note that, this system topology can be used for both network-assisted and user-assisted scenarios.

by a suitable inter-cell interference management scheme, based on power control and resource allocation, such as in [85] and [33].

Fig. 4.1 is an illustrative example of D2D communication system topology, where three cellular users (c_1, c_2, c_3) are transmitting to the BS (uplink), four D2D pairs (d_1, d_6) are operating in dedicated mode, three D2D pairs (d_2, d_4, d_5, d_7, d_8) are operating in reuse mode, and one D2D pair (d_3) is operating in cellular mode. Cellular user c_1 shares it's resources with D2D pair d_7. D2D pairs d_2 and d_5 both reuse the same uplink resources from cellular user c_2, and D2D pairs d_4 and d_8 shares the same uplink resources as cellular user c_3. The intra-cell interference generated, due to resource block sharing (D2D pairs in reuse mode) is also outlined for cellular user c_1 in Fig. 4.1. The example scenario considered in Fig. 4.1 is quite simple and can be extended to a more complex situation with many more users and BSs.

The channel rate of D2D pair i in reuse mode, reusing the uplink resources of cellular user j, is given as follows:

$$r_{di,RM} = \log_2\left(1+ \right. \tag{4.1}$$

$$\left. \frac{p_{di}|g_{di,di}|^2}{p_{cj}|g_{cj,di}|^2 + \sum_{\substack{x\in\mathcal{X} \\ x\neq i}} p_{dx}|g_{dx,di}|^2 + N_0}\right). \tag{4.2}$$

The uplink channel rate of cellular user j who is sharing its uplink resource with a set $\mathcal{X}$ of X D2D pairs, is given by:

$$r_{cj,RM} = \log_2 \left(1 + \frac{p_{cj} |g_{cj,BS}|^2}{\sum_{x \in \mathcal{X}} p_{dx} |g_{dx,BS}|^2 + N_0} \right), \tag{4.3}$$

where p_{di} is the transmit power of D2D pair i; p_{cj} is the transmit power of cellular user j; a set $\mathcal{X}$ of X D2D pairs reusing the same cellular user's resource block; p_{dx} is the interfering transmit power of D2D pair $d_x \in \mathcal{X}$; and N_0 is the AWGN power.

The channel gain, $g_{a,b}$, between transmitter a and receiver b, is considered to be a slowly time-selective flat Rayleigh fading channel model (small-scale fading), given by:

$$g_{a,b} = A_{PL} A_{SSF} \left(\frac{\zeta_0}{\zeta_{a,b}} \right)^{\frac{\alpha}{2}}, \tag{4.4}$$

where A_{PL} is the free space path loss channel attenuation; A_{SSF} is the small-scale fading channel attenuation with fixed normalised Doppler spread; ζ_0 is the reference distance between the transmitter and receiver; and α is the path loss exponent.

Moreover, the channel rate of D2D pair i in dedicated mode, is given by:

$$r_{di,DM} = \log_2 \left(1 + \frac{p_{di} |g_{di,di}|^2}{N_0} \right). \tag{4.5}$$

D2D pairs in cellular mode communicate via the BS, hence the channel rate takes into account the uplink SINR $\gamma_{di,UL}$ (at the BS) and downlink SINR $\gamma_{di,DL}$ (at the D2D pair receiver). The channel rate of D2D pair i in cellular mode is given by:

$$r_{di,CM} = \log_2 \left(1 + \min\{\gamma_{di,UL}, \gamma_{di,DL}\} \right), \tag{4.6}$$

$$\gamma_{di,UL} = \frac{p_{di,UL} |g_{di,BS}|^2}{N_0}, \quad \gamma_{di,DL} = \frac{p_{BS} |g_{BS,di}|^2}{N_0}, \tag{4.7}$$

$p_{di,UL}$ is the uplink transmit power of D2D pair transmitter i; and p_{BS} is the downlink transmit power of the BS. We assume that the BS transmit power is much larger than D2D pair transmit power, as the BS is less resource constrained than D2D users [94, 95]. Thus, the channel rate of D2D pair i in cellular mode is given by: $r_{di,CM} = \log_2 (1 + \gamma_{di,UL})$.

The uplink channel rate of cellular user j whose resources are not being shared

with any D2D pair/s, is given as follows:

$$r_{cj} = \log_2 \left(1 + \frac{p_{cj} |g_{cj,BS}|^2}{N_0} \right). \tag{4.8}$$

Even though D2D communications offloads traffic and demand on the BS, D2D users require assistance in selecting a particular transmission mode to operate in. In fact, D2D pairs operating in reuse mode can cause severe intra-cell interference, which is a major problem in D2D communications. Thus, D2D pairs in reuse mode must find a suitable cellular user's uplink resource block to reuse, while ensuring cellular user performance does not degrade. Power control is an effective tool to reduce intra-cell interference, which also reduces user transmit power. To ensure interference within the network will not degrade any user's QoS, we place a threshold on channel rate for both cellular users and D2D pairs, as follows:

$$r_{cj} \geq \overline{r}_c \; \forall \; j \in \mathcal{C}, \tag{4.9}$$

$$r_{di} \geq \overline{r}_d \; \forall \; i \in \mathcal{D}, \tag{4.10}$$

where channel rate and SINR are dependent on one another, such that threshold SINR is given as follows, $\overline{\gamma} = 2^{\overline{r}} - 1$. Hence, we assume that, cellular users must achieve target channel rate $\overline{r}_c$ before supporting any D2D pairs in reuse mode. If any cellular user is operating below target channel rate, this will cause network performance to degrade and intra-cell interference to significantly increase.

4.2 Non-cooperative Cross-Layer Repeated Game

Here, we propose a non-cooperative cross-layer repeated game. In this game, we adopt two games to make up our proposed game, which are a non-cooperative power control game and a two-armed bandit game. Both games are applied simultaneously to solve the cross-layer mode selection and power control algorithm, such that D2D users can minimise their transmit power and operate in ideal transmission mode. The proposed game first solves the non-cooperative power control game for both cellular and D2D users, i.e. maximising each player's utility function for each transmission mode with respect to their selected action (transmit power), thus obtaining updated transmit powers for all cellular and D2D users. D2D users

in the non-cooperative power control game calculate the utility function for both transmission modes. The game then derives ideal mode selection using the results of D2D users from the non-cooperative power control game in the two-armed bandit game. Note that, in the proposed non-cooperative cross-layer repeated game, we consider only reuse mode and cellular mode. Hence, depending on the outcome of the two-armed bandit game for each D2D user, this will determine which transmission mode is selected and the corresponding transmit power that will be updated for the next time slot.

The non-cooperative cross-layer repeated game is defined as:

Definition 4.1 *The proposed non-cooperative cross-layer repeated game is defined by* $G_{MS\text{-}NPC} = (\mathcal{N}, \mathcal{A}_i, u_i(\cdot))$*, where* $\mathcal{N}$ *is the finite set of players;* $\mathcal{A}_i$ *is the action set for each player* $i \in \mathcal{N}$*; and* $U_i(\cdot)$ *is the individual utility function for D2D pairs and cellular users.*

The proposed non-cooperative game is finitely repeated for T times, $0 \leq t \leq T$, and contains imperfect information, which means that all cellular and D2D users select their actions simultaneously (i.e. players do not know what their neighbours selected actions are). The finite set of players for the proposed non-cooperative game consists of cellular users and D2D pairs, i.e., $\mathcal{N} = \mathcal{C} \cup \mathcal{D}$. The finite strategy set for each D2D pair $i \in \mathcal{D}$ consists of two strategies, i.e., one strategy from each game adopted, $\mathcal{A}_i = [A_{\mathrm{pc}_i}, A_{\mathrm{arm}_i}]$. Hence, the strategy set for the non-cooperative power control game is transmit power, which is denoted as, $A_{\mathrm{pc}_i} = [p_{\min}, p_{\max}]$. On the other hand, the strategy set for the two-armed bandit game is D2D mode selection, where D2D pairs can select to operate in either reuse mode or cellular mode. Each adopted game has their own utility, thus, for each player $i \in \mathcal{N}$ the utility for the proposed non-cooperative game, is given by, $u_i(\cdot) = [u_{\mathrm{pc}_i}(\cdot), u_{\mathrm{arm}_i}(\cdot)]$. Each utility function best describes the performance of the D2D pair for the respective game.

4.2.1 Non-cooperative Power Control Game

The non-cooperative power control game $G_{pc} = [\mathcal{N}, A_{\mathrm{pc}_i}, u_{\mathrm{pc}_i}(\cdot)]$, focuses on minimising transmit power for both cellular and D2D users such that the game produces a power efficient outcome and minimises interference further. The action set implemented for each player is transmit power, which is a finite set and is bounded by

a minimum transmit power and maximum transmit power, $p_i^{min} \leq p_i \leq p_i^{max}$. The utility function for the non-cooperative power control game for both cellular and D2D users, is defined in [167]. The chosen utility function is a trade-off between maximising SINR and minimising transmit power, which will ensure good signal quality at the receiver without large latency periods. The utility function, $u_{\mathrm{pc}_{cj}}$ for cellular user $j \in \mathcal{C}$, and the utility function, $u_{\mathrm{pc}_{di,\diamond}}$, for D2D pair $i \in \mathcal{D}$, are given by:

$$u_{\mathrm{pc}_{cj}}(p_{cj}) = -w_i p_{cj} - v_i \left(\overline{\gamma}_c - \gamma_{cj}\right)^2, \tag{4.11}$$

$$u_{\mathrm{pc}_{di,\diamond}}(p_{di,\diamond}) = -w_i p_{di,\diamond} - v_i \left(\overline{\gamma}_d - \gamma_{di,\diamond}\right)^2, \tag{4.12}$$

where $\diamond$ represents a D2D pair in either cellular mode (CM), $u_{\mathrm{pc}_{di,CM}}$ or reuse mode (RM) $u_{\mathrm{pc}_{di,RM}}$); $w_i > 0$ and $v_i > 0$ are positive constants to ensure all parts of (4.11) and (4.12) have the same magnitude; and $\overline{\gamma} - \gamma_i$ is the SINR error, which is the difference between target SINR and received SINR of user $i \in \mathcal{N}$.

The transmit power for p_i of each user $i \in \mathcal{N}$ belongs to a non-empty, convex, and compact subset of Euclidean space $\mathbb{R}^{|\mathcal{N}|}$. Each user i finds the best response transmit power, by maximising its utility function, u_{pc_i}, with respect to transmit power, as follows:

$$\begin{aligned} \underset{p_i}{\text{maximise}} \quad & u_{\mathrm{pc}_i}(p_i) \quad \forall i \in \mathcal{N} \\ \text{subject to} \quad & p_{\min} \leq p_i \leq p_{\max}. \end{aligned}$$

Therefore, once optimal transmit power has been found for each user $i \in \mathcal{N}$, each D2D pair can then determine a transmission mode, using the two-armed band game theoretic approach.

4.2.2 Two-Armed Bandit Game

The two-armed bandit game $G_{\mathrm{arm}} = [\mathcal{N}, A_{\mathrm{arm}_i}, u_{\mathrm{arm}_i}(\cdot)]$, focuses on selecting an arm using a Poisson distribution, as defined in [90]. In this game, we considered only two arms, as D2D pairs can either select to operate in cellular mode or reuse mode. Note that, in this game, cellular users are always fixed in cellular mode, i.e., traditional cellular communications. The arm strategy for D2D pair i selecting reuse mode, is given by, $\alpha_{di,RM}(t) \in \{0, 1\}$, whereas the arm strategy for D2D pair i selecting cellular mode, is dependent on the outcome of the reuse mode arm strategy and is given by, $\alpha_{di,CM}(t) = 1 - \alpha_{di,RM}(t)$ [90]. To solve the proposed

two-armed bandit game, a balancing method is adapted from [90], which is based on a Levy-bandit approach. Hence, a D2D pair i will determine its arm strategy, as follows:

$$\alpha_{di,RM}(t) = \begin{cases} 0, & \text{if } \rho(t) < \rho^*, \\ 1, & \text{if } \rho(t) > \rho^*, \end{cases} \tag{4.13}$$

where $\rho(t)$ is the likelihood of reuse mode being selected more than cellular mode (i.e., the belief of a D2D pair at any time), and ρ^* is the threshold belief. Thus, the belief of D2D pair i, is defined as [90]:

$$\rho(t) = \frac{\rho(t-1)\exp(-\alpha_{di,RM}(t)\beta\delta)}{1-\rho(t-1)+\rho(t-1)\exp(-\alpha_{di,RM}(t)\beta\delta)}, \tag{4.14}$$

where $\rho(t-1)$ is the prior belief of D2D pair i, β is an intensity factor, and δ is the time step duration. Furthermore, the threshold belief of D2D pair i, ρ^*, is a function of D2D pair i's utility from the non-cooperative power control game, for both cellular mode and reuse mode, and is given by [90]:

$$\rho^\star = \frac{\omega \max(u_{\mathrm{pc}_{di,CM}})}{(\omega+1)(\max(u_{\mathrm{pc}_{di,RM}}) - \max(u_{\mathrm{pc}_{di,CM}})) + \omega \max(u_{\mathrm{pc}_{di,CM}})}, \tag{4.15}$$

where $\omega = \frac{r}{\beta}$ and r is a discounted reward; $\max(u_{\mathrm{pc}_{di,CM}})$ is the maximisation of the utility for D2D pair i in cellular mode with respect to transmit power; and $\max(u_{\mathrm{pc}_{di,RM}})$ is the maximisation of the utility for D2D pair i in reuse mode with respect to transmit power.

In the two-armed bandit game, the utility function u_{arm_i} is referred to as an expected accumulated discounted reward, as defined in [90] and [172]. The expected accumulated discounted reward for all D2D users in any transmission mode given a prior belief, is as follows [90, 172]:

$$\begin{aligned} u_{\mathrm{arm}_i} &= \text{maximise}(E[R_i]), \\ \therefore u_{\mathrm{arm}_i} &= E\left[\int_{t=0}^{T} r\exp(-rt)\left((1-\alpha_{di,RM})u_{\mathrm{pc}_{di,CM}} + \alpha_{di,RM}u_{\mathrm{pc}_{di,RM}}\rho\right)dt|\rho(0)\right], \end{aligned} \tag{4.16}$$

where the expected reward is defined in terms of the current selected mode, and utility functions from the non-cooperative power control game for both cellular mode and reuse mode. We have modified the upper limit of the integral from ∞

to T, as t is finite and bounded between 0 and T in our proposed model.

In addition to maximising the expected accumulated discounted reward for all D2D users, the regret bounds should also be minimised. Regret is defined as the difference between the expected accumulated discounted reward and the optimal arm strategy [173,174]. In the context of the proposed two-armed bandit algorithm, when a D2D user is determining the best transmission mode (arm strategy) based on the reward in (4.16), the D2D user will regret choosing a transmission mode (arm strategy) that is not the optimal strategy.

4.2.3 Non-Cooperative Cross-Layer Repeated Game Stability Analysis

To solve the proposed non-cooperative cross-layer repeated game, we consider the concept of a Nash equilibrium. The Nash equilibrium is a stable point in the game, where no player has incentive to increase their payoffs without changing their strategy unilaterally, given all other player's strategies. The Nash equilibrium for our proposed non-cooperative cross-layer repeated game is defined as:

Definition 4.2 The combined transmit power and mode strategy vector $\mathbf{A}^* = [\mathbf{A}_{\text{pc}}, \mathbf{A}_{\text{arm}}] = [(p_1^*, \alpha_1^*), (p_2^*, \alpha_2^*), \ldots, (p_N^*, \alpha_N^*)]$ for every player $i \in \mathcal{N}$, there exists a Nash Equilibrium for each stage of the proposed non-cooperative repeated game, if the following is satisfied:

$$u_i([p_i^*, \alpha_i^*], [\boldsymbol{p}_{-\boldsymbol{i}}^*, \boldsymbol{\alpha}_{-\boldsymbol{i}}^*]) \geq u_i([p_i, \alpha_i], [\boldsymbol{p}_{-\boldsymbol{i}}^*, \boldsymbol{\alpha}_{-\boldsymbol{i}}^*]) \qquad \forall i \in \mathcal{N}, \tag{4.17}$$

where p_i^* is an element of $\boldsymbol{p}^*$ which is the best response transmit power of user $i \in \mathcal{N}$; $\boldsymbol{p}_{-\boldsymbol{i}}^*$ are elements of $\boldsymbol{p}^*$ which are the best response transmit powers of all other users, other than user i; and α_i^* is an element of $\boldsymbol{\alpha}^*$ which is the best response mode selection of D2D pair $i \in \mathcal{D}$.

Remark 4.1 *Applying both utility functions, for the non-cooperative power control game (at the physical layer) and the two-armed bandit game (at the network layer), results in a unique Nash Equilibrium.*

Proposition 4.1 *The unique Nash Equilibrium, i.e., the best response for each player, is fully defined over all game stages.*

Based on the above definition and proposition, each user $i \in \mathcal{N}$ aims to maximise their corresponding utility functions, by choosing optimal transmit power for cellular users, and optimal transmit power and mode selection for D2D pairs. We will analyse each utility function from the non-cooperative power control game and the two-armed bandit game, to show that there exists a Nash Equilibrium for the proposed non-cooperative cross-layer repeated game.

Proposition 4.2 *The utility function defined in the non-cooperative power control game for any user $i \in \mathcal{N}$ is strictly concave and continuous with respect to user i's transmit power.*

Proof: See Appendix B.1. ■

Proposition 4.3 *The utility function defined in the two-armed bandit game for any user $i \in \mathcal{N}$ is strictly concave and continuous with respect to user i's belief.*

Proof: See Appendix B.2 ■

Remark 4.2 *For the game to have a Pareto-efficient outcome, both the non-cooperative power control game and the two-armed bandit game need to be included.*

Proposition 4.4 *The game is Pareto-efficient when transmit power is minimised and optimal mode is selected in our proposed non-cooperative cross-layer repeated game.*

Proof: A Pareto-efficient outcome is fully defined over the entire action set and players, according to the proposed algorithm, where after a finite number of stages T, operating at the unique Nash Equilibrium for both power and mode, implies that the outcome is Pareto-efficient across all players. ■

4.3 Cross-Layer Coalitional Game Formulation

Here, we propose a fully distributed cross-layer resource utilisation framework based on coalitional game theory [35, 91–96, 99, 101] for D2D communications. The proposed cross-layer coalitional game will enable D2D pairs to cooperatively form clusters (coalitions) based on mode selection and resource allocation, in a fully distributed manner. Within a cluster, D2D pairs cooperate and share resources fairly

amongst one another, in order to improve D2D user performance and in particular cellular user performance. The proposed coalitional game model is applied to network-assisted D2D communications.

However, since network-assisted D2D communications assumes that the BS has global knowledge (complete CSI) of the network, this approach may lead to signalling overhead and network complexity to significantly increase. Additionally, due to the exponential increase in the number of users in the network, thus, the BS will be required to assist more D2D pairs in selecting a transmission mode. In fact, in practical scenarios, the BS may not be able to obtain global knowledge of the network (complete CSI) in a timely manner, or when the network is large scale. Therefore, we extend the proposed coalitional game to also consider user-assisted D2D communications, which enables D2D pairs to select a suitable transmission mode to operate in without BS assistance (i.e., partial CSI). In the user-assisted scenario, the role of the BS is to assist D2D pairs with authentication processes, scheduling, and resource allocation, that is, the fundamental mechanisms of communications with partial CSI [16,31]. Although user-assisted D2D communications may cause D2D pairs to increase power consumption and latency [16], the proposed user-assisted scenario will enable a realistic clustering approach for D2D pairs, such that signalling overhead and network complexity is not increased.

4.3.1 Cross-Layer Coalitional Game Overview

In this section, we introduce the proposed CLC game framework to cooperatively cluster D2D pairs, in order to reduce transmission power, increase channel rate, and utilise resources more efficiently. The proposed CLC game combines a mode selection and resource allocation coalitional game (main game), along with a dedicated mode resource allocation sub-problem and an interference management sub-problem. Moreover, we outline how the main coalitional game interacts with each sub-problem (and vice versa).

Let $\mathcal{N} = \mathcal{C} \cup \mathcal{D}$ be the set of players, which consists of all cellular users and D2D pairs within the network. Within the proposed game, D2D pairs aim to form cooperative clusters based on mode selection and resource allocation, where D2D pairs and cellular users form a set $\mathcal{F}$. We assume a fixed cluster structure, where the set of clusters $\mathcal{F}$ is a collection of F disjoint clusters. That is, the number of clusters considered, equals the number of orthogonal resource blocks allocated to

users in the network. A cluster $\mathcal{F}_\pi \in \mathcal{F}$ is defined as a subset of $\mathcal{N}$, such that $\bigcup_{\pi=1}^{R} \mathcal{F}_\pi = \mathcal{N}$, where $\forall \mathcal{F}_\pi, \mathcal{F}_{\pi'} \in F$, if $\mathcal{F}_{\pi'} \neq \mathcal{F}_\pi$, then $\mathcal{F}_{\pi'} \cap \mathcal{F}_\pi = \emptyset$. Moreover, the set of clusters $\mathcal{F}$ can be broken down into smaller subsets based on D2D pair mode selections, i.e., $\mathcal{F} = \mathcal{D}_{RM} \cup \mathcal{D}_{CM} \cup \mathcal{D}_{DM}$ where $\mathcal{D}_{RM}$ is the set of D2D pairs selecting reuse mode, $\mathcal{D}_{CM}$ is the set of D2D pairs selecting cellular mode, and $\mathcal{D}_{DM}$ is the set of D2D pairs selecting dedicated mode.

Definition 4.3 A *cross-layer coalition formation game* is defined by $G_{CLC} = (\mathcal{N}, v, \mathcal{F})$, where $\mathcal{N}$ is the finite set of players; v is the non-transferable coalition value that captures the individual utilities of D2D pairs and cellular users; and $\mathcal{F}$ is the set of coalitions.

Initially, cellular users are partitioned into singleton clusters (that is, the BS allocates each cellular user an orthogonal resource block), where all D2D pairs within close proximity (satisfying the threshold distance for direct communication and QoS constraints) are randomly partitioned across all clusters (that is, clustering based on resource sharing). Thus, forming the set $\mathcal{D}_{RM}$, such that $\bigcup_{\mathcal{F}_\pi \in \mathcal{D}_{RM}} \mathcal{F}_\pi = \mathcal{C}$. Any remaining D2D pairs not within close proximity will operate in cellular mode and are partitioned into singleton clusters (that is, the BS allocated D2D pairs in cellular mode an orthogonal resource block). Thus, forming the set $\mathcal{D}_{CM} \subset \mathcal{F}$.

Fig. 4.2 shows the proposed CLC game framework, and the relationships between the main game (red block) and the sub-problems (blue blocks). Within the CLC game, optimal D2D mode selection is determined in the mode selection and resource allocation coalitional game and the dedicated mode resource allocation sub-problem. Moreover, the coalitional game determines if D2D pairs are to operate in reuse mode or cellular mode, and the dedicated mode resource allocation sub-problem determines which D2D pairs in reuse mode will be allocated orthogonal spectrum.

Firstly the BS checks if there are any idle orthogonal resource blocks available within the network to allocate to D2D pairs in reuse mode (ensuring that there is a sufficient amount of resource blocks available if a cellular user joins the network or if D2D pairs switch to cellular mode). If resource blocks are available, the dedicated mode resource allocation sub-problem (Section 4.3.3.1) determines suitable D2D pairs operating in reuse mode to allocate an orthogonal resource block to. Thus, D2D pairs switch transmission mode from reuse to dedicated, which directly maps to partitioning D2D pairs into singleton clusters and forming the set of D2D

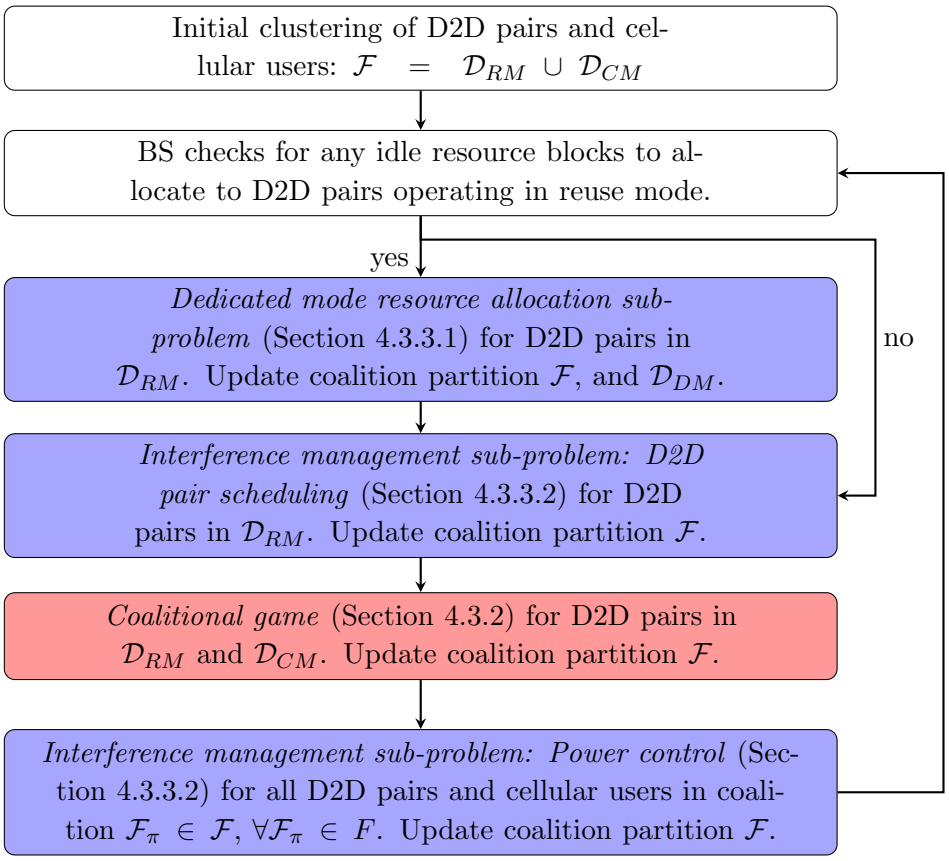


Figure 4.2: Proposed CLC game framework to solve resource utilisation in D2D communications. The red block indicates where the network-assisted or user-assisted schemes are considered.

pairs operating in dedicated mode $\mathcal{D}_{DM} \subset \mathcal{F}$. Next, the interference management sub-problem schedules D2D pairs in reuse mode across all clusters, in order to reduce intra-cell interference within the network (Section 4.3.3.2). Then, all D2D pairs operating in cellular mode $\mathcal{D}_{CM}$ and reuse mode $\mathcal{D}_{RM}$, determine their most preferred transmission mode or resource sharing cellular user, in the mode selection and resource allocation coalitional game (Section 4.3.2). Finally, the interference management sub-problem also considers a power control sub-game (Section 4.3.3.2), to determine optimal transmit power for each user in cluster $\mathcal{F}_\pi \in \mathcal{F}$, $\forall \mathcal{F}_\pi \in F$, while guaranteeing (4.9) and (4.10). Thus, after each game/sub-problem in the proposed CLC game, the coalition partition $\mathcal{F}$ is updated accordingly. The process is repeated for a finite number of stages.

The red block in Fig. 4.2, indicates a different implementation scheme, to accommodate for the network-assisted and user-assisted scenarios.

Remark 4.3 There are three types of clusters out of a total of R disjoint clusters, which reflect the three transmission modes for D2D pairs: (i) D2D pair i in dedicated mode forms a singleton set, $\{i\}$; (ii) D2D pair i in cellular mode forms a singleton set, $\{i\}$; (iii) D2D pair i in reuse mode forms a cluster with a total of X D2D pairs sharing the same resource block, which makes up the set $\mathcal{X}$, $\mathcal{X} \subset \mathcal{D}$, plus one cellular user.

4.3.2 Mode Selection and Resource Allocation Coalitional Game

In this section, we propose a D2D mode selection and resource allocation/sharing coalitional game to cooperatively cluster D2D pairs into coalitions, in order to reduce transmission power, increase channel rate, and utilise resources more efficiently. We consider the proposed coalitional game for both network-assisted and user-assisted D2D communication scenarios. This approach is fully distributed and autonomous, where D2D pairs self-organise into coalitions based on mode selection (reuse mode and cellular mode), resource allocation, reduced transmit power, and increased channel rate. The proposed mode selection and resource allocation coalitional game is formally defined as follows:

Definition 4.4 A *mode selection and resource allocation coalitional game* is defined by $G_C = (\mathcal{N}_s, v, \mathcal{S})$, where $\mathcal{N}_s$ is the finite set of rational players; v is the

coalition value; and $\mathcal{S}$ is the set of coalitions.

In our proposed mode selection and resource allocation coalitional game, the finite set of players $\mathcal{N}_s$ consists of cellular users, D2D pairs in reuse mode, and D2D pairs in cellular mode, $\mathcal{N}_s = \mathcal{C} \cup \mathcal{D}_{RM} \cup \mathcal{D}_{CM}$. We assume that cellular users are fixed in their allocated coalition and cannot change coalitions throughout the game. The value of coalition $\mathcal{S}_\pi \in \mathcal{S}$, $v(\mathcal{S}_\pi)$, is non-transferable and is a set of vectors capturing the individual D2D pair and cellular user utilities within a coalition $\mathcal{S}_\pi$. The set of coalitions $\mathcal{S}$ is a collection of S disjoint coalitions, such that $\mathcal{S} \subseteq \mathcal{F}$. A coalition $\mathcal{S}_\pi \in \mathcal{S}$ is defined as a subset of $\mathcal{N}_s$, such that $\bigcup_{\mathcal{S}_\pi \in \mathcal{S}} \mathcal{S}_\pi = \mathcal{N}_s$, where $\forall \mathcal{S}_\pi, \mathcal{S}_{\pi'} \in S$ if $\mathcal{S}_{\pi'} \neq \mathcal{S}_\pi$ then $\mathcal{S}_{\pi'} \cap \mathcal{S}_\pi = \emptyset$.

Each user within a coalition determines its utility, with respect to all other members in that coalition. In particular, each D2D pair must decide to either leave their current coalition $\mathcal{S}_\pi$ and joint a new coalition $\mathcal{S}_{\pi'}$, such that $\mathcal{S}_\pi \neq \mathcal{S}_{\pi'}$, or stay in their current coalition. This choice of deciding to leave-and-joint or stay, maps directly to the D2D mode selection problem for reuse mode and cellular mode. For example, if D2D pair i is currently operating in reuse mode, they have two choices, either to switch to cellular mode, switch to another cellular user's resource block to reuse, or stay in their current coalition. The proposed utility function for each user $i \in \mathcal{N}_s$, models the trade-off between maximising channel rate and minimising transmit power. The utility function $u_i(\mathcal{S}_\pi)$ for each user i in coalition $\mathcal{S}_\pi$, is given by:

$$u_i(\mathcal{S}_\pi) = \lambda_i r_i - \omega {p_i}^2, \tag{4.18}$$

where i is either a D2D pair or cellular user from coalition $\mathcal{S}_\pi \in \mathcal{S}$, $\mathcal{S}_\pi \neq \emptyset$; r_i is the channel rate of user i in coalition $\mathcal{S}_\pi$; p_i is the transmit power of user i in coalition $\mathcal{S}_\pi$; $\lambda_i > 0$ is a weighting factor that depends on network parameters and user $i \in \mathcal{S}_\pi$; and ω is a positive constant. Moreover, the utility for each user in coalition $\mathcal{S}_\pi$ depends only on the users in that coalition, this is due to intra-cell interference only being generated between users utilising the same resource block which are in the same coalition. Thus, each user's utility is not affected by members outside of their coalition (characteristic form).

Remark 4.4 The proposed utility function in (4.18) is strictly concave and continuous with respect to transmit power for each user. This property of the utility function is very crucial to the interference management sub-problem in Section 4.3.3.2.

The coalition value $v(\mathcal{S}_\pi)$ is a set of vectors capturing each members utility within coalition $\mathcal{S}_\pi$, and is a mapping over a nonempty closed convex subset of $\mathbb{R}^{|\mathcal{S}_\pi|}$. The value of coalition $\mathcal{S}_\pi$ is given as follows:

$$v(\mathcal{S}_\pi) = \left\{ \boldsymbol{u}(\mathcal{S}_\pi) \in \mathbb{R}^{|\mathcal{S}_\pi|} | u_i(\mathcal{S}_\pi), \forall i \in \mathcal{S}_\pi \right\}, \tag{4.19}$$

where $\boldsymbol{u}(\mathcal{S}_\pi)$ is a nonempty vector of space $\mathbb{R}^{|\mathcal{S}_\pi|}$; and $u_i(\mathcal{S}_\pi)$ is the utility of user i in coalition $\mathcal{S}_\pi$, defined in (4.18), and is an element of $\boldsymbol{u}(\mathcal{S}_\pi)$.

Remark 4.5 The coalition value defined for the coalition partition $\mathcal{S}$, can be extended to the proposed CLC game coalition partition $\mathcal{F}$. The characteristic form for each coalition $\mathcal{F}_\pi$ is a mapping over a nonempty closed convex subset of $\mathbb{R}^{|\mathcal{F}_\pi|}$. The value of coalition $\mathcal{F}_\pi$ is given by, $v(\mathcal{F}_\pi) = \left\{ \boldsymbol{u}(\mathcal{F}_\pi) \in \mathbb{R}^{|\mathcal{F}_\pi|} | u_i(\mathcal{F}_\pi), \forall i \in \mathcal{F}_\pi \right\}$, where $\boldsymbol{u}(\mathcal{F}_\pi)$ is a nonempty vector of space $\mathbb{R}^{|\mathcal{F}_\pi|}$; and $u_i(\mathcal{F}_\pi)$ is the utility of user i in coalition $\mathcal{F}_\pi$, defined in (4.18).

Remark 4.6 In the coalitional game, if all D2D pairs decided to operate in reuse mode and reuse only one cellular user's uplink resource block at the same time, this would result in a *resource sharing grand coalition*. This grand coalition would cause channel rate to reduce and transmit power to increase, which, in turn, would result in severe intra-cell interference in the network. Furthermore, the grand coalition would degrade cellular user QoS, and reduce system capacity, throughput, and D2D battery lifetime. Thus, a grand coalition is far from optimal for D2D communications.

4.3.2.1 Network-Assisted Scenario Coalitional Algorithm

In the network-assisted scenario the BS has global knowledge of the network, and is able to assist D2D pairs in selecting an appropriate transmission mode to operate in (either cellular or reuse mode), or which cellular users resource block to reuse (if reuse mode is selected). That is, the BS assists D2D pairs in deciding to either leave-and-join a new coalition or stay in their current coalition.

The proposed coalition algorithm for the network-assisted scenario is based on two rules, which are, *leave-and-join* and *Preference-order*. The leave-and-join rule allows D2D pairs to decided to switch coalitions (transmission mode or cellular user's resource block) if a set of conditions are satisfied. In order for the D2D pairs to make such a decision, the BS assists each D2D pair in generating a

Preference-order over all coalitions. A Preference-order is an ordered list of coalition preferences that each D2D pair prefers being a member of, a concept presented in [67]. The network-assisted Preference-order is defined as follows:

Definition 4.5 *Preference-order* is defined as a complete, reflexive and transitive binary relation over the set $\{\mathcal{S}_\pi \subseteq \mathcal{D} : i \in \mathcal{S}_\pi\}$ [72]. Preference-order is denoted by $\succ_i$, for any D2D pair $i \in \mathcal{D}$.

Within the coalitional game, D2D pair i has a Preference-order across all coalitions in the set $\mathcal{S}$. For D2D pair i to determine which coalition they prefer being a member of, they must calculate their utility in both coalitions, as follows:

$$\mathcal{S}_\pi \succ_i \mathcal{S}_{\pi'} \Leftrightarrow u_i(\mathcal{S}_\pi) > u_i(\mathcal{S}_{\pi'}), \tag{4.20}$$

where $\mathcal{S}_\pi, \mathcal{S}_{\pi'} \in \mathcal{S}$; $\mathcal{S} \subset \mathcal{F}$. Equation (4.20) states that the selected D2D pair i, strictly prefers being a member of coalition $\mathcal{S}_\pi$ over $\mathcal{S}_{\pi'}$, as D2D pair i obtains reduced transmit power and intra-cell interference, plus increased channel rate (that is, improved utility) in $\mathcal{S}_\pi$. The network-assisted leave-and-join rule is defined as follows:

Definition 4.6 Network-assisted leave-and-join - D2D pair i in coalition $\mathcal{S}_\pi \in \mathcal{S}$, leaves its current coalition and joins a newly preferred coalition, $\mathcal{S}_\pi \setminus \{i\} \in \mathcal{S}$, $\mathcal{S}_{\pi'} \cup \{i\} \in \mathcal{S}$, where $\mathcal{S}_\pi \neq \mathcal{S}_{\pi'}$ and $\pi \neq \pi'$, if and only if:

(i) The utility of D2D pair i satisfies the Preference-order rule in (4.20): $\mathcal{S}_{\pi'} \succ_i \mathcal{S}_\pi$.

(ii) The average utility of all members in the current coalition and the newly preferred coalition, must be unchanged or improved:

$$\overline{u}(\mathcal{S}_\pi \setminus \{i\}) \geq \overline{u}(\mathcal{S}_\pi),\ \overline{u}(\mathcal{S}_{\pi'} \cup \{i\}) \geq \overline{u}(\mathcal{S}_{\pi'}),$$

$$\therefore \overline{u}(\mathcal{S}_\pi) = \frac{\sum_{i \in \mathcal{S}_\pi} u_i(\mathcal{S}_\pi)}{|\mathcal{S}_\pi|}.$$

According to Condition (ii) in Definition 4.6, for D2D pair i to change coalitions, no user can achieve worse performance on average in coalitions $\mathcal{S}_\pi$ or $\mathcal{S}_{\pi'}$, that is, either increased transmit power or reduced channel rate. Average utility is

considered instead of sum utility or individual utility (as used in literature [93–95, 99]), as this causes D2D pairs in reuse mode to distribute themselves more evenly amongst the resource sharing coalitions (that is, the cellular users).

4.3.2.2 User-Assisted Scenario Coalitional Algorithm

The coalition algorithm for the user-assisted scenario, considers modified versions of network-assisted Preference-order and leave-and-join rules, as defined in Definition 4.5 and 4.6 respectively. In the user-assisted scenario, the BS has partial knowledge of the network and is unable to assist D2D pairs to select a transmission mode to operate in, or which cellular users resource block to reuse.

The user-assisted Preference-order for D2D pair i, considers its current coalition $\mathcal{S}_\pi$ with the highest preference, and all other coalitions in the network with equal and lower preference. Thus, due to the limited assistance of the BS the user-assisted scenario only considers one coalition at a time, which is D2D pair i's current coalition $\mathcal{S}_\pi$. Hence, D2D pair i must determine their utility for both reuse mode ($u_{i,RM}$) and cellular mode ($u_{i,CM}$), in order to detremine which transmission mode to operate in.

Using this modified Preference-order, we propose a user-assisted leave-and-join rule, which is divided in two parts depending on the transmission mode D2D pair i is currently in. If D2D pair i is currently operating in reuse mode, user-assisted leave-and-join rule part-I is defined as follows:

Definition 4.7 User-assisted leave-and-join part-I - D2D pair i in reuse mode coalition $\mathcal{S}_\pi \in \mathcal{S}$, leaves their current coalition and joins a new coalition, $\mathcal{S}_\pi \setminus \{i\} \in \mathcal{S}$, $\mathcal{S}_{\pi'} \cup \{i\} \in \mathcal{S}$, where $\mathcal{S}_\pi \neq \mathcal{S}_{\pi'}$ and $\pi \neq \pi'$, if and only if, one of the following conditions is satisfied for D2D pair i:

(i) If $u_{i,RM} > u_{i,CM}$ and $u_{i,RM} < u_{trgt}$, where $u_{i,RM} = u_i(\mathcal{S}_\pi)$, D2D pair i will stay in reuse mode and move to the next consecutive reuse mode coalition $\mathcal{S}_{\pi'}$. Since D2D pair i's utility in $\mathcal{S}_\pi$ is below a utility threshold u_{trgt}, QoS for the cellular and D2D user is not guaranteed, and thus D2D pair i must find another cellular user's resource block to reuse, $\mathcal{S}_{\pi'}$.

(ii) If $u_{i,CM} > u_{i,RM}$, where $u_{i,RM} = u_i(\mathcal{S}_\pi)$, D2D pair i will switch transmission mode from reuse mode ($\mathcal{S}_\pi$) to cellular mode ($\mathcal{S}_{\pi'}$), and will be allocated a free orthogonal resource block from the BS.

On the other hand, if D2D pair i is currently operating in cellular mode, the user-assisted leave-and-join rule part-II is defined as follows:

Definition 4.8 User-assisted leave-and-join part-II: D2D pair i in coalition (cellular mode) $\mathcal{S}_\pi \in \mathcal{S}$, leaves their current coalition and joins a new coalition (reuse mode), $\mathcal{S}_\pi \setminus \{i\} \in \mathcal{S}$, $\mathcal{S}_{\pi'} \cup \{i\} \in \mathcal{S}$, where $\mathcal{S}_\pi \neq \mathcal{S}_{\pi'}$ and $\pi \neq \pi'$, if and only if, $u_{i,CM} < u_{i,RM}$, where $u_{i,CM} = u_i(\mathcal{S}_\pi)$ and $u_{i,RM} = u_i(\mathcal{S}_{\pi'})$.

4.3.3 Sub-Problems in the CLC Game

In the following, we analyse each sub-problem in the proposed CLC game, in order to minimise intra-cell interference in the network, while improving system capacity and network performance. We assume a given coalition partition for each sub-problem, which is updated accordingly after each sub-problem is solved. The sub-problems in the CLC game include a dedicated resource allocation sub-problem, D2D pair scheduling sub-problem, and a power control sub-problem.

4.3.3.1 Dedicated Mode Resource Allocation Sub-Problem

The dedicated mode resource allocation sub-problem aims to find D2D pairs in reuse mode whose utility (payoff) is the worst, where the BS allocates R_{DM} orthogonal (dedicated) resource blocks that have been idle for some period of time, from the set $\mathcal{R}_{DM} \subset \mathcal{R}$, to those D2D pairs. D2D pairs in reuse mode with worst utility (payoff) will experience high transmit power and low channel rate, due to receiving severe intra-cell interference. Allocating orthogonal resource blocks to D2D pairs in reuse mode, maps to D2D pairs switching from reuse to dedicate mode. The BS ensures that there will still be a sufficient amount of resource blocks available within the network, for newly joined cellular users, or D2D pairs switching to cellular mode. Thus, allocating these D2D pairs to dedicated mode will improve overall network performance and system capacity, while reducing intra-cell interference in the network.

To determine which D2D pair $i \in \mathcal{D}_{RM}$ is allocated an orthogonal resource block from the set $\mathcal{R}_{DM}$, the following minimisation must be satisfied, as follows:

$$i^* = \min(\boldsymbol{u}(\mathcal{D}_{RM})), \tag{4.21}$$

where i^* is the selected D2D pair from $\mathcal{D}_{RM}$, and $\boldsymbol{u}(\mathcal{D}_{RM})$ is a vector of all utilities

of D2D pairs from $\mathcal{D}_{RM}$. Algorithm 2 solves the dedicated resource allocation sub-problem, where the dedicated mode resource allocation algorithm is the same for both network-assisted and user-assisted scenarios.

In particular, the dedicated mode resource allocation sub-problem increases the number of singleton coalitions, such that $\mathcal{D}_{DM} \subset \mathcal{F}$, $\bigcup_{\mathcal{F}_\pi \in \mathcal{D}_{DM}} \mathcal{F}_\pi = \mathcal{R}_{DM}$. Thus, in the CLC game, D2D pairs operating in dedicated mode are fixed in their singleton coalitions for the remainder of the game. This will ensure that all D2D pairs in $\mathcal{D}_{DM}$ will exhaust the benefits of the allocated resource blocks, and thus ensuring network performance is not degraded.

4.3.3.2 Interference Management Sub-Problem

D2D pairs in reuse mode, reuse/share a particular cellular user's resource block. This results in intra-cell interference between D2D pairs and cellular users, and other D2D pairs reusing the same resource block. Therefore, we propose a D2D pair uplink resource scheduling scheme along with a power control sub-game, to minimise this intra-cell interference.

The proposed D2D pair uplink scheduling scheme for D2D pairs operating in reuse mode, extends on the current work on ITLinQ by incorporating cellular user interference and QoS into the scheduling scheme. Thus, the proposed scheme is called revised-information-theoretic link scheduling ($RITLinQ$), which aims to reduce the number of transmitting D2D pairs in reuse mode at any one time, where all scheduled D2D pairs will transmit in the same time slot. Thus, reducing the intra-cell interference within the network, while guaranteeing cellular and D2D user performance.

To determine which D2D pairs are scheduled within each resource sharing coalition, $\mathcal{F}_\pi \setminus \{j\} \in \mathcal{D}_{RM}$, $j \in \mathcal{C}$, the D2D pairs are first randomly prioritised. D2D pair i with the highest priority is always scheduled (active D2D pair), $i \in \mathcal{A}$, $\mathcal{A} \subset \mathcal{F}_\pi$. The remaining inactive D2D links form the set $\mathcal{Q} \subset \mathcal{F}_\pi$, where each D2D pair $q \in \mathcal{Q}$ is selected one at a time in descending order from highest priority, and tests to see if it satisfies a set of conditions in order to be scheduled for transmission. Hence, the set of conditions D2D pair $q \in \mathcal{Q}$ must satisfy, is given as follows:

(i) The interference generated to the cellular user, from all scheduled D2D pairs in $\mathcal{A}$ and the selected D2D pair $q \in \mathcal{Q}$, does not cause the cellular user's QoS to degrade, thus satisfying (4.9).

Algorithm 2 Dedicated mode resource allocation sub-problem

for $i = 1$ to $|\mathcal{D}|$ **do**
  if $\zeta_{i,i} > \overline{\zeta}$ **then**
    BS allocates D2D pair $i \in \mathcal{D}_{CM}$ a orthogonal resource block;
  else
    BS allocates D2D pair $i \in \mathcal{D}_{RM}$ a non-orthogonal resource block;
  end if
end for
BS has R_{DM} idle orthogonal resource blocks for D2D pairs, which form the set $\mathcal{R}_{DM} \subset \mathcal{R}$;
while $\mathcal{R}_{DM} \neq \emptyset$ **do**
  $i^* = \min(\boldsymbol{u}(\mathcal{D}_{RM}))$;
  $R^*_{DM} = \mathcal{D}_{RM}(i^\star)$;
  ▷ Resource block $R^*_{DM} \in \mathcal{R}_{DM}$ has been assigned to D2D pair $i^* \in \mathcal{D}_{RM}$
  $\mathcal{R}_{DM} = \mathcal{R}_{DM} \setminus \{R^*_{DM}\}$;
  $\mathcal{D}_{RM} = \mathcal{D}_{RM} \setminus \{i^*\}$;
  $\mathcal{D}_{DM} = \mathcal{D}_{DM} \cup \{i^*\}$;
end while

(ii) The selected D2D pair $q \in \mathcal{Q}$ must ensure that it does not receive too much interference from the already scheduled D2D pairs in $\mathcal{A}$ and the cellular user, by satisfying:

$$\xi(SNR_{dq})^{\eta} > INR_{dq,di}, \quad \forall i \in \mathcal{A}, i \neq q, \tag{4.22}$$

$$INR_{dq,di} = \frac{p_{cj}|g_{cj,dq}|^2 + \sum_{\substack{i=1 \\ i \neq q}}^{A} p_{di}|g_{di,dq}|^2}{N_{0,d}}, \tag{4.23}$$

$SNR_{dq} = \frac{p_{dq}|g_{dq,dq}|^2}{N_{0,d}}$ is the signal-to-noise ratio of D2D pair q; $INR_{dq,di}$ is the interference-to-noise ratio from D2D pair q to D2D pair i; ξ is a tuning parameter; and η is an exponent.

(iii) The selected D2D pair $q \in \mathcal{Q}$ must also ensure that they do not cause too much interference to already scheduled D2D pairs in $\mathcal{A}$ and the cellular user, by satisfying:

$$\xi(SNR_{dq})^{\eta} > INR_{di,dq}, \quad \forall i \in \mathcal{A}, i \neq q, \tag{4.24}$$

$$INR_{di,dq} = \frac{p_{dq}|g_{dq,BS}|^2 + \sum_{\substack{i=1 \\ i \neq q}}^{A} p_{dq}|g_{dq,di}|^2}{N_{0,d}}. \tag{4.25}$$

Furthermore, if D2D pair $q \in \mathcal{Q}$ is not scheduled in a particular time slot, they will remain inactive (silent) and wait till the next time slot, where D2D pairs are

randomly prioritised again within each coalition.

Meanwhile, the power control sub-game aims to minimise transmit power for all users in coalition $\mathcal{F}_\pi \in \mathcal{F}$, $\forall \mathcal{F}_\pi \in F$, which also reduces intra-cell interference and increases power efficiency and D2D user battery lifetime. The set of players in the power control sub-game, is equal to the set of users within coalition $\mathcal{F}_\pi$. The finite action set for each user i in coalition $\mathcal{F}_\pi$ is transmit power, p_i, which is bounded by a minimum and maximum. We observe that p_i belongs to a nonempty, convex, and compact subset of Euclidean space $\mathbb{R}^{|\mathcal{F}_\pi|}$. The power control sub-game maximises each user's utility function, defined in (4.18), with respect to all other members in the same coalition, in order to solve for optimal transmit power, and is expressed as follows:

$$\begin{aligned} &\underset{p_i \in \mathcal{F}_\pi}{\text{maximise}} \quad u_i(\mathcal{F}_\pi) \quad \forall i \in \mathcal{F}_\pi,\ \forall \mathcal{F}_\pi \in F \\ &\text{subject to} \quad p_{min} \leq p_i \leq p_{max}. \end{aligned} \tag{4.26}$$

Proposition 4.5 The power control sub-game contributes to the stability and convergence of the overall CLC game.

Proof: See Appendix B.3. ■

Thus, the proposed utility function, defined in (4.18), is strictly concave and continuous with respect to transmit power p_i for each user in the proposed CLC game. From (4.26), the best response (optimal) transmit power p_i^* for user i, can be solved by setting the first derivative of the utility function in (4.18) to zero, i.e., $\frac{\partial u_i(\mathcal{F}_\pi)}{\partial p_i} = 0$. The best response transmit power p_i^* for user i, is given by:

$$\begin{aligned} p_i^* =& \frac{-\left(I_i + N_0\right)}{2|g_{i,i'}|^2} \\ &+ \frac{\sqrt{2\lambda_i \omega \left(|g_{i,i'}|^2\right)^2 \ln(2) + \omega^2 \left(I_i + N_0\right)^2 \ln(2)^2}}{2\omega |g_{i,i'}|^2 \ln(2)}, \end{aligned} \tag{4.27}$$

where i is either a D2D pair or cellular user from coalition $\mathcal{F}_\pi$; I_i is the interference that user i receives in coalition $\mathcal{F}_\pi$; and $|g_{i,i'}|^2$ is the channel gain of transmitter i to receiver i'. Hence, the best response transmit power p_i^* for each user i, is searched within $\{p_{\min}, p_i^*, p_{\max}\}$.

4.4 Distributed Cross-Layer Coalitional Algorithm

To solve the proposed CLC game, we propose a fully distributed algorithm to find a stable coalition formation for resource utilisation in D2D communications. The proposed distributed algorithm shown in Algorithm 3 considers the network-assisted scenario, whereas the proposed distributed algorithm in Algorithm 4 considers the user-assisted scenario. Both algorithms, enables D2D pairs to select optimal transmission mode (optimal cluster formation), while utilising resources efficiently and minimising intra-cell interference within the network. During the proposed algorithms, it is assumed that the D2D pairs are established (and the number of D2D pairs in the cell is constant) over a finite time horizon of the game. If any additional D2D pairs are established in the network or any D2D pairs in the network are terminated, then the new set of D2D pairs will be considered in the next occasion when the game is played. Note that, the proposed CLC game can be re-initiated any-time after convergence, with the new set of D2D pairs. Hence, Algorithm 3 and Algorithm 4 outputs the stable coalition partition and the best response transmit power for D2D pairs and cellular users.

4.5 Cross-Layer Coalition Game Stability and Computational Cost Analysis

Next, we analyse the convergence and stability of the coalition partition in the proposed CLC algorithms for the network-assisted and user-assisted D2D communication scenarios. The computational cost of the CLC algorithms is also analysed.

4.5.1 Stability Analysis

The stability of the proposed CLC game is defined by the core of the game and Nash stability. Note that, core stability does not imply that the game is also Nash stable[1]. Nash stability for our proposed CLC game is defined as follows:

Definition 4.9 The coalition partition $\mathcal{F}^* = \{\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_\pi\}$, is *Nash stable* for every player $i \in \mathcal{D}$ if $i \in \mathcal{F}_\pi$, $\mathcal{F}_\pi \in \mathcal{F}^*$, and $\mathcal{F}_\pi \succeq_i \mathcal{F}_{\pi'} \cup \{i\}$ for all other coalitions

[1]Nash Stability is an efficiency measure for coalitional games, which is different to that of Nash Equilibrium, which is a best response condition for non-cooperative games.

Algorithm 3 CLC game: Network-assisted D2D communications

Input: Set of D2D pairs $\mathcal{D}$ and cellular users $\mathcal{C}$, initial formed coalitions: $\mathcal{F}^{t=0} = \{\mathcal{D}_{RM} \cup \mathcal{D}_{CM} \cup \mathcal{C}\}$;
while $t \leq T$ **do**
  if $\mathcal{R}_{DM} \neq \emptyset$ **then**
    Run Algorithm 2. Update $\mathcal{F}^t$, $\mathcal{D}_{DM}$, $\mathcal{D}_{RM}$, and $\mathcal{R}_{DM}$;
  end if
  Run *RITLinQ* protocol (see Sec. 4.3.3.2) for each resource sharing coalition $\mathcal{F}_\pi$, where $\mathcal{F}_\pi \setminus \{j\} \subset \mathcal{D}_{RM}$, $j \in \mathcal{C}$. Update $\mathcal{F}^t$;
  Selected coalition: $\mathcal{F}_\pi \in \mathcal{S}$;
  for $(i = 1$ to $|\mathcal{F}_\pi|)$ & $(i \in \mathcal{D})$ **do**
    ▷ Use leave-and-join Def. 4.6;
    if i satisfies Def. 4.6 **then**
      D2D pair i leaves current coalition $\mathcal{F}_\pi \setminus \{i\}$ and joins coalition $\mathcal{F}_\phi \cup \{i\}$, where $\mathcal{F}_\pi \neq \mathcal{F}_\phi$ and $\mathcal{F}_\phi \in \mathcal{F}^t$;
    else
      D2D pair i stays in current coalition $\mathcal{F}_\pi$;
    end if
  end for
  Run *Power control sub-game* $\forall n \in \mathcal{F}_\pi$, $n \in \mathcal{N}$, $\forall \mathcal{F}_\pi \in F$;
  Update coalition value: $\forall \mathcal{F}_\pi \in \mathcal{F}^t$;
  Update coalition partition: $\mathcal{F}^{t+1} = \mathcal{F}^t$;
  $t = t + 1$;
end while
$p_n^* = p_n(T) \forall n \in \mathcal{N}$;
Output: Coalition partition $\mathcal{F}^T$ and $p_n^* \forall n \in \mathcal{N}$.

$\mathcal{F}_{\pi'} \in \mathcal{F}^*$, $\mathcal{F}^* \cup \{\emptyset\}$, $i \notin \mathcal{F}_{\pi'}$ [72].

Individual stability and contractual individual stability are additional stability conditions [72, 73] which need to be considered when analysing the stability of the coalition partition in the proposed CLC game. Nash stability is the strongest notion of stability for coalition partition, which ensures that no player has incentive to leave a coalition. On the other hand, individual stability states that no player has incentive to move from their current coalition to a new coalition, if all members of the new coalition are negatively impacted [72, 73]. Furthermore, contractual individual stability consists of even tighter bounds, which states that no player has incentive to move from their current coalition to a new coalition, if all members in the current and new coalition are negatively impacted [72, 73].

Remark 4.7 If the coalition partition $\mathcal{F}$ satisfies the concept of Nash stability, then this implies individual stability and contractual individual stability are also satisfied [72].

Algorithm 4 CLC game: User-assisted D2D communications

Input: Set of D2D pairs $\mathcal{D}$ and cellular users $\mathcal{C}$, initial formed coalitions: $\mathcal{F}^{t=0} = \{\mathcal{D}_{RM} \cup \mathcal{D}_{CM} \cup \mathcal{C}\}$;
while $t \leq T$ **do**
  if $\mathcal{R}_{DM} \neq \emptyset$ **then**
    Run Algorithm 2. Update $\mathcal{F}^t$, $\mathcal{D}_{DM}$, $\mathcal{D}_{RM}$, and $\mathcal{R}_{DM}$;
  end if
  Selected coalition: $\mathcal{F}_\pi \in \mathcal{F}^t$;
  if $\mathcal{F}_\pi \in \mathcal{S}$ **then**
    Run *RITLinQ* protocol (see Sec. 4.3.3.2) for resource sharing coalition $\mathcal{F}_\pi$, where $\mathcal{F}_\pi \setminus \{j\} \subset \mathcal{D}_{RM}$, $j \in \mathcal{C}$. Update $\mathcal{F}^t$;
    for ($i = 1$ to $|\mathcal{F}_\pi|$) & ($i \in \mathcal{D}$) **do**
      ▷ Use leave-and-join Def. 4.7 or Def. 4.8;
      if i satisfies Def. 4.7 or Def. 4.8 **then**
        D2D pair i leaves current coalition $\mathcal{F}_\pi \setminus \{i\}$ and joins coalition $\mathcal{F}_\phi \cup \{i\}$, where $\mathcal{F}_\pi \neq \mathcal{F}_\phi$ $\mathcal{F}_\phi \in \mathcal{F}^t$;
      else
        D2D pair i stays in current coalition $\mathcal{F}_\pi$;
      end if
    end for
    Run *Power control sub-game* $\forall k \in \mathcal{F}_\pi$, $k \in \mathcal{N}$;
  else
    Run *Power control sub-game* $\forall k \in \mathcal{F}_\pi$, $k \in \mathcal{N}$;
  end if
  Update coalition value: $\mathcal{F}_\pi \in \mathcal{F}^t$;
  Update coalition partition: $\mathcal{F}^{t+1} = \mathcal{F}^t$.
  $t = t + 1$;
end while
$p_n^* = p_n(T) \forall n \in \mathcal{N}$;
Output: Coalition partition $\mathcal{F}^T$ and $p_n^* \forall n \in \mathcal{N}$.

Theorem 4.1 The coalition partition $\mathcal{F}$ in our proposed CLC game is Nash stable across all cellular and D2D users.

Proof: See Appendix B.4. ■

Remark 4.8 The proof for Nash stability and social efficiency for the coalition partition for the network-assisted or user-assisted D2D communication scenarios, yields the same outcome.

Moreover, the core of a coalitional game, is a coalition partition such that no player wishes to deviate [175]. To prove that the core of the proposed CLC game exists and is non-empty, we define two properties, that is, the common ranking property and the weak top-coalition property, for the network-assisted and user-assisted CLC games, respectively. The common ranking property for our proposed

network-assisted CLC game, is defined as follows:

Definition 4.10 The network-assisted CLC game G_{CLC} satisfies the *common ranking property* iff there exists a Preference-order (as defined in Definition. 4.5), such that for any D2D pair $i \in \mathcal{D}$ and any $\mathcal{F}_\pi, \mathcal{F}_{\pi'} \in \mathcal{F}$, where $\mathcal{F}_\pi \neq \mathcal{F}_{\pi'}$, we have $\mathcal{F}_\pi \succeq_i \mathcal{F}_{\pi'}$ [175].

Remark 4.9 Since the proposed network-assisted CLC game satisfies the common ranking property defined in Definition 4.10, this property guarantees that the core of the network-assisted CLC game G_{CLC} is non-empty.

On the other hand, the weak-top coalition property for our proposed user-assisted CLC game, is defined as follows:

Definition 4.11 Given a non-empty set $\mathcal{F} \subseteq \mathcal{N}$, a non-empty subset $\mathcal{F}_\pi \subseteq \mathcal{F}$ is a *weak top-coalition* of $\mathcal{F}$ iff coalition $\mathcal{F}_\pi$ has an ordered partition such that (i) for any D2D pair $i \in \mathcal{F}_\pi$ and any $\mathcal{F}_{\pi'} \subseteq \mathcal{F}$ with $i \in \mathcal{F}_{\pi'}$, where $i \in \mathcal{D}$, we have $\mathcal{F}_\pi \succeq_i \mathcal{F}_{\pi'}$; and (ii) for any D2D pair $i \in \mathcal{F}_\pi$ and any $\mathcal{F}_{\pi'} \subseteq \mathcal{F}$ with $i \in \mathcal{F}_{\pi'}$, such that $i \in \mathcal{D}$, we have $\mathcal{F}_{\pi'} \succ_i \mathcal{F}_\pi$ where $\mathcal{F}_{\pi'} \cap \mathcal{F}_\pi = \emptyset$. The user-assisted CLC game G_{CLC} satisfies the *weak top-coalition property* iff for a non-empty set of players $\mathcal{F} \subseteq \mathcal{N}$, there exists a weak top-coalition of $\mathcal{F}$ [175].

Remark 4.10 Since the proposed user-assisted CLC game satisfies the weak top-coalition property defined in Definition 4.11, this property ensures that the core of the user-assisted CLC game G_{CLC} is non-empty.

From the proof of Definition 4.9, social efficiency of the game directly implies that the coalition partition is *Pareto efficient* [176]. Pareto efficiency is defined as follows:

Definition 4.12 The coalition partition $\mathcal{F} = \{\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_\pi\}$, is Pareto efficient if there is no partition $\mathcal{F}'$, $\mathcal{F} \neq \mathcal{F}'$, such that $\mathcal{F}'(i) \succeq_i \mathcal{F}(i) \ \ \forall i \in \mathcal{D}$, and for at least one player $m \in \mathcal{D}$, $\mathcal{F}'(m) \succ \mathcal{F}(m)$ [72, 177].

4.5.2 Computational Cost

The computational cost for the proposed CLC game is different for network-assisted and user-assisted D2D communication scenarios. In the network-assisted scenario,

due to evaluating all D2D pairs and cellular users in their coalitions simultaneously, this results in a computational cost of $O(\mathcal{N})$, where $\mathcal{N} = \mathcal{C} \cup \mathcal{D}$. However, the network-assisted scenario will cause an increase in signalling overhead and network complexity, due to the BS needing global knowledge of the network. The computational cost for the game in the user-assisted scenario is $O(R\ \mathcal{N})$, due to evaluating each coalition separately. Hence, the computational cost, for the user-assisted scenario, scales linearly with the number of resource blocks R in the network, when compared to the network-assisted algorithm. However, note that, signalling overhead and network complexity, is different to computational cost for the game. Moreover, there is a trade-off between computational cost, signalling overhead, network complexity, and the BS's knowledge of the network.

4.6 Simulation Results

For our simulations, we consider a single BS located at the center of a circular area with a 500 meter radius. D2D pairs and cellular users are distributed randomly throughout the network. All users encounter a time-selective slow and flat Rayleigh fading channel, with a normalised Doppler spread of 0.002, which, for example, could correspond to a maximal Doppler frequency of 4 Hz with a game iteration occurring at 0.5 ms intervals. A fixed number of resource blocks within the network is considered, where the number of resource blocks maps directly to the number of coalitions in the game. Thus, the total number of available resource blocks considered in the CLC game is $R = 36$. The utility function for each player in (4.18) has the following constants $w = 1$ and $b = 2$. The rest of the simulation parameters are outlined in Table 4.1.

In the simulation results, instead of plotting each transmission mode for D2D pairs (cellular mode, dedicated mode, and reuse mode), we plot the average of all transmission modes combined, which is referred to as D2D pairs in the simulation plots. The proposed CLC game is evaluated using Monte Carlo simulations with varying cellular users C and resource blocks R. We compare our proposed CLC algorithm against a baseline coalitional algorithm, given in [94], and the proposed non-cooperative cross-layer repeated (NCL) algorithm from Section 4.2, i.e., based on our work in [77]. The baseline coalitional algorithm [94] and the NCL algorithm [77], both consider D2D pairs only in cellular mode and reuse mode.

Table 4.1: Simulation Parameters

Parameter	**Value**
Cell radius	500 m
Carrier frequency	2 GHz
Maximum distance between D2D pairs	50 m
D2D pair reference distance	25 m
Cellular user reference distance	250 m
Doppler spread ($f_D T_s$)	0.002
Cellular free space path loss attenuation (A_{PL})	$10^{-4.32}$
D2D free space path loss attenuation (A_{PL})	$10^{-3.32}$
Path loss exponent for cellular users (α)	3.76
Path loss exponent for D2D pairs (α)	4
Number of cellular users per cell (C)	21
Number of D2D pairs per cell (D)	80
Cellular and D2D transmit power range	0-23 dBm
Cellular user threshold SINR ($\overline{\gamma}_c$)	6 dB
D2D pair threshold SINR ($\overline{\gamma}_d$)	3 dB
AWGN power (N_0)	-110 dBm
RITLinQ tuning parameter (ξ)	21
RITLinQ exponent (η)	0.3

Fig. 4.3 shows the network sum power of all users in the network over time. We observe that the proposed network-assisted CLC algorithm and user-assisted CLC algorithm, minimises network sum power on average and converges to an optimal outcome. Due to considering RITLinQ in the CLC algorithm, the number of D2D pairs reusing a cellular user's resource is reduced, thus causing intra-cell interference to reduce. Hence, there is considerable benefit to the individual user performance and overall network performance, when users form cooperative groups and aim to maximise such a common objective. For the network-assisted CLC algorithm, on average, the network sum power is reduced by around 3.8% compared to the user-assisted CLC algorithm, 12.6% compared to the NCL algorithm, and 23.9% compared to the baseline coalition algorithm. Whereas, the user-assisted CLC algorithm reduces network sum power by around 9.2% compared to the NCL algorithm, and 20.9% compared to the baseline coalition algorithm. In addition, the proposed CLC algorithm shows some slight oscillation, due to the number of scheduled D2D pairs in reuse mode changing in each game iteration. Thus, the network-assisted CLC algorithm outperforms the user-assisted CLC algorithm in terms of network sum power, due to D2D pairs in the user-assisted scenario having to select their transmission mode without the assistance of the BS.

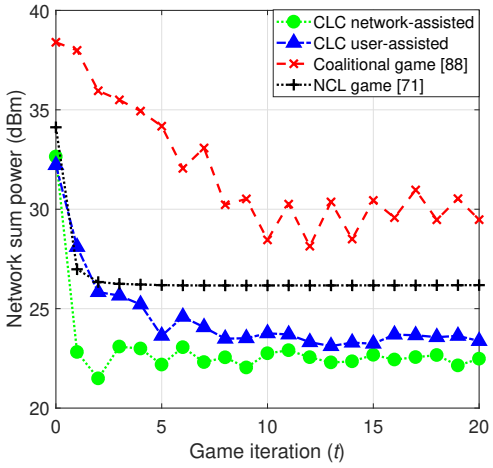


Figure 4.3: Network sum power across all users in the network.

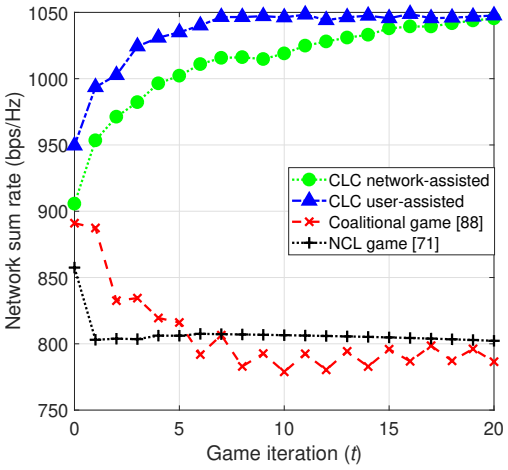


Figure 4.4: Network sum rate across all users in the network.

Fig. 4.4 shows the dynamics of network sum rate for all users in the network. We observe that the proposed network-assisted CLC algorithm and user-assisted CLC algorithm, maximises network sum rate on average and converges to an optimal outcome. As shown in Fig. 4.4, the user-assisted CLC algorithm and network-assisted CLC algorithm maximise the network sum rate, compared to the baseline coalitional algorithm, and the NCL algorithm. Hence, the user-assisted CLC algorithm, on average, increases network sum rate by around 2.1% compared to the network-assisted CLC algorithm, 30.6% compared to the NCL algorithm, and 31.8% compared to the baseline coalition algorithm. On the other hand, the network-assisted CLC algorithm increases network sum rate by around 30.3% compared to the NCL algorithm, and 31.5% compared to the baseline coalition algorithm.

Fig. 4.5 shows the network sum rate across all users, under different numbers of cellular users C, while fixing the number of available resource blocks at, $R = 44$. We observe that as the number of cellular users increases under the current network conditions, there comes a point where increasing the number of cellular users is no longer beneficial, due to reducing the number of available resource blocks for D2D pairs to operate in dedicated mode. For $R = 44$ and $D = 80$, the proposed network-assisted CLC algorithm increases network sum rate by around 3.8% compared to the user-assisted CLC algorithm, by around 39.1% compared to the NCL algorithm, and by around 39.4% compared to the baseline coalition algorithm. The user-assisted CLC algorithm increases network sum rate by around 33.9% compared to the NCL algorithm, and by around 34.3% compared to the baseline coalition

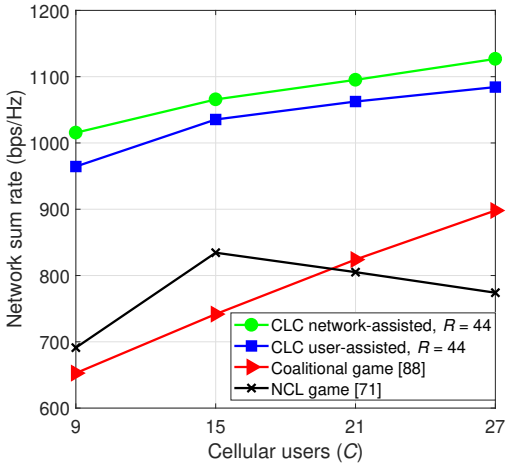


Figure 4.5: Network sum rate as a function of cellular users C.

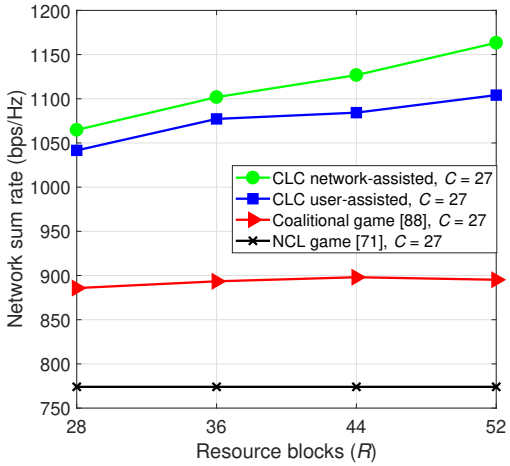


Figure 4.6: Network sum rate as a function of resource blocks R.

algorithm.

Moreover, in Fig. 4.6, we show the network sum rate across all users, under different numbers of resource blocks R, while fixing the number of cellular users at $C = 27$ and D2D pairs at $D = 80$. Increasing the number of resource blocks within the network directly maps to increasing the number of coalitions within the proposed CLC game. As the number of coalitions within the proposed CLC game is increased, more dedicated resource blocks become available for D2D pairs to operate in dedicated mode, which, in turn, reduces the number of D2D pairs reusing cellular user's resources (spreading the load more evenly across the coalitions). The proposed network-assisted CLC algorithm increases network sum rate by around 3.4% compared to the user-assisted CLC algorithm, by around 24.7% compared to the baseline coalition algorithm, and by around 43.9% compared to the NCL algorithm. Whereas, the user-assisted CLC algorithm increases network sum rate by around 20.6% compared to the baseline coalition algorithm, and by around 39.1% compared to the NCL algorithm.

Fig. 4.7 shows the number of active (scheduled) D2D pairs in reuse mode, under different numbers of cellular users C. We observe that as the number of cellular users increase, the number of active D2D pairs in reuse mode without RITLinQ increases at a constant rate, while the number of active D2D pairs in reuse mode with RITLinQ increases at a slightly faster rate. Due to increasing the number of cellular users within the network, the D2D pairs in reuse mode have more cellular user resource block's to reuse, thus the number of D2D pairs reusing

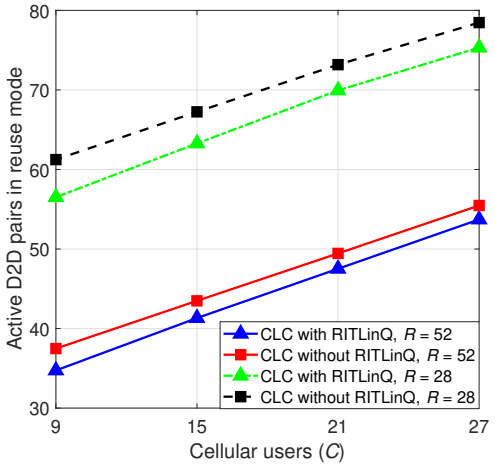


Figure 4.7: Number of active D2D pairs in reuse mode as a function of cellular users C.

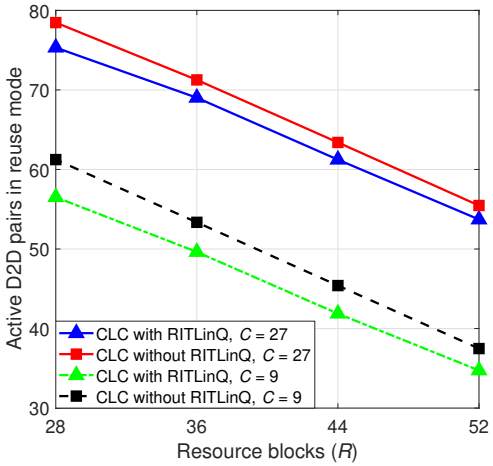


Figure 4.8: Number of active D2D pairs in reuse mode as a function of resource blocks R.

a particular cellular user's resource block is reduced, which, in turn, reduces the intra-cell interference within the network. Furthermore, as the number of resource blocks increase from $R = 28$ to $R = 52$, we observe that the number of D2D pairs in reuse mode decreases and thus the number of active D2D pairs also decrease, due to more resource blocks becoming available for D2D pairs to operate in dedicated mode. On average, the proposed CLC algorithm with RITLinQ reduces the number of active D2D pairs in reuse mode at a given time by around 5.5% for $R = 28$, compared to the CLC algorithm without RITLinQ. While the CLC algorithm with RITLinQ reduces the number of active D2D pairs in reuse mode at a given time by around 4.8% for $R = 52$, compared to the CLC algorithm without RITLinQ.

Fig. 4.8 shows the number of active (scheduled) D2D pairs in reuse mode, under different numbers of resource blocks R. We observe that as the number of resource blocks increase, the number of active D2D pairs in reuse mode without RITLinQ decreases at a constant rate, while the number of active D2D pairs in reuse mode with RITLinQ decreases at a slightly slower rate. Since dedicated mode was considered in the D2D mode selection, as the number resource blocks within the network increases, the number of D2D pairs switching from reuse to dedicated mode will also increase, thus, reducing intra-cell interference and the number reusing D2D pairs. Additionally, as the number of cellular users increase from $C = 9$ to $C = 27$, we observe that the number of D2D pairs in reuse mode increase, due to reducing the number of free resource blocks for D2D pairs to

operate in dedicated mode. Thus, on average, the proposed CLC algorithm with RITLinQ reduces the number of active D2D pairs in reuse mode at a given time by around 7.45% for $C = 9$, compared to the CLC algorithm without RITLinQ. Whereas, the CLC algorithm with RITLinQ reduces the number of active D2D pairs in reuse mode at a given time by around 3.4% for $C = 27$, compared to the CLC algorithm without RITLinQ. Moreover, in Fig. 4.7 and Fig. 4.8, the CLC algorithm with and without RITLinQ is very efficient and consistent with respect to the number of active D2D pairs in reuse mode and varying resources within the network. The proposed CLC algorithm with RITLinQ provides further benefit to the network, by allocating and utilising resources more efficiently.

4.7 Summary of Contributions

In this chapter, we jointly optimised D2D mode selection, resource allocation, and interference management in a fully distributed and autonomous manner. In addition, we also investigated two network scenarios, i.e., a network-assisted scenario and a user-assisted scenario. Hence, in this chapter we answered questions Q3, Q4, and Q5, posed in Section 1.2.

Firstly, we proposed a network-assisted non-cooperative mode selection and power control game to solve this joint optimisation, by modelling the problem using a non-cooperative power control game and a two-armed bandit game. Analysis, of the proposed non-cooperative game, demonstrated convergence to a unique Nash Equilibrium. Then, we extended the non-cooperative cross-layer repeated game to a cross-layer coalitional game with a non-transferable utility, where D2D pairs partitioned themselves into coalitions in a fully distributed manner, based on mode selection and resource allocation. In order to further reduce intra-cell interference within the network, our proposed coalitional game also incorporated scheduling D2D pairs reusing cellular user resources, along with power control for all users in the network. Hence, the proposed cross-layer coalitional game based algorithm, was shown to converge to a Nash stable coalition partition, which is also socially efficient and has a non-empty core. Furthermore, the proposed network-assisted coalitional game was extended and modelled using a user-assisted approach, where the BS only has partial knowledge of the network. The user-assisted coalitional game based algorithm was also shown to converge to a Nash stable outcome.

Overall, the proposed coalitional game (for either network-assisted or user-assisted scenarios) was the most suited approach for D2D mode selection, compared to the non-cooperative game approach. Indeed, the coalitional game enabled D2D users to dynamically switch and continuously change their transmission mode, whereas in the proposed non-cooperative game, once a mode selection change had occurred for a D2D user then they would no longer be able to switch transmission modes again. This is due to the non-cooperative and pure strategy nature of the proposed non-cooperative cross-layer game. In addition, the coalitional mode selection approach also took into consideration the effects of other players when a D2D pair has sought to switch to a different transmission mode, such that if one user was to be negatively impacted due to this change, then a D2D pair was not permitted to change their transmission mode. Thus, this improved overall network performance and increased throughput for the coalitional game approach.

Simulation results showed that, on average, the proposed cross-layer coalitional algorithm minimises network sum power, whilst maximising network sum rate, compared to the proposed non-cooperative cross-layer repeated game. In addition, our simulation results showed that, by increasing the number of cellular users or resource blocks within the network, there is an increase in network sum rate, in conjunction with allocating and utilising resources more efficiently within the network. Thus, the proposed coalition game based algorithm significantly increased system capacity and throughput, compared to the non-cooperative cross-layer game based algorithm.

Chapter 5

Correlation-Aware Clustering in M2M Networks

Chapters 3 and 4 have focused on resource management and end-user satisfaction for D2D communications in current and future cellular networks, where such D2D communications enable users who are within close proximity to transmit directly between one another, without relaying through a BS. Thus, D2D communications is an integral part of alleviating network congestion, traffic, and demand on the BS. Importantly, the concept of D2D communications can be extended from CTDs to MTDs, which is referred to as M2M communications, which is a fundamental implementation domain for emerging IoT systems. Such M2M communications enables MTDs to share data via a direct link to one another.

In this chapter, we consider the problem of self-organising, correlation-aware clustering for a massive number of MTDs densely deployed in a cellular network. In such dense networks, MTDs sense an environment and transmit their data to the BS via a cellular uplink. However, MTDs are typically located within close proximity and will gather correlated data, and, thus, large amounts of redundant bits can be transmitted to the BS. In this chapter, we formulate this problem as a dynamic evolutionary game, where data correlation and transmission power are modelled as dynamics. In the formulated game, MTDs can self-organise in a distributed manner to form clusters, based on data correlation. To derive the utility function for the proposed evolutionary game model, we use stochastic geometry to accurately model and characterize the distance distributions between MTDs which, in turn, allows the derivation of a closed-form upper bound (worst-case) expression for the inter-cluster interference. The utility function per cluster, is defined as a function of MTD distance distributions, inter-cluster interference, and cluster size.

The proposed evolutionary game based algorithm converges to a stable cluster formation, which is robust to a small portion of MTDs changing their strategy at the stable outcome, due to stochastic changes in the M2M network environment, such as the joining of new MTDs, MTD loss of battery, or rapid fluctuations in the sensing environment. In this regard, we also derive the maximum portion of MTDs that can deviate from the stable cluster formation, while still maintaining a stable cluster formation. Hence, simulation results show that the proposed approach can effectively cluster MTDs with highly correlated data which, in turn, enables those MTDs to eliminate a large number of redundant bits.

This chapter is organised as follows. In Section 5.1, we define the system model and problem formulation for correlation-aware clustering of MTDs. In Section 5.2, an evolutionary game based on data correlation and transmission energy is proposed to cluster a massive number of MTDs, and in the system design for this game stochastic geometry is used to characterize distances and inter-cluster interference. Section 5.3 outlines the proposed algorithm. Within Section 5.4, the proposed evolutionary game stability is analysed. Simulation results are illustrated in Section 5.5. Finally, Section 5.6 concludes the chapter.

5.1 System Model

Consider the uplink of a wireless cellular network having a single BS serving a massive number of MTDs engaged in M2M communications. All MTDs in the set $\mathcal{M}$ are randomly deployed in a 2D space $\mathbb{R}^2$, where the location of each MTD m is denoted by $\boldsymbol{y}_m$, $\forall m \in \mathcal{M}$, and is based on a Poisson point process (PPP) $\Phi_M = \{\boldsymbol{y}_m \in \mathbb{R}^2 | m \in \mathcal{M}\}$ with density λ_m. Since in practical M2M networks, two or more MTDs cannot be placed at the same given point in the network, where the number of MTDs are placed within a finite area $A < \infty$ [178]. Hence, the PPP model for the massive number of MTDs is said to be *locally finite*, that is, $\mathcal{M}_f$ is the set of locally finite MTDs where $\mathcal{M}_f \subset \mathcal{M}$ and $|\mathcal{M}_f| < \infty$ [178]. We assume that MTDs gather data from a Gaussian random field, as this model represents an upper bound for the number of bits needed to encode a source field [44, 152, 179]. Considering the Gaussian distribution will result in maximum entropy across all distributions, which means that MTDs will transmit with maximum packet size [179]. Thus, the data source s_m for each MTD m, is a Gaussian random variable with mean

μ_m and variance σ_m^2 [44]. Here, we use entropy to model the information of each MTD's data. Let each MTD quantize its continuous Gaussian data source with a sufficiently small quantization step Δ. We use the entropy H_m to measure the number of bits each MTD's quantized data source [44]:

$$H_m = \frac{1}{2}\log_2\left(\frac{2\pi e}{\Delta^2}\sigma_m^2\right). \tag{5.1}$$

In this network, each MTD m sends its data via a cellular link to the BS. We use a given resource allocation mechanism for the cellular links. The BS allocates orthogonal resource blocks to each MTD for the cellular link, based on an orthogonal frequency-division multiple access (OFDMA) system. A finite set $\mathcal{Z}$ of Z resource blocks are used for cellular link transmission, with fixed bandwidth B Hz per resource block, during each time slot t with a fixed duration T. Each MTD m is allocated one resource block $z \in \mathcal{Z}$ and a transmit power $p_m \in [0, P_{\max}]$ for cellular link transmission. The transmission power per resource block $z \in \mathcal{Z}$ that MTD m requires to send H_m bits over the cellular link during each time slot is:

$$p_m(t) = \left(\frac{BN_0}{g_m(t)^{(z)}}\right)\left(2^{\frac{H_m}{TB}} - 1\right), \tag{5.2}$$

where $g_m(t)^{(z)}$ is the channel gain of MTD m over the cellular link on resource block z; and N_0 is the noise power spectral density.

In the uplink of the cellular networks, the number of available orthogonal resource blocks have been designed to suit the needs of CTDs. Thus, the number of available resource blocks for each MTD is limited, and as a result, not all MTDs will be allocated orthogonal resources to transmit sensing data to the BS, which may incur data/information to be lost. However, due to the dense deployment of MTDs, MTDs within close proximity will gather correlated data [44]. For example, sensors in the same room usually record similar measurements, and thus, the data correlation is high. Hence, multiple MTDs will send a large number of redundant bits (same information) to the BS. Next, we introduce a *correlation-aware clustering scheme* that will enable a massive number of MTDs to cooperatively form clusters by sharing their data via M2M links. Particularly, clustering allows a reduction in the number of MTDs accessing the cellular uplink, thus saving orthogonal resources in the uplink for CTDs. Here, we note that, since massive M2M networks are locally finite, then the average number of MTDs per cluster is also

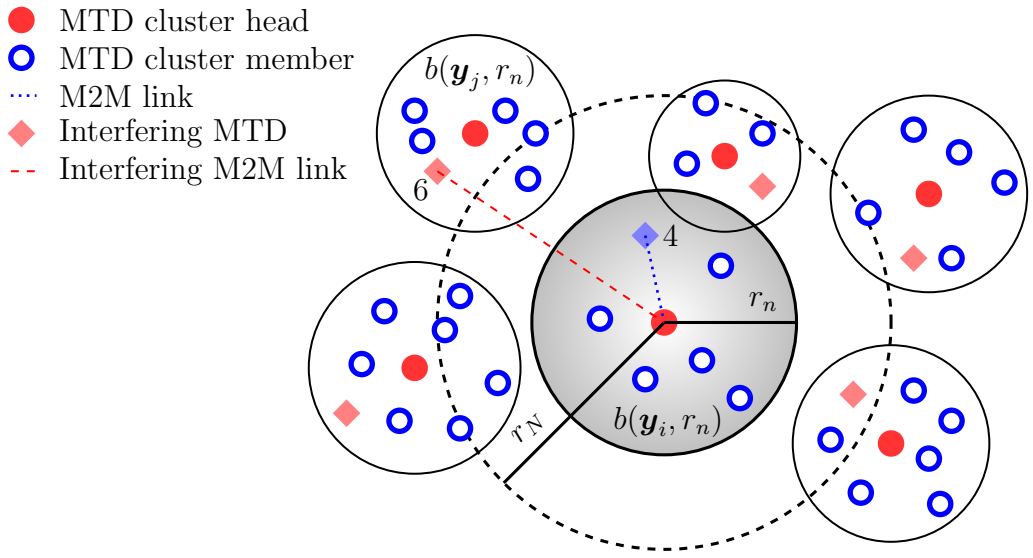


Figure 5.1: Illustration example of clustering in an M2M network.

finite [180, 181].

5.1.1 Problem Formulation

The goal of correlation-aware clustering is to decrease the number of redundant bits sent to the BS, so as to decrease transmission power, increase energy-efficiency and prolong MTD battery lifetime. Within each cluster, an MTD is designated as a *cluster head* and all cluster members send their data via an M2M link to the cluster head. The cluster head relays the collected data to the BS over the cellular link. Before transmission, the cluster head coordinates and eliminates the redundant bits of the shared data (i.e., perform data aggregation), in order to reduce the size of the transmitted packets [45, 182, 183]. We assume the cluster head is selected by an existing scheduling algorithm, e.g., [140, 184], or [185].

Since the massive number of MTDs are located in a finite area, the number of MTDs within this area are locally bounded, and the BS can allocate orthogonal resource blocks to each cluster head for the cellular link transmission. Thus, to improve network efficiency and reduce the number of allocated orthogonal resource blocks for each cluster, the BS can allocate the same set of resource blocks for M2M communications. Note that, as we focus on the clustering process, we assume that the resource allocation within each cluster is given and that the cellular and M2M transmission use orthogonal resources. In order to ensure that sufficient orthogonal resources can be allocated to each cluster head when the number of

clusters in the network is large, our system model incorporates the fast uplink grant scheme discussed in [186, 187], and [188]. Hence, within each cluster, the resource blocks are allocated equally and orthogonally to all cluster members, excluding the cluster head (e.g., see [18, 147] and [150]). Consequently, we only need to deal with inter-cluster interference caused by MTDs in different clusters simultaneously reusing the same resource block. Furthermore, this resource block allocation is suitable for M2M links since they typically consist of low power and small data transmission [18].

For a reference cluster head, e.g., MTD i located at y_i, is equipped with omni-directional antennas, i.e., we assume an MTD can transmit in any direction. Hence, MTD i has an M2M link coverage area of radius r_n. We model the area around MTD i as a circular ball (sphere) $b(\boldsymbol{y}_i, r_n)$, which forms a cluster $\mathcal{K}_{i,n}$ of n MTDs on average, where $n = \lfloor \lambda_m \pi r_n^2 \rfloor$ and λ_m is the network density. Note that, MTDs selected as a cluster head will follow an independent thinning process of the MTD PPP, which will generate another PPP based on a smaller density than compared to the network density [52, 178]. Those MTDs within the ball (i.e., cluster members) are independently and uniformly distributed. The maximum average number of MTDs in a cluster is N, where $N = \lfloor \lambda_m \pi r_N^2 \rfloor$, and r_N is the maximum distance from the reference cluster head i, such that the communication between the MTDs and the cluster head is reliable.

A typical cluster, determines a bounded threshold region for short range communications, which is defined as follows:

Definition 5.1 A *typical cluster* $\mathcal{K}_{i,n}$ is defined as a disjoint cluster centered around reference cluster head i located at $\boldsymbol{y}_i$ (the same point as the ball) and consists of n MTDs on average. The typical cluster $\mathcal{K}_{i,n}$ is a subset of the set of MTDs, $\mathcal{K}_{i,n} \subseteq \mathcal{M}$, that covers a bounded region with radius r_n, and an average number of $n = \lfloor \lambda_m \pi r_n^2 \rfloor$ MTDs.

Based on Definition 5.1, the average number of MTDs within the typical cluster $\mathcal{K}_{i,n}$ is a function of the cluster radius r_n and the location of cluster head $\boldsymbol{y}_i$. We assume that the MTDs form a collection of massive disjoint clusters $\mathfrak{K}$, where no other cluster overlaps with any other cluster in the network. A cluster $\mathcal{K}_{i,n} \in \mathfrak{K}$ is a subset of $\mathcal{M}$, where $\bigcup_{\mathcal{K}_{i,n} \in \mathfrak{K}} \mathcal{K}_{i,n} = \mathcal{M}$ and $\mathcal{K}_{i,n} \cap \mathcal{K}_{j,n} = \emptyset \ \forall i \neq j$. The MTDs within each cluster can be marked, which essentially links the MTD to a certain cluster. Therefore, the MTD locations within each cluster can be modeled

Table 5.1: The maximum average number of MTDs for different densities and radii.

Density (λ_m)	**Radius** (r_N)	$N = \lfloor \lambda_m \pi r_N^2 \rfloor$
0.36	5 m	28
0.27	7 m	41
0.18	11 m	68
0.09	18 m	91

by the marked PPP [178]. Fig. 5.1 represents an illustrative example of a ball and a typical cluster, in the M2M network. Within the ball, we consider a typical cluster $\mathcal{K}_{i,n} \in \mathfrak{K}$ with average cluster size n, and a reference cluster head located at the center of the ball and the typical cluster. Thus, the index i of the typical cluster $\mathcal{K}_{i,n}$ directly relates to the average cluster size n, which also denotes the average number of MTDs within the cluster (including the reference cluster head). As shown in Fig. 5.1, MTDs outside of the ball will cause inter-cluster interference to the reference cluster head i over the M2M link. As illustrated in Fig. 5.1, we consider an example scenario, where cluster $\mathcal{K}_{i,n}$ has a radius r_n and a maximum cluster radius of r_N. In cluster $\mathcal{K}_{i,n}$ and cluster $\mathcal{K}_{j,n}$, MTDs 4 and 6 respectively, share the same allocated resource block and are transmitting to their respective cluster head simultaneously. Thus, MTD 6 from cluster $\mathcal{K}_{j,n}$ will cause inter-cluster interference to the M2M link from MTD 4 to cluster head i in cluster $\mathcal{K}_{i,n}$. Furthermore, since the location of MTDs are based on the realization of the PPP, we outline the average number of MTDs for a given network density λ_m and radius r_N, using $N = \lfloor \lambda_m \pi r_N^2 \rfloor$. If we consider a circular area with a 2000 m radius and a network density $\lambda_m = 0.09$, then the average number of MTDs within the network would be $N = 1.13 \times 10^6$. Meanwhile, if the network density is $\lambda_m = 0.36$ then the average number of MTDs within the network would be $N = 4.52 \times 10^6$. Table 5.1 outlines examples for the actual number of MTDs within a cluster, given the radius r_N and network density λ_m.

We assume that, within any typical cluster $\mathcal{K}_{i,n} \in \mathfrak{K}$ during time slot T, each MTD m must be able to send H_m bits, as given by (5.1), to the reference cluster head i. Therefore, the transmission rate R_{mi} of the M2M link between each cluster member $m \in \mathcal{K}_{i,n} \setminus \{i\}$ and cluster head i, must be greater than or equal to the threshold bit rate $\frac{H_m}{T}$, that is, $\forall m \in \mathcal{K}_{i,n} \setminus \{i\} : \; R_{mi} \geq \frac{H_m}{T}$. Based on the Gaussian random field model, the collection of data streams generated by n MTDs, $\boldsymbol{s}_i = [s_m]_{n \times 1}$, from all of the MTDs in cluster $\mathcal{K}_{i,n}$ will follow a multivariate Gaussian distribution [44]. To model the joint information of data within

Table 5.2: Summary of the notations used throughout this work.

Notation	Description
Φ_M	PPP modelling the locations of MTDs.
λ_m	Density of ϕ_M.
Φ_{I_n}	PPP modelling the locations of interfering MTDs.
λ_{In}	Density of ϕ_{I_n}.
$g_{mi}(t)$	Channel gain power between MTD m and MTD i at time slot t.
ν	Path loss exponent for all M2M and cellular links, where $\nu > 2$.
$\mathcal{K}_{i,n}$	Typical cluster with radius r_n and MTD i as the cluster head; subset of MTDs $\mathcal{K}_{i,n} \subset \mathcal{M}$.
n	Average size of a typical cluster $\mathcal{K}_{i,n}$, where $n = \lfloor \lambda_m \pi r_n^2 \rfloor$.
N	Maximum average number of MTDs per cluster.
r_N	Maximum distance from the reference cluster head i, i.e., the maximum cluster radius.
p_m; q_m	Transmit power via the: cellular link, $p_m \in [0, P_{\max}]$, and the M2M link, $q_m \in [0, Q_{\max}]$.
H_m	Individual joint entropy of MTD m.
$H_{\mathcal{K}_{i,n}}$	Joint entropy of cluster $\mathcal{K}_{i,n}$.

each cluster, we use the notation of a joint entropy derived in [189]. Since the number of MTDs considered in our system model is massive and follows a PPP, the location of MTDs are random and deriving the exact distance between the cluster members and the cluster head can be computationally expensive. In order to simplify our model, we assume the worst-case scenario of joint information in each cluster, i.e., the maximum transmission distance r_n is considered to derive the maximum number of transmission bits (joint entropy) $H_{\mathcal{K}_{i,n}}$ for cluster $\mathcal{K}_{i,n}$ is given by [189]:

$$H_{\mathcal{K}_{i,n}} = H_m + H_m(n-1)\left(1 - \frac{\alpha}{\frac{r_n}{c} + 1}\right), \tag{5.3}$$

where $\alpha = \frac{\log_2(e)}{\log_2(2\pi e)}$ and c is a correlation constant that describes the statistical and spatial relationship between the MTDs. The correlation constant c characterizes the spatial correlation in the data [189], i.e., the correlation constant normalises the maximum radius of the cluster, r_n.

The M2M link transmission power $q_{mi} \in [0, Q_{\max}]$ from MTD m to MTD i in cluster $\mathcal{K}_{i,n}$, depends on the inter-cluster interference, number of transmission bits H_m, and channel gain. Thus, the transmission power q_{mi} that is used to send H_m bits from MTD $m \in \mathcal{K}_{i,n}$ to the reference cluster head $i \in \mathcal{K}_{i,n}$, is given by:

$$q_{mi}(t) = \left(\frac{I_{-i}(t)}{g_{mi}(t)}\right)\left(2^{\frac{H_m}{TB}} - 1\right), \tag{5.4}$$

where $I_{-i}(t)$ is the inter-cluster interference from MTDs in other clusters $\mathcal{K}_{-i} = \{\mathcal{K}_j | \mathcal{K}_j \in \mathfrak{K}, i \neq j\}$ that are transmitting concurrently to their respective cluster heads over the same resource blocks. The channel gain between MTD m and MTD i is given by $g_{mi}(t) = \xi(t)A_t\left(\frac{4\pi d_{mi}}{\mu}\right)^{-\nu}$, where ν is the path loss exponent, d_{mi} is the distance between MTD m and i, μ is the wavelength of an electromagnetic wave, $\xi(t)$ is the time-varying fading channel attenuation, and A_t is the antenna gain of the transmitter and receiver. For simplicity, we ignore the thermal noise based on the assumption that it is negligible compared to the interference. For a quick reference, the notation used in this paper is listed in Table 5.2.

5.1.1.1 M2M Link Distance

We derive an expression for M2M link distance within the typical cluster $\mathcal{K}_{i,n}$ and the reference cluster head i, by substituting the channel gain expression, $g_{mi}(t)$ into (5.4), and rearranging for d_{mi}, as follows:

$$\begin{aligned} q_{mi}(t) &= \left(\frac{I_{-i}(t)}{\xi(t)A_t\left(\frac{4\pi d_{mi}}{\mu}\right)^{-\nu}}\right)\left(2^{\frac{H_m}{TB}} - 1\right), \\ \therefore d_{mi} &= \frac{\mu\left(q_{mi}(t)\xi(t)A_t\right)^{\frac{1}{\nu}}}{4\pi\left(\left(2^{\frac{H_m}{TB}} - 1\right)I_{-i}(t)\right)^{\frac{1}{\nu}}}. \end{aligned} \tag{5.5}$$

Given (5.5) and assuming maximum M2M link transmission power is used $q_{mi}(t) = Q_{\max}$ for all MTDs within cluster $\mathcal{K}_{i,n}$, the maximum M2M link distance is:

$$D_{mi} = \frac{\mu\left(Q_{\max}\xi(t)A_t\right)^{\frac{1}{\nu}}}{4\pi\left(\left(2^{\frac{H_m}{TB}} - 1\right)I_{-i}(t)\right)^{\frac{1}{\nu}}}, \tag{5.6}$$

where the maximum radius for reference cluster head i is $r_N = D_{mi}$.

Since the number of MTDs within the network is massive, instead of obtaining exact locations of all MTDs, we characterize the data correlation within each cluster based on the distance distributions of MTDs. Also, as the distance between MTDs decreases, the covariance between MTD data will increase, which in turn, causes

the joint entropy of the cluster to decrease.

Based on the stochastic geometry [178, 180] and PPP distribution assumption, we consider a k-closest MTD approach, where MTD $m \in \mathcal{K}_{i,n} \setminus \{i\}$ is the k^{th} closest MTD to cluster head i. For doing so, we sort the distances ("order" the distances) from cluster head i to all MTDs in cluster $\mathcal{K}_{i,n}$, in ascending order. For cluster $\mathcal{K}_{i,n}$, we define an *ordered set* of distances $\mathcal{D}_i = \{d_{(k)i}\}_{k=1:n-1}$, such that $d_{(1)i} \leq d_{(2)i} \leq \ldots \leq d_{(k)i} \leq \ldots \leq d_{(n-1)i}$. Since the random variables of the sequence $\mathcal{D}_i$ are independent and identically distributed (i.i.d.), the probability distribution function of the distance from the k^{th} closest MTD to cluster head i, will be given by [190]:

$$f_{\mathcal{D}_i}^{(k)}(d) = \frac{(n-1)!}{(k-1)!(n-1-k)!} F_{\mathcal{D}_i}(d)^{k-1} f_{\mathcal{D}_i}(d) \times \left(1 - F_{\mathcal{D}_i}(d)\right)^{n-1-k}, \tag{5.7}$$

where $d = d_{(k)i}$, $f_{\mathcal{D}_i}(d_{(k)i}) = \frac{2d_{(k)i}}{r_n^2}$ for $0 \leq d_{(k),i} \leq r_n$, and $F_{\mathcal{D}_i}(d_{(k)i}) = \frac{d_{(k)i}^2}{r_n^2}$ for $0 \leq d_{(k),i} \leq r_n$.

5.1.1.2 Total Transmission Power per Cluster

The total transmission power of a typical cluster $\mathcal{K}_{i,n}$ can be defined as the summation of transmission powers over all M2M links and the cellular link within cluster $\mathcal{K}_{i,n}$. Following (5.2) and (5.4), the total transmission power $P_{\mathcal{K}_{i,n}}(t)$ of cluster $\mathcal{K}_{i,n}$ is given by:

$$\begin{aligned} P_{\mathcal{K}_{i,n}}(t) &= \sum_{m \in \mathcal{K}_{i,n} \setminus \{i\}} q_{mi}(t) + p_i(t), \\ &= \sum_{m \in \mathcal{K}_{i,n} \setminus \{i\}} \left(\frac{I_{-i}(t)}{g_{mi}(t)}\right) \left(2^{\frac{H_m}{TB}} - 1\right) + \\ &\quad \left(\frac{BN_0}{g_i(t)}\right) \left(2^{\frac{H_{\mathcal{K}_{i,n}}}{TB}} - 1\right). \end{aligned} \tag{5.8}$$

Equation (5.8) shows how the correlation-aware clustering scheme works. As the correlation between MTD data within the cluster decreases, the total transmission power of the typical cluster $\mathcal{K}_{i,n}$ will increase. This is because the cluster head has to send more bits via the cellular link. From (5.8), with decreasing joint entropy, the cellular link transmit power will also decrease.

However, finding an optimal and centralized cluster formation for a massive number of MTDs in a densely deployed M2M network is not realistic due to the high complexity and signalling overheads. In addition, the signalling overhead can further increase due to the dynamics of the MTD network that can stem from various factors, such as the joining of new MTDs, MTD loss of battery, or rapid fluctuations in the sensing environment.

Next, we propose a distributed correlation-aware MTD clustering framework based on *evolutionary game theory* [191–193] for a massive number of MTDs. Accurate modelling of the M2M network topology becomes a key step towards meaningful performance analysis of correlation-aware clustering. To perform such modelling, prior to formulating the game, in Section 5.1.1.1 we characterized the distance distributions between MTDs, then we will use stochastic geometry tools [178, 180] to determine inter-cluster interference. Based on these system parameters, we propose an evolutionary game model that can effectively capture the dynamics pertaining to a massive number of possible MTD cluster formations. At the convergence point, an evolutionary game is robust to potential deviations of MTDs that can result from factors such as, MTDs leaving/entering the network, or rapid fluctuations in the sensing environment. The proposed model based on stochastic geometry and evolutionary game theory, can help network designers predict how design parameters, such as the density of MTDs, transmission power, resource block bandwidth, and duration time slot, will affect the evolutionary stable clusters of MTDs.

5.2 Correlation-Aware Evolutionary Game in a Massive M2M network

In this section, we propose an evolutionary game [61, 71, 191–195] for clustering a massive and locally finite number of MTDs, in order to reduce transmission power for each MTD as well as the number of redundant bits sent to the BS. This approach is distributed, as it allows a massive number of MTDs to self-organise into clusters based data correlation. In order to define a tractable utility expression for this evolutionary game, we use stochastic geometry to determine the distribution of the inter-cluster interference to all cluster members, excluding the reference cluster head.

We consider an upper bound for inter-cluster interference[1] for all MTDs in cluster $\mathcal{K}_{i,n}$. As explained before, inter-cluster interference is caused by when an MTD in $\mathcal{K}_{-i} = \{\mathcal{K}_j | \mathcal{K} \in \mathfrak{K}, i \neq j\}$ simultaneously reusing the same resource blocks. The distance between an interfering MTD in $\mathcal{K}_{-i}$ and cluster head $i \in \mathcal{K}_{i,n}$ is greater than the cluster radius, r_n. We consider the worst case interference by assuming at least one interfering MTD is located in each cluster outside of $\mathcal{K}_{i,n}$, i.e., the number of interfering MTDs is equal to $\mathfrak{K} \setminus \mathcal{K}_{i,n}$. Then, based on the distance distributions between MTDs, the mean inter-cluster interference for a typical cluster head, is given as follows [178]:

$$\begin{aligned}
\mathbb{E}(I_n) &= \mathbf{E}\left(\sum_{\boldsymbol{y}\in\Phi_{I_n}} qA_t\xi(t)\left(\frac{4\pi}{\mu}\right)^{-\nu}\|\boldsymbol{y}\|^{-\nu}\right), \\
&= qA_t\xi(t)\left(\frac{4\pi}{\mu}\right)^{-\nu}\int_{\mathbb{R}^2\setminus b(\boldsymbol{y}_i,r_n)}\lambda_{I_n}\|\boldsymbol{x}\|^{-\nu}dx, \\
\therefore \mathbb{E}(I_n) &= qA_t\xi(t)\left(\frac{4\pi}{\mu}\right)^{-\nu}\left(\frac{1}{\pi r_n^2}\right)\int_{r_n}^{\infty}\int_0^{2\pi} r^{-\nu} r d\theta dr, \qquad (5.9)
\end{aligned}$$

where interfering MTDs follow a PPP Φ_{I_n} with density $\lambda_{I_n} = \frac{1}{\pi r_n^2}$, r_n is the radius of a typical cluster, and $\nu > 2$. Solving the integration in (5.9) for radius r_n, we have [178]:

$$\mathbb{E}(I_n) = 2qA_t\xi(t)\left(\frac{4\pi}{\mu}\right)^{-\nu}\left(\frac{1}{r_n^{\nu}(\nu-2)}\right). \qquad (5.10)$$

Based on (5.10), as the network density increases, this causes the inter-cluster interference to also increase due to the increasing number of MTDs in the network. Additionally, as the radius of the typical cluster r_n increases, the number of interfering MTDs is reduced and so is the inter-cluster interference.

5.2.1 Evolutionary Game

We define the cluster type in the evolutionary game as follows:

Definition 5.2 A *cluster type* $j \in \mathcal{S}$ represents the average number of MTDs within a cluster, that is, an average cluster of size j.

Thus, the proposed evolutionary game can be formally defined as follows:

[1]Here, we note that, we consider the worst case inter-cluster interference.

Definition 5.3 An *evolutionary game* is defined by, $G_E = (\mathcal{M}, \mathcal{S}, \boldsymbol{x}, u)$, where $\mathcal{M}$ is a massive set of MTDs (population), $\mathcal{S} = \{1, \ldots, N\}$ is the finite set of cluster types (pure strategy set), $\boldsymbol{x}$ is the population state, and u is the utility function of a cluster.

In our proposed evolutionary game, each cluster belongs to a finite set of cluster types $\mathcal{S} = \{1, 2, \ldots, N\}$, where a cluster of type 1 is considered to be a singleton cluster. Here, $N = \lfloor \lambda_m \pi r_N^2 \rfloor$ is the maximum average cluster size for any cluster head within the ball, with r_N being the maximum M2M range which is equal to D_{mi} in (5.6).

The population state vector $\boldsymbol{x} \in \mathbb{R}^N$, where $\sum_{j \in \mathcal{S}} x_j = 1$, captures the percentage of MTDs forming different average cluster sizes. Thus, each element x_j of $\boldsymbol{x}$ represents the average percentage of MTDs forming a cluster size j. The utility achieved by a given cluster $\mathcal{K}_{i,j}$ at time slot t is denoted as $u_{\mathcal{K}_{i,j}}(t)$. Note that, within the ball $b(\boldsymbol{y}_i, r_j)$, we consider one cluster whose cluster head is located at $\boldsymbol{y}_i$ in the network.

We derive a utility function that captures the average transmission power per MTD per cluster $\mathcal{K}_{i,j}$ and globally reflects the joint actions of all MTDs within the cluster. We assume that, MTDs are cooperative. The closed-form expression of the proposed utility function at time slot t, $u_{\mathcal{K}_{i,j}}(t)$, can be derived as follows.

Definition 5.4 The utility function $u_{\mathcal{K}_{i,j}}(t)$ when typical cluster $\mathcal{K}_{i,j}$ chooses type j, at t is:

$$
\begin{aligned}
u_{\mathcal{K}_{i,j}}(t) = &- \frac{1}{j} \left(\frac{1}{A_t \xi(t)} \right) \left(\frac{4\pi r_j}{\mu} \right)^{\nu} \left(\frac{2\,(j-1)\,\mathbb{E}(I_j)}{2+\nu} \right) \times \\
&\left(2^{\frac{H_m}{TB}} - 1 \right) - \frac{1}{j} \left(\frac{BN_0}{g_i} \right) \times \\
&\left(2^{\frac{H_m + H_m(j-1)\left(1 - \frac{\alpha}{\frac{r_j}{c}+1}\right)}{TB}} - 1 \right) - \delta j, \qquad (5.11)
\end{aligned}
$$

where $j = \lfloor \lambda_m \pi r_j^2 \rfloor$ is the cluster type and the average number of MTDs within a typical cluster $K_{i,j}$, i.e., the average cluster size..

The proposed utility function in (5.11) includes three main terms. In the first term, we model the average transmission power of M2M links across all MTDs $m \in \mathcal{K}_{i,j} \setminus \{i\}$ to the reference cluster head i, as defined in (5.4). In the second

term of the utility function, we capture the cellular link transmit power of cluster head i to the BS, as defined in (5.2). Lastly, the third term captures the cluster size, where $\delta \in [-1, 1]$ is a tunable parameter that leverages the costs associated with the preference of forming smaller/larger sized clusters: $\delta < 0$ puts more emphasis on forming larger sized clusters on average to eliminate more redundant bits and reduce MTD transmission power, and $\delta > 0$ puts more weight on minimising the cost of the signalling overhead in the proposed utility. Varying δ between -1 and 1 shows the performance switching from a larger to a smaller sized cluster when δ crosses above 0. Since we are considering a massive number of MTDs, the utility function also takes into account the MTD distance distributions, as defined in (5.7), and inter-cluster interference, as defined in (5.10), in order to accurately model MTD average transmission power. See Appendix C.1 for the derivation of (5.11). Hence, depending on the cluster type of the typical cluster and the value of δ, this will affect the average number of MTDs within the cluster and the average transmission power per MTD per cluster. In general, MTDs will prefer to form a larger sized cluster on average (that is, a high cluster type), with higher data correlation as the number of redundant bits will increase (corresponding to the send part of the utility function in (5.11)), which in turn will reduce MTD's energy consumption.

5.2.2 Dynamics of Cluster Formation

Evolutionary game theory uses biologically-inspired dynamics to model how individual MTDs form different types of clusters (i.e., cluster of different sizes) over time [61, 71]. We assume that the percentage of MTDs that select cluster type $j \in \mathcal{S}$ is x_j, where $x_j \in [0, 1]$ is an element of the population state vector $\boldsymbol{x}$.

To update the percentage of MTDs selecting cluster type j, we adopt properties from continuous-time replicator dynamics (see [61, 71], and [196]). Over time, the MTDs in a cluster $\mathcal{K}_{i,j}$ will update their preference in $\boldsymbol{x}$ and will become more certain about what cluster type they would prefer to form [71, 196]. In continuous-time replicator dynamics, the rate of MTDs selecting cluster type j is proportional to the difference between the fitness of cluster type j and the average expected fitness of the population [61, 195]. The fitness of a type is defined as the average payoff of that type, and is a function of the population state $\boldsymbol{x}$. The evolution of

cluster type j, is given by:

$$\dot{x}_j(t) = x_j(t)(\overline{u}_{\mathcal{K}_{i,j}}(\boldsymbol{x}, t) - U_{\mathcal{K}_i}(\boldsymbol{x}, t)), \tag{5.12}$$

where $\overline{u}_{\mathcal{K}_{i,j}}(\boldsymbol{x}, t)$ is the fitness of cluster $\mathcal{K}_{i,j}$ with average size j, and $U_{\mathcal{K}_i}(\boldsymbol{x}, t)$ is the average expected fitness of the population. The fitness of cluster type j, is defined as:

$$\overline{u}_{\mathcal{K}_{i,j}}(\boldsymbol{x}, t) = \sum_{j' \in \mathcal{S}} u_{\mathcal{K}_{i,jj'}}(t) x_{j'}(t), \tag{5.13}$$

where for all $j, j' \in \mathcal{S}$, $u_{\mathcal{K}_{i,jj'}}(t) = u_{\mathcal{K}_{i,j}}$ if $u_{\mathcal{K}_{i,j}} \geq u_{\mathcal{K}_{i,j'}}$, or $u_{\mathcal{K}_{i,jj'}}(t) = u_{\mathcal{K}_{i,j'}}$ if $u_{\mathcal{K}_{i,j'}} > u_{\mathcal{K}_{i,j}}$, and $u_{\mathcal{K}_{i,j}}$ is given by (5.11). The average expected fitness of the population, $U_{\mathcal{K}_i}(\boldsymbol{x}, t)$, is given by:

$$U_{\mathcal{K}_i}(\boldsymbol{x}, t) = \sum_{j \in \mathcal{S}} \overline{u}_{\mathcal{K}_{i,j}}(\boldsymbol{x}, t) x_j(t). \tag{5.14}$$

The evolution of cluster type j for the MTD population in (5.12) can be either greater than, less than, or equal to zero. If $\dot{x}_j$ is greater than 0, this implies that the fitness of cluster type j is greater than the average expected fitness of the population, and thus the percentage of the population selecting cluster type j is increasing. If $\dot{x}_j$ is less than 0, this indicates that the percentage of the population selecting cluster type j is decreasing, that is, cluster type j is growing extinct. When $\dot{x}_j$ is equal to 0 then, we have a stationary point for the percentage of MTDs selecting cluster type j (an evolutionary equilibrium for cluster type j). Thus, no MTD has incentive to change their cluster type from j, as the fitness of cluster type j is equal to the average expected fitness of the population.

Within the proposed evolutionary game, conflicts between MTDs may arise, which are reflected in the replicator dynamics, in (5.12) [197, 198]. For example, a conflict may arise between MTDs within cluster $\mathcal{K}_{i,j}$, where $j < N$, and other MTDs outside of the cluster that can potentially become a cluster member. For instance, the conflict could be pertaining to the cluster size preference that maximises the fitness for a particular cluster size. If such a conflict occurs, then this may cause a decrease in the fitness for certain cluster types, which may lead to certain cluster sizes becoming extinct [61, 197, 198]. In order to resolve such a conflict, MTDs involved in the conflict will update/change their cluster type preference (membership) over time, until the conflict is resolved. The proposed evolutionary

game approach can naturally enable such updates.

5.3 Proposed Algorithm

To solve the proposed evolutionary game, we propose a distributed algorithm to find a stable cluster formation for correlated-aware clustering in M2M communications. The proposed distributed algorithm, shown in Algorithm 5, enables a massive number of MTDs to autonomously update their type to form the best cluster. To clarify how Algorithm 5 works, we have included a flow chart in Fig. 5.2. The network operator defines the cluster head selection policy, as outlined in Phase I in Fig. 5.2. Then during Phase III, at each iteration, MTDs update their preferences for each cluster type, where the cluster head is selected based on the defined policy. Thus, as we can see from Fig. 5.2 and (5.11), the cluster formation in the proposed evolutionary algorithm depends on the defined cluster head selection policy. Here, we assume a cluster head selection policy whereby cluster members share the role of cluster head for an equal amount of time. During the dynamic cluster formation process in Algorithm 1, MTDs will determine a consensus of the data sampled from MTDs within each cluster. If for instance, an MTD loses its battery power at time slot $t+1$, during the cluster formation process, its cluster type preference changes from $x_j(t) > 0$ and becomes $x_j(t+1) = 0$, $\forall j \in \mathcal{S}$. Thus, the MTD has no cluster type preference over the set $\mathcal{S}$. Meanwhile, during the cluster formation process, if an MTD is unable to update its status/preference within a cluster, then the MTD will need to wait till the next round or iteration, before it can be updated. Hence, Algorithm 5 outputs the population state, $\boldsymbol{x}^{\circ}$, which includes the percentage of MTDs in the population selecting a cluster type from the strategy set.

5.4 Evolutionary Game Stability Analysis

To solve the evolutionary game, we consider the concept of an ESS, which is defined as:

Definition 5.5 The population state $\boldsymbol{x}^* \in \mathbb{R}^N$ is an *ESS*, if there exists a portion of MTDs $\epsilon_j^* > 0$ for each cluster size $j \in \mathcal{S}$, such that for all $0 < \epsilon < \epsilon_j^*$ and for all

$j \in \mathcal{S}$:

$$U_{\mathcal{K}_i}(x_j^*(t),(1-\epsilon)\boldsymbol{x}_{-j}^*(t)+\epsilon\boldsymbol{x}'_{-j}(t)) > U_{\mathcal{K}_i}(x'_j(t),(1-\epsilon)\boldsymbol{x}_{-j}^*(t)+\epsilon\boldsymbol{x}'_{-j}(t)) \quad (5.15)$$

where $\boldsymbol{x}'(t) \in \mathbb{R}^N$ is any population state which is different from $\boldsymbol{x}^*(t)$, i.e., $\boldsymbol{x}'(t) \neq \boldsymbol{x}^*(t)$ and $x'_j(t)$ is an element of $\boldsymbol{x}'(t)$. In particular, $\epsilon\boldsymbol{x}'(t)$ represents portion ϵ of MTDs from the population, that will choose a strategy from the population state $\boldsymbol{x}'(t)$ instead of $\boldsymbol{x}^*(t)$.

The ESS is robust to a small portion $\epsilon_x \in (0,1), \forall\ \epsilon \in (0,\epsilon_x)$, of MTDs changing their cluster type. This change in cluster type could be due to the joining of new MTDs, rapid fluctuations in the sensing environment, or MTD loss of battery. The replicator dynamics in (5.12) capture the dynamics of distributed MTD clustering in our proposed algorithm. Hence, at the convergence of the ESS, the number of possible cluster sizes is constant, and will range from 1 to N. Next, we prove that the proposed Algorithm 5 converges to a population state $\boldsymbol{x}^*$, and we find the maximum portion of MTDs, ϵ, that may deviate from an ESS, based on Definition 5.5.

Theorem 5.1 The proposed Algorithm 5 converges to a population state, $\boldsymbol{x}^*$, which is an ESS for the proposed evolutionary correlation-aware clustering game in G_E. At the convergence, the maximum portion of MTDs that can deviate from ESS, ϵ^*, is given as follows:

$$\epsilon^* = \min_{j\in\mathcal{S}} \frac{1}{2\overline{u}_{\mathcal{K}_{i,j}}(\boldsymbol{x}^*,t) - u_{\mathcal{K}_{i,j}}(\boldsymbol{x}^*,t)}. \quad (5.16)$$

Proof: See Appendix C.2. ■

5.5 Simulation Results

For our simulations, we consider a single BS located at the center of a circular area with a 2 km radius. We consider a Poission-based distribution for distributing MTDs around a randomly placed MTD within the network. Given the network radius $r_N = 2$ km and a network density $\lambda_m = 0.09$, the average number of MTDs within the network is $N = 1.13 \times 10^6$, whereas the average number of MTDs within

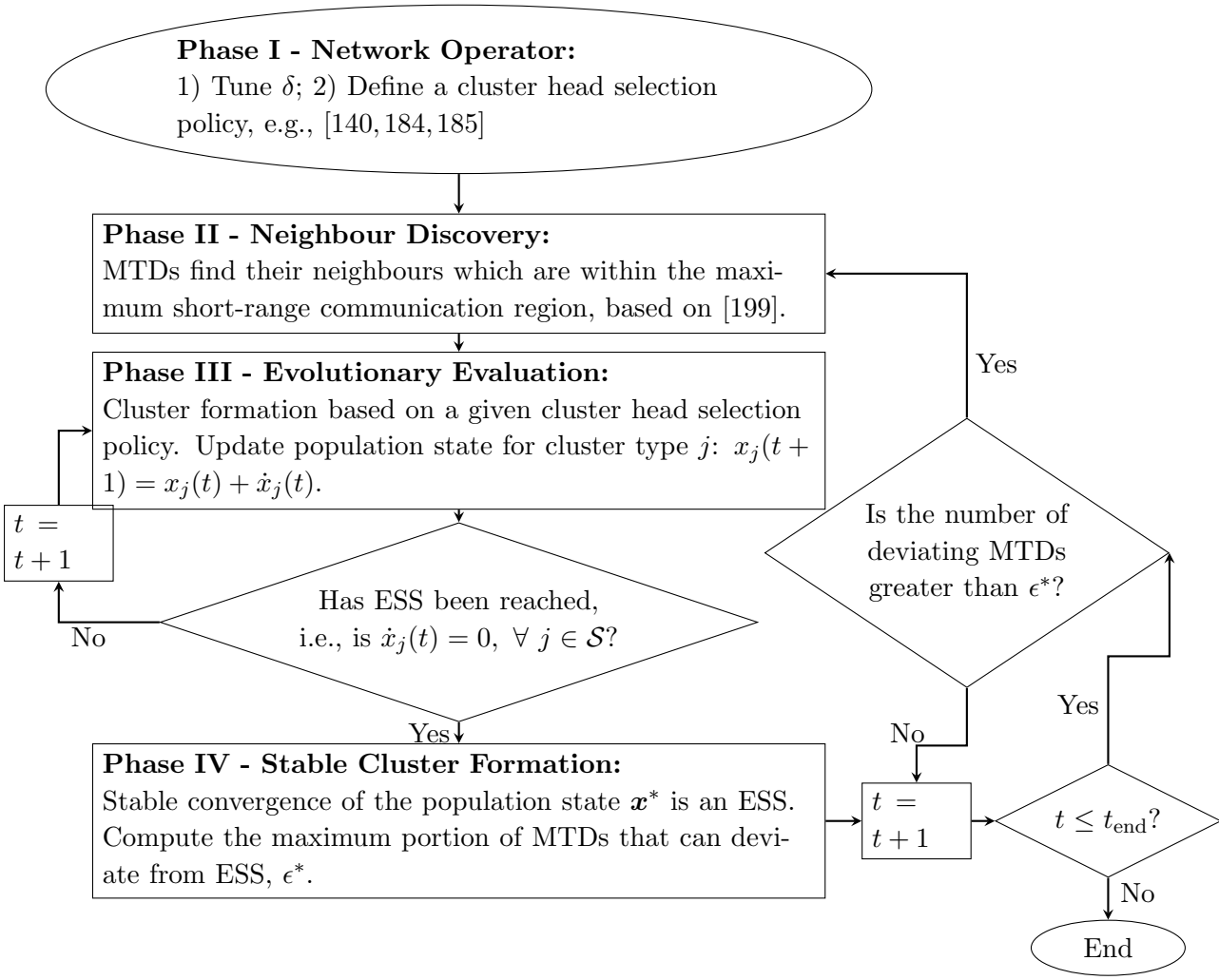


Figure 5.2: Flow chart of the proposed algorithm.

Algorithm 5 Evolutionary Game for Correlation-Aware Clustering in Massive M2M Network

Input: Network density: λ_m; Set of MTDs: $\mathcal{M}$; Maximum radius: r_N;
Phase I - Network Operator
Tune δ to either minimise or maximise the cluster size (which is equivalent to minimising or maximising the signalling overhead);
Define a cluster head selection policy, e.g., [140, 184], or [185].
Phase II - Neighbour Discovery
A ball $b(\boldsymbol{y}_i, r_N)$ is randomly placed in the network, centered at $\boldsymbol{y}_i$, $i \in \mathcal{M}$;
A neighbour discovery protocol based on [199] is used to determine the number of MTDs in the ball, N;
Phase III - Evolutionary Evaluation
Define the set of cluster types for a cluster as $\mathcal{S} = \{1, 2, \ldots, N\}$, and set the initial population state for all cluster types as $\boldsymbol{x}(t = 0)$;
repeat
    for cluster type $j \in 1 \ldots N$ **do**
        Find the inter-cluster interference, $\mathbb{E}(I_j)$, given the network density λ_m and radius of the ball r_j;
        Based on a defined cluster head selection policy, the utility function of MTD i in a cluster with radius r_j, $u_{\mathcal{K}_{i,j}}(t)$, is calculated using (5.11);
        MTDs in a cluster with radius r_j, determines the evolution of their preferences, $\dot{x}_j(t)$, using (5.12), (5.13), and (5.14);
        MTDs in a cluster with radius r_j, which forms cluster type j, updates their population state: $x_j(t+1) = x_j(t) + \dot{x}_j(t)$;
    end for
    $t = t + 1$;
until Convergence to ESS and $\dot{x}_j(t) = 0 \ \forall j \in \mathcal{S}$
Phase IV - Stable Cluster Formation
Output: Population state, $\boldsymbol{x}^\circ = \{x_j | \forall j \in \mathcal{S}\}$, that represents the formed MTD clusters, where the cluster heads are selected based on the chosen policy;

the network, for a network density of $\lambda_m = 0.36$, will be $N = 4.52 \times 10^6$. We focus on a small section of the circular area around the reference cluster head, that has a 500 m radius. The BS allocates orthogonal resource blocks to cluster heads that they can use to send the data of the clusters to the BS. Within each cluster, each MTD is also allocated an orthogonal resource block, to send its data to the cluster head. However, M2M links in different clusters can simultaneously reuse the same resource blocks, resulting in inter-cluster interference. Each resource block has a fixed bandwidth of $B = 180$ kHz, and the maximum transmission power over the cellular link is $P_{\max} = 35$ dBm and the M2M link is $Q_{\max} = 20$ dBm. The duration of each time slot t is fixed to $T = 1$ ms. We consider a carrier frequency of 2 GHz. Furthermore, we assume MTDs are static, where the cellular and M2M links have a path loss exponent ν of 2.5. The cellular and M2M links have a transmit and

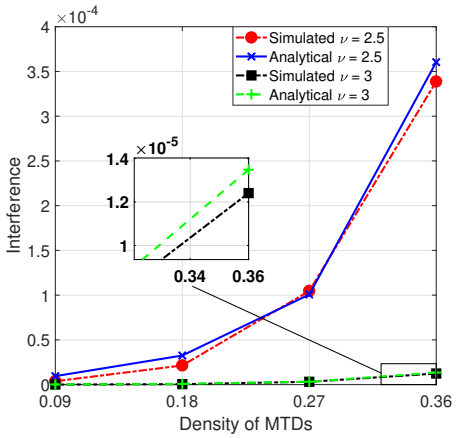


Figure 5.3: Simulated and analytical interference as a function of MTD density, λ_m, where $c = 6$.

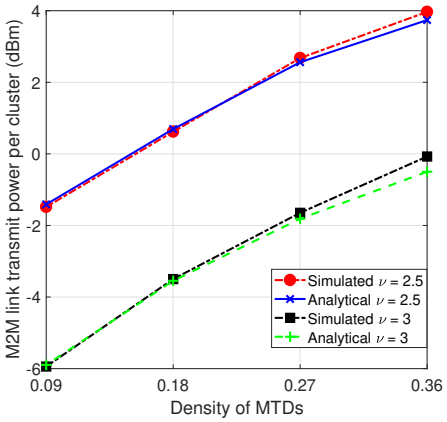


Figure 5.4: Simulated and analytical M2M link transmit power as a function of MTD density, λ_m, where $c = 6$.

receive antenna gain of 9.54 dB. The noise power spectral density over a cellular link is -174 dBm/Hz. The data source of each MTD m has $\mu_m = 0$ and $\sigma_m = 10$. The quantization step for each MTD is set to $\Delta = \frac{1}{256}$ [44].

The evolutionary game is evaluated using Monte Carlo simulations with varying MTD density λ_m, path loss exponent ν, and correlation constant c. Since the number of MTDs considered is massive, it is not possible to use conventional techniques such as coalition formation games [67] that can exhibit exponential complexity in large systems. Thus, at convergence we compare our proposed evolutionary game to two benchmarks, which are: (i) a pure cluster type in which the BS directly-partitions the population of MTDs to form the maximum possible cluster size N, with preference 1; and (ii) uniform cluster type in which an equal percentage of the population will be uniformly distributed across all cluster types, that is, each cluster type has preference $\frac{1}{N}$.

Fig. 5.3 shows the inter-cluster interference for different network densities and path loss exponents. From this figure, we can see that the analytical results derived using (5.10) closely match the corresponding simulation results. Also, in Fig. 5.3, we observe that, as the network density increases, this causes the inter-cluster interference to increase, due to the ball radius, r_N, decreasing as shown in Fig. 5.9. Furthermore, by increasing the path loss exponent from $\nu = 2.5$ to $\nu = 3$ the inter-cluster interference will be reduced due to the higher propagation losses.

Fig. 5.4 shows the M2M link transmit power per cluster for different network densities and path loss exponents. From this figure, we can see that the analytical

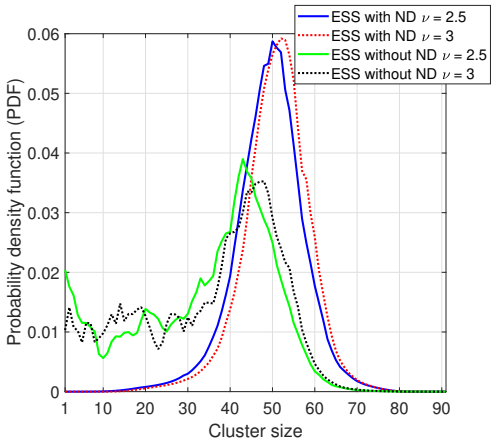


Figure 5.5: Percentage of MTDs forming a particular cluster size as a function of path loss exponents, ν, where $\lambda_m = 0.09$ and $c = 6$.

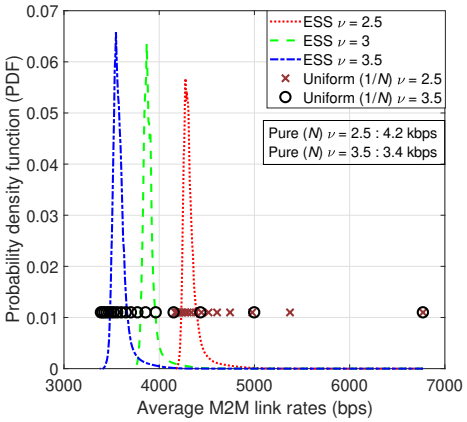


Figure 5.6: Percentage of MTDs selecting a particular transmission M2M link rate as a function of path loss exponents, ν, where $\lambda_m = 0.09$ and $c = 6$.

results derived using (5.11) closely match the corresponding simulation results. Also, in Fig. 5.4, we observe that, as the network density increases, this causes the M2M link transmit power per cluster to increase, due to the increasing interference as shown in Fig. 5.3. Hence, increasing the path loss exponent from $\nu = 2.5$ to $\nu = 3$ will reduce the M2M link transmit power due to the higher propagation losses.

In Fig. 5.5, we show the probability density function (PDF) of MTD preference for forming certain cluster sizes within the ball, with and without neighbour discovery (ND), under different path loss exponents, ν. On average, our proposed evolutionary algorithm with ND, has a mean cluster size of 49.3 and a variance of 65.2 for $\nu = 2.5$, and a mean cluster size of 50.9 and a variance of 59.9 for $\nu = 3$. On the other hand, our proposed evolutionary algorithm without ND, has a mean cluster size of 34.3 and a variance of 253.6 for $\nu = 2.5$, and a mean cluster size of 35.4 and a variance of 262.5 for $\nu = 3$. Therefore, our proposed algorithm with ND improves the performance of MTDs, as they form clusters with greater preference and certainty.

Fig. 5.6 shows the PDF of MTD preference for using certain M2M transmission rates within a cluster, under different path loss exponents ν. We observe that for the proposed evolutionary algorithm, MTDs prefer to transmit with an M2M link rate of 4.28 kbps for 5.7 % of the time for $\nu = 2.5$, whereas MTDs prefer to transmit with an M2M link rate of 3.55 kbps for 6.58 % of the time for $\nu = 3.5$. Thus,

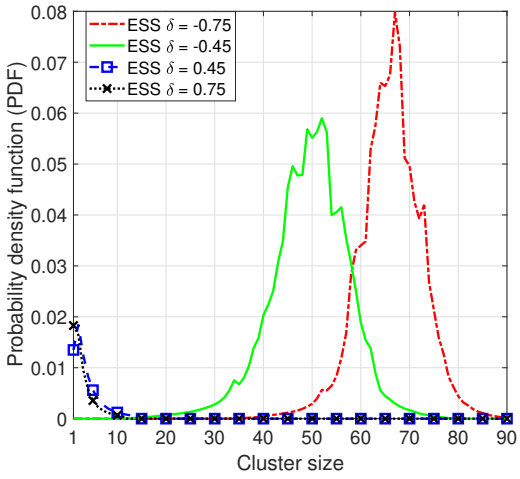


Figure 5.7: Percentage of MTDs forming a particular cluster as a function of δ, where $\lambda_m = 0.09$, $c = 6$, and $\nu = 2.5$.

increasing the path loss exponent will lead to a decrease in M2M link rates and a decrease in M2M link communication range. Moreover, the uniform benchmark has a similar mean and variance to the proposed evolutionary algorithm for $\nu = 2.5$ and $\nu = 3.5$. However, the proposed evolutionary algorithm has a higher maximum preference for particular M2M link rates. In addition, the pure benchmark has similar maximum preferences for M2M link rates as well. Note that, the uniform and pure benchmarks are not robust to potential changes in the M2M network.

Fig. 5.7 shows the PDF of MTD preference for forming certain cluster sizes, under different values for δ. We observe that, as the value of δ decreases below 0, the MTD population prefers to form larger sized clusters on average. The results show that in the proposed evolutionary algorithm the MTD population prefers to form a cluster of size 52 for 6 % of the time for $\delta = -0.45$, and a cluster size of 67 for 8 % of the time for $\delta = -0.75$. On the other hand, the results show that, in the proposed evolutionary algorithm, the MTD population prefers to form a cluster of size 2 for 1.85 % of the time for $\delta = 0.45$, and a cluster size of 1 for 1.83 % of the time for $\delta = 0.75$. Hence, these results validate that δ can change an MTD's objective. A network operator can then use these results to decide on how to optimise δ for their purposes.

Fig. 5.8 shows the PDF of MTD preference for forming certain cluster sizes within the ball, under different network densities, λ_m, and correlation constants, c. We observe that, as the network density and correlation constant increase, the MTDs' preferences for forming particular cluster sizes does not change significantly.

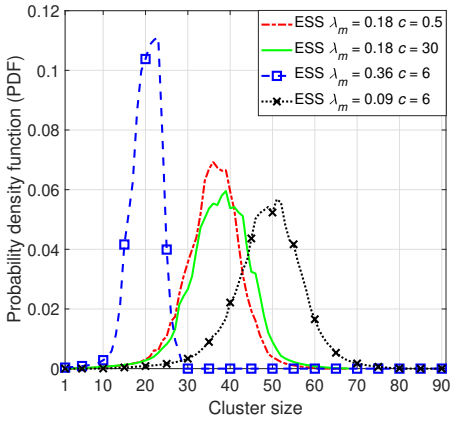


Figure 5.8: Percentage of MTDs forming a particular cluster size as a function of correlation constants, c, and MTD density, λ_m, where $\nu = 2.5$.

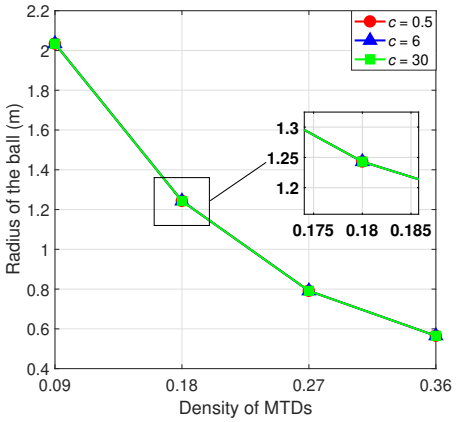


Figure 5.9: Radius of the ball as a function of MTD density, λ_m, and correlation constants, c, where $\nu = 2.5$.

This is due to the fact that the maximum radius of the ball, defined in (5.6), is a function of MTD density. As illustrated in Fig. 5.9, the radius of the ball changes according to the density, thus maintaining the number of MTDs within the ball consistent. Furthermore, we observe that, in the proposed evolutionary algorithm, the MTD population prefers to form a cluster of size 36 for 6.93% of the time for $c = 0.5$ and $\lambda_m = 0.18$, whereas the MTD population prefers to form a cluster of size 39 for 5.95% of the time for $c = 30$ and $\lambda_m = 0.18$. As the correlation constant increases within the cluster, the MTD population prefers to form slightly larger clusters. The results also show that, in the proposed evolutionary algorithm, the MTD population prefers to form a cluster of size 23 for 11.13% of the time for $\lambda_m = 0.36$ and $c = 6$, whereas the MTD population prefers to form a cluster of size 51 for 5.67% of the time for $\lambda_m = 0.09$ and $c = 6$.

Fig. 5.10 shows the average transmission power per MTD per cluster over time, with and without ND, under different correlation constants, c. In this figure, we can see that any increase in the correlation constant will lead to a decrease in the average transmission power per MTD per cluster, due to the increase in data correlation within the cluster. Over time, the proposed evolutionary algorithm with ND converges to an average transmit power per MTD per cluster of 29.12 dBm at $t = 232$ for $c = 0.5$, 26.6 dBm at $t = 269$ for $c = 6$, and 26.06 dBm at $t = 267$ for $c = 30$. On the other hand, the proposed evolutionary algorithm without ND converges to an average transmit power per MTD per cluster of 26.83 dBm at

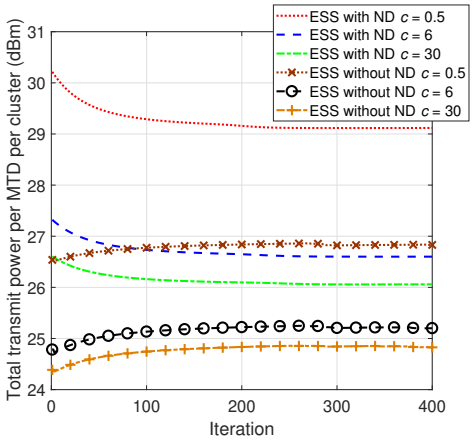


Figure 5.10: Dynamics of average transmit power per MTD per cluster, where $\lambda_m = 0.09$ and $\nu = 2.5$.

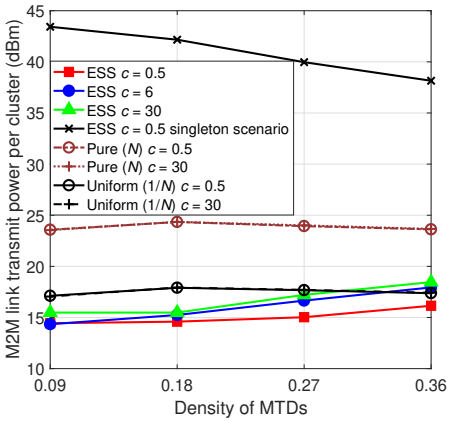


Figure 5.11: Total M2M link transmit power per MTD per cluster as a function of MTD density, λ_m, where $\nu = 2.5$.

$t = 319$ for $c = 0.5$, 25.2 dBm at $t = 381$ for $c = 6$, and 24.83 dBm at $t = 380$ for $c = 30$. Overall, the scenario without ND reduces energy consumption of MTDs, due to forming smaller sized clusters on average. On the other hand, the scenario with ND converges faster to the ESS than the scenario without ND.

Moreover, in Fig. 5.11, we show the total M2M link transmit power per MTD per cluster at convergence, under different network densities and correlation constants. We observe that, as the network density and correlation constant increase, the cluster size and data correlation will increase. As cluster size and data correlation increase, the total M2M link transmit power per MTD per cluster will also increase. We compare the proposed evolutionary-based correlation aware clustering algorithm for different correlation constants to the singleton scenario, and two benchmarks. The singleton scenario, assumes all MTDs within the ball transmit their own data via the cellular link to the BS, thus no clustering within the ball. On average, the transmit power per cluster is reduced of around 63.1% for $c = 0.5$, 60.5% for $c = 6$, and 59.8% for $c = 30$, compared to the singleton scenario. The proposed evolutionary algorithm decreases the transmission power per cluster of around 36.9% for $c = 0.5$, 32.8% for $c = 6$, and 31.6% for $c = 30$, compared to the pure cluster type benchmark for $c = 0.5$, whereas transmission power per cluster decreases of around 37% for $c = 0.5$, 32.8% for $c = 6$, and 31.7% for $c = 30$, compared to the pure cluster type benchmark for $c = 30$. On the other hand, the proposed evolutionary algorithm also decreases transmission power per cluster of around 14% for $c = 0.5$, 8.4% for $c = 6$, and 6.8% for $c = 30$, compared to the

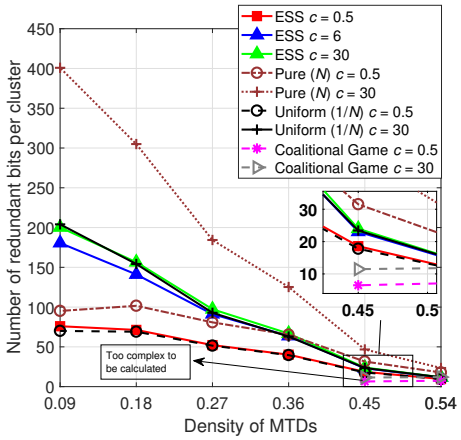


Figure 5.12: Number of redundant bits per cluster vs. MTD density, λ_m, where $\nu = 2.5$.

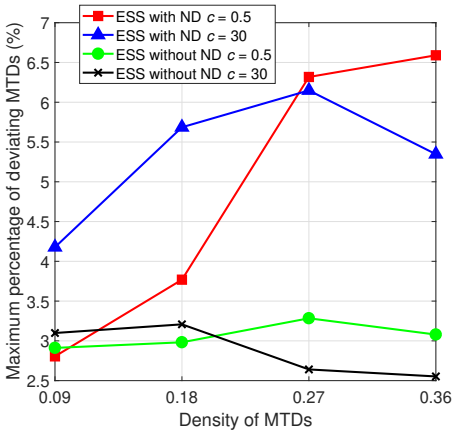


Figure 5.13: Maximum percentage of deviating MTDs from ESS vs. MTD density, λ_m, where $\nu = 2.5$.

uniform cluster type benchmark for $c = 0.5$, whereas transmission power per cluster decreases around 14.1% for $c = 0.5$, 8.5% for $c = 6$, and about 7% for $c = 30$, compared to the uniform cluster type benchmark for $c = 30$. On average, the proposed evolutionary algorithm minimises transmit power by 44.1% and 15.25% across all correlation constants and network densities, compared to the pure cluster type benchmark and the uniform cluster type benchmark respectively.

Fig. 5.12 shows the average number of redundant bits per cluster, for different network densities and correlation constants. The average number of redundant bits per cluster, is the amount of bits that could be removed by the cluster head, before transmitting the cluster data to the BS. Thus, decreasing network density and increasing the correlation constant, as in Fig. 5.12, leads to an increase in data correlation, cluster size, and the number of redundant bits. On average, in Fig. 5.12, the proposed evolutionary-based correlation aware clustering algorithm increases the number of redundant bits by more than double, that is, 50.6 %, when increasing the correlation constant from $c = 0.5$ to $c = 30$. In Table 5.3, we show the average gains (in percentage) resulting from the proposed evolutionary clustering algorithm compared to the three benchmarks, i.e., the pure cluster type benchmark, uniform cluster type benchmark, and a coalitional game approach. As shown in Fig. 5.12 and Table 5.3, the pure benchmark with correlation constant $c = 30$ maximises the number of redundant bits per cluster across all network densities, as this benchmark considers all MTDs preferring to form the maximum possible size cluster within the ball. Thus, considering the results of Figs. 5.10 and

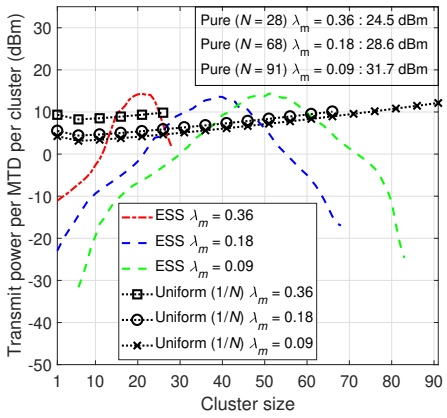


Figure 5.14: Total transmit power per cluster vs. cluster sizes and MTD density, λ_m, where $\nu = 2.5$ and $c = 6$.

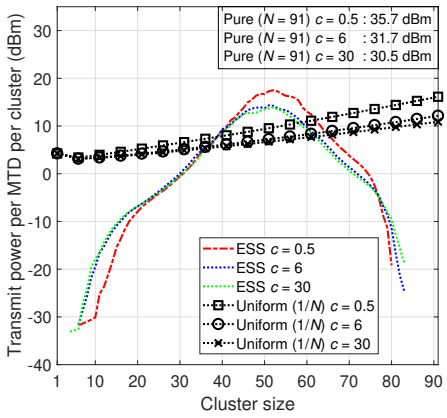


Figure 5.15: Total transmit power per cluster vs. cluster sizes and correlation constant, c, where $\nu = 2.5$ and $\lambda_m = 0.09$.

Table 5.3: Average gains for the number of redundant bits per cluster, Fig. 5.12.

ESS	**Pure** $c = 0.5$	**Pure** $c = 30$	**Uniform** $c = 0.5$	**Uniform** $c = 30$	**Coalition** $c = 0.5$	**Coalition** $c = 30$
$c = 0.5$	-30.9%	-74.3%	3.9%	-49%	108.1%	21.1%
$c = 6$	34.7%	-52.1%	100.3%	-5.1%	155.8%	48.7%
$c = 30$	47.1%	-48.1%	118.3%	2.7%	164.3%	53.6%

5.11, both benchmarks yielded, on average, a larger transmit power and number of redundant bits per cluster, compared to the proposed distributed algorithm. However, the pure benchmark requires a centralized approach in order to achieve these results, whereas the uniform benchmark requires a distributed approach. On the other hand, we have observed that, as the network density increases, the computational complexity needed for the coalitional game approach exponentially increases. With regards to the performance of the coalitional game approach, at convergence, the coalitional game approach forms smaller clusters compared to the proposed evolutionary algorithm. In addition, the proposed approach significantly reduces the number of redundant bits per cluster by up to 41 % compared to the coalitional game solution for $c = 30$. Furthermore, these benchmarks and the coalitional game are not robust to stochastic changes in the M2M network environment.

In Fig. 5.13, we show the maximum percentage of deviating MTDs at an ESS for our proposed evolutionary game with and without ND, under different network

densities, λ_m, and correlation constants, c. For the proposed algorithm with ND, as the density and correlation constant increase (that is, as the network becomes denser and data in the cluster becomes more correlated) the maximum percentage of deviating MTDs also increases. On the other hand, for the proposed algorithm without ND, as the density and correlation constant increase the maximum percentage of deviating MTDs stays constant.. Hence, for a correlation constant, $c = 0.5$, when the network density is $\lambda_m = 0.09$, the maximum portion of MTDs that will deviate, due to stochastic changes in the M2M environment, is about 2.8 % with ND and about 2.9 % without ND. However, as the density of the network increases to $\lambda_m = 0.36$, the maximum percentage of MTDs that will deviate is about 6.59 % with ND and about 3.1 % without ND. Therefore, at the ESS with ND, as the network becomes denser and more correlated, MTDs have more options to deviate compared to when the network density is sparse. Additionally, when the network becomes more correlated, there is a point beyond which increasing the network density no longer increases the maximum percentage of MTDs that may deviate within the network. Overall, the proposed evolutionary algorithm with ND is more robust at the ESS.

Fig. 5.14 shows the average transmit power per MTD per cluster as a function of cluster size, for different network densities, where $c = 6$ and $\nu = 2.5$. We observe a concave relationship between cluster size and the transmit power per MTD per cluster. Depending on the network density, the cluster size changes, when the cluster radius changes. As the network density decreases, the ball radius will also decrease, leading to an increase in the potential number of MTDs per cluster. On the average, the proposed evolutionary algorithm reduces transmit power per MTD per cluster around 87.3% for $\lambda_m = 0.36$, 101.8% for $\lambda_m = 0.18$, and 100.8% for $\lambda_m = 0.09$, compared to the uniform cluster type benchmark. On the other hand, the proposed evolutionary algorithm reduced transmit power per MTD per cluster around 69% for $\lambda_m = 0.36$, 122.4% for $\lambda_m = 0.18$, and 146.9% for $\lambda_m = 0.09$, compared to the pure cluster type benchmark.

Fig. 5.15 shows the average transmit power per MTD per cluster as a function of the cluster size, for different correlation constants, where $\lambda_m = 0.09$ and $\nu = 2.5$. We observe a concave relationship between cluster size and average transmit power per MTD per cluster, that is similar to the one shown in Fig. 5.14. In Fig. 5.15, as the correlation constant increases, the cluster sizes generally remain constant,

but the maximum average transmit power per MTD per cluster decreases. As a given cluster becomes more correlated, the MTDs will then send less information over the cellular link. On the average, the proposed evolutionary algorithm reduces transmit power per MTD per cluster around 98.7% for $c = 0.5$, 98.8% for $c = 6$, and 98.7% for $c = 30$, compared to the uniform cluster type benchmark. On the other hand, the proposed evolutionary algorithm reduced transmit power per MTD per cluster around 158.5% for $c = 0.5$, 142.2% for $c = 6$, and 140% for $c = 30$, compared to the pure cluster type benchmark.

5.6 Summary of Contributions

In this chapter, we have proposed a novel distributed correlation-aware clustering scheme for reducing the number of redundant bits being sent to the BS, as well as reducing transmission power for each MTD in a massive and locally finite M2M network. In the proposed model, MTDs self-organise in a fully distributed manner to form clusters, based on data correlation and transmission power savings for a given resource block allocation mechanism. Hence, in this chapter we have answered questions Q6 and Q7, posed in Section 1.2.

We have modelled the problem using evolutionary game theory and stochastic geometry. Stochastic geometry has been used to accurately model and characterize the distance distributions between MTDs, in order to derive a closed-form expression for average inter-cluster interference. Then, we derived a closed-form expression for average power consumption per cluster, as a function of MTD location, cluster size, and inter-cluster interference. Based on evolutionary game theory, we have proposed a distributed clustering algorithm, in which MTDs autonomously seek to minimise the average MTD transmission power per cluster. We have shown that the proposed distributed algorithm converges to a stable state, that is, an ESS which is robust to a small portion of MTDs deviating from the stable cluster formation. Moreover, we have also derived a maximum portion of MTDs that can deviate from the ESS, while still maintaining a stable cluster formation. Simulation results showed that the proposed evolutionary algorithm decreases transmission power per cluster, and increases the number of redundant bits that can be eliminated in a given cluster.

Chapter 6

Utilizing UAVs in Mission Critical M2M Communications

Chapters 5 focused on distributed correlation-aware clustering for a massive number of MTDs densely deployed in a cellular network. However, some multimedia IoT applications may require MTDs to have ultra-reliability and low latency as in URLLC. Ensuring URLLC, for mission critical M2M communications, will drastically decrease end-to-end latency and increase communication reliability [42]. However, MTDs may be deployed in areas that experience intermittent or poor coverage from cellular BSs, or MTDs may be deployed in environments with no cellular BS infrastructure, and thus, will not be able to ensure URLLC constraints.

In this chapter, we consider the problem of energy-aware scheduling for mission critical M2M communications, where UAVs are utilized as flying BSs. In particular, we consider an IoT network, where MTDs are clustered and aggregators are employed as cluster heads to collect the sensed data from MTDs. In the considered IoT network, UAVs must optimise their schedule in order to collect the data from each aggregator in a distributed manner, without requiring a-priori knowledge of the probabilities associated with the event-driven arrival requests from MTDs. Hence, we formulate the ultra reliable energy-aware scheduling problem by combining both Lyapunov optimisation and matching theory. The use of Lyapunov optimisation is appropriate as it does not require a-priori knowledge of the event-driven traffic of MTDs. Using the technique of Lyapunov optimisation, the problem is transformed into a linear weighted function, which ensures that both the maximum queue length and the queue length variance across all aggregators in the network are bounded. Based on the Lyapunov optimisation of our problem, matching theory is used to formulate and model the energy-aware scheduling problem as

a one-to-many matching game. In the formulated game, the interactions between a number of aggregators that seek to be served by a UAV is modeled. In fact, each UAV will determine a subset of aggregators that it will serve and sufficient dwelling time in order to satisfy the URLLC constraint. Meanwhile, aggregators aim to minimise their energy consumption when transmitting their data to their serving UAV. In particular, based on the Lyapunov optimisation, we define a utility function for the set of UAVs, which captures the URLLC constraint, i.e., high reliability and low-latency constraints, as well as the network stability requirements for mission critical M2M communications. We also define a utility function for the set of aggregators, which captures a trade-off between maximising throughput and minimising energy. Therefore, we propose a distributed energy-aware scheduling and power control algorithm, scheduling UAVs and minimising aggregator energy consumption respectively, which converges to a two-sided exchange-stable matching. Hence, simulation results show that the maximum queue length mean and variance are well below the finite threshold values, which ensures that the UAVs deployed in the network are able to satisfy the URLLC constraints for mission critical M2M communications. In addition, a trade-off is observed between reducing aggregator transmission power and increasing network throughput, when varying the tunable Lyapunov parameter.

This chapter is organised as follows. In Section 5.1, we define the system model and problem formulation for energy-aware scheduling for UAVs in mission critical M2M communications. In Section 6.2, a conditional Lyapunov drift-plus-penalty optimisation is used to find an approximate solution for the joint scheduling and energy optimisation. Then, in Section 6.3, we take the approximate solution from the Lyapunov optimisation and propose a fully distributed matching game to efficiently schedule the UAVs and reduce aggregator transmission power, while guaranteeing URLLC. Section 6.4 outlines the proposed algorithm. Within Section 6.5, the proposed matching game computational cost and convergence is analysed. Simulation results are illustrated in Section 6.6. Finally, Section 6.7 concludes the chapter.

6.1 System Model

Consider an IoT system consisting of M MTDs in a set $\mathcal{M}$, distributed over a given geographical area. In this area, a set $\mathcal{K}$ of K UAVs (low altitude platform)

are also deployed as flying BSs to collect data from the MTDs via *the uplink*. In this network, a finite set $\mathcal{A}$ of G aggregators is deployed. The network's MTDs form a collection $\mathcal{G}$ of G disjoint clusters, where $\mathcal{G}_g \subseteq \mathcal{M}$. In each cluster $\mathcal{G}_g$, the aggregator g is designated as the cluster head (CH) who is responsible for relaying all the received data packets via a uplink to a UAV. The location of each aggregator g within the network is denoted by $\boldsymbol{w}_g = (x_g, y_g)$. Thus, MTDs within each cluster $\mathcal{G}_g$ transmit their data packets at each time slot to the aggregator (i.e., the CH). In order for each UAV $k \in \mathcal{K}$ to collect the uplink data, they must dynamically move and stop above each aggregator. The number of time slots that UAV k must spend above CH g is called the *dwelling time*. The dwelling time depends on the number of packets that CH g wants to transmit to UAV k.

We consider discrete time slots, t, with fixed duration τ. At each time slot t, the location of each UAV k in the network is given by $\boldsymbol{f}_{kt} = (x_{kt}, y_{kt}, z_{kt})$, respectively. In our model, blocks of consecutive discrete time slots are referred to as a *decision epoch*, which has finite length H. Each decision epoch is indexed by h. We assume that the length of the decision epoch H is given, and is the same across all UAVs in the network. For example, if a decision epoch consists of 3 consecutive time slots (i.e., a finite length of $H = 3$), then the duration of the decision epoch will be 3τ. At the beginning of each decision epoch h, each UAV k needs to decide on: 1) The subset of CHs $\mathcal{S}_{kh} \subset \mathcal{A}$ it will serve; 2) The path $\mathcal{P}_{kh}$ it will take to cover the set of selected CHs in $\mathcal{S}_{kh}$; and 3) The set $\mathcal{T}_{kh} = \{\mathcal{T}_{gk,h} | g \in \mathcal{S}_{kh}\}$ of dwelling times for UAV k to cover all CHs $g \in \mathcal{S}_{kh}$ during a decision epoch h, where $\mathcal{T}_{gk,h}$ is a set of $T_{gk,h}$ time slots.

The path $\mathcal{P}_{kh}$ is an ordered pair set, where $\{(i, j)\} \in \mathcal{P}_{kh}$ implies that the UAV will move from CH i to CH j at decision epoch h, where $i, j \in \mathcal{S}_{kh}$. We assume that the UAV will take the shortest distance between CHs $\mathcal{S}_{kh}$. Moreover, in each UAV path, a UAV will not revisit a CH during a decision epoch. Thus, the duration of time that UAV k needs to fly from the coverage area of CH i to CH j where $\{(i, j)\} \in \mathcal{P}_{kh}$, is given by:

$$\tau_{ij,k} = \frac{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}{v_k}, \tag{6.1}$$

where v_k is the speed of UAV k.

Fig. 6.1 represents an illustrative example of five clusters being serviced by two

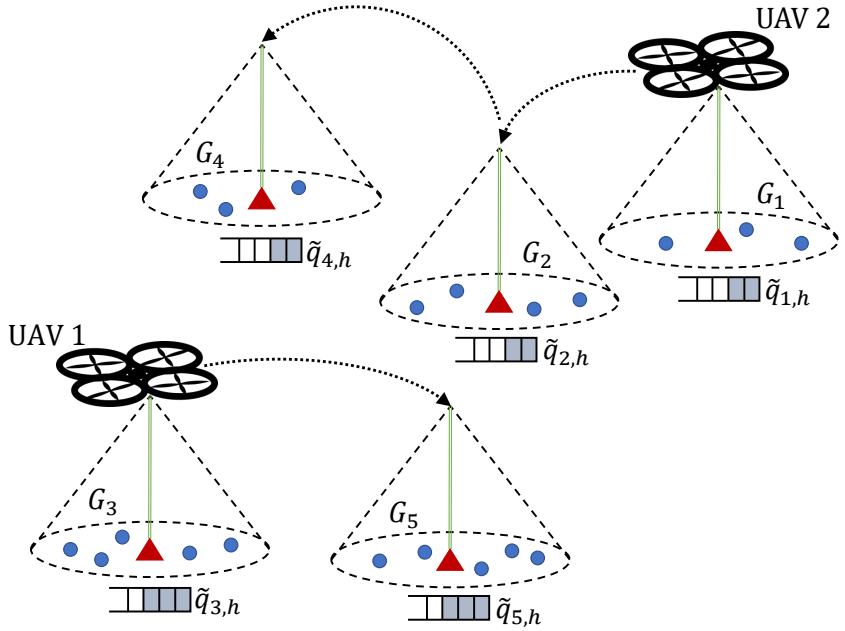
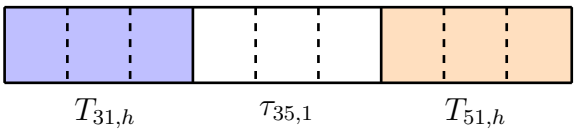
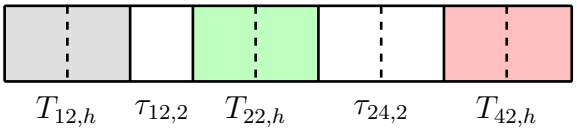


Figure 6.1: System model of a clustered M2M network, with UAVs as flying BSs.



a) UAV 1 decision epoch event sequence



b) UAV 2 decision epoch event sequence

Figure 6.2: Decision epoch example based on system model example in Fig. 6.1.

UAVs. As illustrated in Fig. 6.1, the UAVs will stop directly above each aggregator to collect the data packets from the cluster. In this example scenario, UAV 1 provides service to clusters $\mathcal{G}_3$ and $\mathcal{G}_5$, i.e., the subset of CHs $\mathcal{S}_{1h} = \{\mathcal{G}_3, \mathcal{G}_5\}$, whereas UAV 2 provides service to clusters $\mathcal{G}_1$, $\mathcal{G}_2$, and $\mathcal{G}_4$, i.e., the subset CHs $\mathcal{S}_{2h} = \{\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_4\}$. Thus, in this example scenario, the aggregators for clusters $\mathcal{G}_1$, $\mathcal{G}_2$, and $\mathcal{G}_4$ have less packets waiting in their queue to be transmitted, which, in turn, allows UAV 2 to service an extra cluster during decision epoch H than compared to UAV 1. In addition, for each UAV in Fig. 6.1, the path/trajectory is also outlined.

During each decision epoch h, each UAV k has a maximum quota μ_k which

represents the maximum number of CHs that UAV k can serve. The quota μ_k depends on the number of CHs that UAV k can serve, the queue length of each CH, the dwelling time of each CH, and the flying/transition time from one CH to another. The best case quota for each UAV $k \in \mathcal{K}$ is given by, $\mu_k = \lceil \frac{H}{2} \rceil$. This case occurs whenever the dwelling time for each CH is just 1 time slot, and the flying time of UAV k between two aggregators is also 1 time slot. For example, if the decision epoch for all UAVs has a length $H = 9$, then the best case quota for UAV k will be given by, $\lceil \frac{H}{2} \rceil = \lceil \frac{9}{2} \rceil = 5$. Furthermore, Fig. 6.2 outlines an illustrative example of how the decision epoch can be broken down, which is based on the system model example scenario in Fig. 6.1. In particular, during any decision epoch, the following events occur sequentially: Dwelling time, followed by flying time, then dwelling time, and so on, as illustrated in Fig. 6.2.

6.1.1 Uplink Transmission Model

We consider an orthogonal frequency division multiple access (OFDMA) uplink scheme, with a finite set of resource blocks (RBs) each of which having a bandwidth B. One RB is assigned to each CH $g \in \mathcal{G}_g$ to transmit its set of packets to UAV k, when the UAV hovers directly above the CH. The transmit power of CH g during decision epoch h for reliable data transmission over any RB is given by, $p_{gh} \in [0, P_{\max}]$, where $P_{\max}$ is the maximum transmission power. Let $\xi_{gk,t}$ denote the channel gain between CH g and UAV k on each RB during time slot t. The number of packets that each CH g can send to UAV k during the dwelling time $T_{gk,h}$ at decision epoch h, is given by:

$$d_{gk,h} = \frac{\sum_{t \in \mathcal{T}_{gk,h}} \log_2 \left(1 + \frac{p_{gh}\beta\xi_{gk,t}}{\sigma^2}\right) \tau}{D_g} \text{ for } g \in \mathcal{S}_{kh}, \tag{6.2}$$

where $\beta = \frac{-1.5}{\ln(5P_e)}$ is the SNR gap for M-QAM modulation with P_e being the maximum acceptable error probability [200]. In (6.2), D_g is the size of each data packet transmitting from CH g, and σ^2 is the variance of the additive white Gaussian noise.

The channel gain between CH g and UAV k follows a ground-to-air (G2A) channel model, that depends on the type of environment (e.g., rural, urban, suburban, dense urban), location of the CH, and elevation angle between the CH and the UAV [52]. In practical scenarios, the exact location of each UAV and/or CH, as

well as the environment, may be unknown. Thus, the G2A channel model is comprised of a line-of-sight (LoS) link with probability, ρ_{LoS}, and a non-line-of-sight (NLoS) link with probability, $\rho_{\text{NLoS}} = 1 - \rho_{\text{LoS}}$. The typical probability of LoS, ρ_{LoS}, is given by [52] and [30]:

$$\rho_{\text{LoS}} = \frac{1}{1 + \alpha \exp\left(-\varphi(\theta - \alpha)\right)}, \tag{6.3}$$

where α and φ depend on carrier frequency and type of environment, θ is the elevation angle from CH g to UAV k, such that $\theta = \frac{180}{\pi} \sin^{-1}\left(\frac{z_{kt}}{\eta_{gk,t}}\right)$; z_{kt} is the height of UAV k and $\eta_{gk,t}$ is the distance between CH g and UAV k at time slot t, where $\eta_{gk,t} = \sqrt{(x_g - x_{uk})^2 + (y_g - y_{uk})^2 + z_{kt}^2}$. As the UAV will stop directly above the CH, then θ will always be equal to 90°. Because it is hard to determine exactly which portion of the the signal experiences LoS or NLoS during the G2A communication, the average path loss is considered, and is given by [52] and [30]:

$$\overline{L}_{gk,t} = \rho_{\text{LoS}}^{gk,t} \omega_1 \left(\frac{4\pi\eta_{gk,t}}{\lambda}\right)^{\nu} + \rho_{\text{NLoS}}^{gk,t} \omega_2 \left(\frac{4\pi\eta_{gk,t}}{\lambda}\right)^{\nu}, \tag{6.4}$$

where ν is path loss exponent; λ is the wavelength of the electromagnetic wave, i.e., $\frac{f_c}{c}$, where f_c is the carrier frequency and c is the speed of light; and ω_1 and ω_2 are the excessive path loss coefficients for LoS and NLoS, respectively, with $\omega_2 > \omega_1 > 1$. Thus, the average channel gain for the G2A communication link between CH g and UAV k at time slot t, is given by:

$$\overline{\xi}_{gk,t} = \frac{1}{\overline{L}_{gk,t}} = \frac{1}{\left[\omega_1 \rho_{\text{LoS}}^{gk,t} + \omega_2 \rho_{\text{NLoS}}^{gk,t}\right] \left(\frac{4\pi\eta_{gk,t}}{\lambda}\right)^{\nu}}. \tag{6.5}$$

6.1.2 Aggregator Queue Model

In each cluster, a queue of data packets is formed at CH $g \in \mathcal{G}_g$, which contains all of the data packets received from all MTDs within cluster $\mathcal{G}_g$. Let $\boldsymbol{q}_h$ be the vector of G queues at the beginning of decision epoch h with each element q_{gh} being the queue length of CH g at decision epoch h. Let $\widetilde{a}_{gh}$ be the arrival of data packets to CH g during decision epoch h. Due to the stochastic event-driven transmission of MTDs, the arrival of data packets $\widetilde{a}_{gh}$ from MTDs within cluster $\mathcal{G}_g$ to CH g is a positive discrete random variable at each decision epoch. As the the arrival of packets from MTDs within a cluster to the aggregator is a random variable,

the queue length also stochastically changes. During a decision epoch, the queue length of cluster $\mathcal{G}_g$ will change by [201]:

$$q_{g(h+1)} = \max\{q_{gh} - d_{gk,h}, 0\} + \widetilde{a}_{gh} \text{ if } g \in \mathcal{S}_{kh}. \tag{6.6}$$

In our model, the maximal queue length of cluster heads, $Q_h = \max_{g\in\mathcal{G}}\{q_{gh}\}$, reflects the worst-case queuing delay during one decision epoch within the network. As a reliability measure, we leverage the notion of risk where risk is synonymous with losing urgent messages. In our considered M2M network, as the delay (or queue length) becomes higher, this may lead to urgent-messages being lost/dropped. Therefore, to ensure reliable M2M communications, we aim to minimise the *risk of losing messages*. To do this, we use the entropic risk measure $\ln(\mathbb{E}\{e^{\delta Q_h}\})/\delta$ with a risk sensitivity parameter $\delta > 0$ as our reliability metric [42]. Thus, imposing a finite threshold κ on the the entropic risk measure, the M2M network served by UAVs as flying BSs will be *ultra reliable*, if the following requirement is met:

$$\limsup_{h\to\infty} \frac{\ln(\mathbb{E}\{e^{\delta Q_h}\})}{\delta} \leq \kappa. \tag{6.7}$$

The ultra reliability constraint in (6.7), ensures that the risk of the maximum queue length is finite and does not approach infinity. By taking the Maclaurin series expansion, we get $\ln(\mathbb{E}\{e^{\delta Q_h}\}) = \mathbb{E}\{Q_h\} + \frac{\delta}{2}\text{Var}\{Q_h\} + \text{O}\{\delta^2\}$. Next, we focus on the mean and variance of Q_h by considering $0 < \delta \ll 1$.

In this network, the arrival of packets at the aggregator is based on the stochastic event-driven traffic of MTDs. However, the probability model of the event-driven traffic of MTDs is unknown by the UAVs. Thus, each UAV does not know how the queue length of each aggregator will change during the future decision epochs, while the scheduling algorithm of UAVs should guarantee the ultra reliability condition defined in (6.7). Hence, our goal is to design an *energy-aware scheduling scheme guaranteeing ultra reliability*, that makes decisions for UAVs at every decision epoch h, without requiring a-priori knowledge of the probabilities associated with the event-driven arrival requests to the aggregators. In particular, each UAV must determine a subset of CHs that it will serve, while also ensuring sufficient dwelling time is allocated to each CH, in order to collect sufficient packets from CHs to meet the ultra-reliability requirement (6.7). Meanwhile, each CH must determine suitable transmission power required to send the collected data during

the dwelling time, in order to reduce energy consumption. Next, we propose a scheduling framework based on Lyapunov optimisation, as in [201].

6.2 Energy-Aware Scheduling for UAVs and CHs

In this section, we propose a Lyanpunov scheduling algorithm, in order to solve the energy-aware scheduling optimisation problem for UAVs and CHs in an clustered M2M network. We use the Lyapunov optimisation technique because a priori knowledge of the event-driven arrival process for M2M communication is not known. Thus, our goal is to find a scheduling policy for UAV k at the beginning of each decision epoch h, given by, $\pi_{kh} = \{\mathcal{S}_{kh}, \mathcal{T}_{kh}, \mathcal{P}_{kh}\}$, where $\Pi_h = \{\pi_{kh} | \forall k \in \mathcal{K}\}$. In addition, our goal is to also find a suitable transmit power for CH g for uplink transmission to UAV k at the beginning of each decision epoch h to decrease CH energy consumption, which is given by, $\boldsymbol{p}_h$ where each element is the allocated transmit power p_{gh} for CH g. Therefore, at the beginning of each decision epoch h, the scheduling problem for UAVs and CHs can be formulated as follows:

$$\min_{\Pi_h, \boldsymbol{p}_h} \sum_{k \in \mathcal{K}} \sum_{g \in \mathcal{S}_{kh}} \sum_{t \in \mathcal{T}_{gk,h}} p_{gh}\tau, \tag{6.8}$$

s. t.

$$\limsup_{H \to \infty} \frac{1}{H} \sum_{h=0}^{H-1} \sum_{g=1}^{G} \mathbb{E}\{q_{gh}\} < \infty, \tag{6.9}$$

$$\limsup_{H \to \infty} \frac{1}{H} \sum_{h=0}^{H-1} \mathbb{E}\{Q_h\} \leq M_{\text{th}}, \tag{6.10}$$

$$\limsup_{H \to \infty} \frac{1}{H} \sum_{h=0}^{H-1} \mathbb{E}\{(Q_h)^2\} \leq B_{\text{th}}, \tag{6.11}$$

$$\sum_{g \in \mathcal{S}_{kh}} \left(\sum_{t \in \mathcal{T}_{gk,h}} \tau \right) + \sum_{(i,j) \in \mathcal{P}_{kh}} \tau_{ij,k} \leq \tau H, \forall k \in \mathcal{K}, \tag{6.12}$$

$$R_{\text{th}} \leq \log_2 \left(1 + \frac{p_{gh}\beta\overline{\xi}_{gk,t}}{B\sigma^2} \right), \forall k \in \mathcal{K}, \forall g \in \mathcal{S}_{kh}, \forall t \in \mathcal{T}_{gk,h}, \tag{6.13}$$

$$\mathcal{T}_{gk,h} \subseteq \{t+1, ..., t+H\}, \forall g \in \mathcal{G}, \forall k \in \mathcal{K}, \tag{6.14}$$

$$\mathcal{S}_{kh} \cap \mathcal{S}_{wh} = \emptyset, \forall k, w \in \mathcal{K}, \tag{6.15}$$

$$\mathcal{S}_{kh} \subseteq \mathcal{K}, \forall k \in \mathcal{K}, \tag{6.16}$$

$$|\mathcal{S}_{kh}| \leq \mu_k, \forall k \in \mathcal{K}, \tag{6.17}$$

$$\cup_{k\in\mathcal{K}} \mathcal{S}_{k,h} = \mathcal{K}, \tag{6.18}$$

$$0 \leq p_{gh} \leq P_{\max}, \forall g \in \mathcal{G}. \tag{6.19}$$

Here, constraints (6.9), (6.10), and (6.11) are the reliability constraints. Constraint (6.9) ensures that rate stability is guaranteed, i.e., ensuring the queue length of each CH within the network does not get infinitely large. Constraint (6.10) ensures that the average maximum queue length is finitely bounded by threshold M_{th}, whereas constraint (6.11) ensures that the squared average power maximum queue length is finitely bounded by threshold B_{th} where $B_{\text{th}} = M_{\text{th}}^2 + 2\frac{\kappa - M_{\text{th}}}{\delta}$. Constraint (6.13) imposes a rate threshold R_{th} on the link between UAV k and its served CH g. Constraint (6.12) ensures that each UAV can service all CHs in $\mathcal{S}_{kh}$, during the decision epoch. Constraints (6.14) to (6.19) guarantee the feasibility for the decision variables $\mathcal{T}_{gk,h}$, $\mathcal{S}_{kh}$, and p_{gh}.

Due to the stochastic constraints in (6.10), and integer decision variables $\mathcal{T}_{gk,h}$, $\mathcal{S}_{kh}$ and p_{gh}, the optimisation problem in (6.8) is not only stochastic but also combinatorial. As the lengths of the queues in the stochastic constraint (6.10) change with random events, and the probability distribution of these random events is not known in advance, the problem in (6.8) cannot be solved by traditional stochastic optimisation techniques. Next, we propose a tunable minimum-drift-plus-penalty optimisation problem based on Lyapunov optimisation to approximately solve the problem in (6.8), by only knowing the current lengths of queues without knowing any prior knowledge about future random events.

6.2.1 A Tunable Minimum-Drift-Plus-Penalty Optimisation

To solve problem (6.8), we use tools from Lyapunov optimisation theory to dynamically allocate transmit power to the CH and schedule UAVs. In order to ensure (6.10) and (6.11), we respectively introduce two virtual queues which evolve as follows [201]:

$$z_{1(h+1)} = \max\{z_{1h} + Q_h - M_{\text{th}}, 0\}, \tag{6.20}$$

$$z_{2(h+1)} = \max\{z_{2h} + Q_h^2 - B_{\text{th}}, 0\}. \tag{6.21}$$

Let the concatenated vector of all actual and virtual queues be given by, $\Theta_h \triangleq$

$[\boldsymbol{q}_h, Z_{1h}, Z_{2h}]$, and the update equations (6.6), (6.20) and (6.21) be the *state* of the network at the beginning of each decision epoch h. Thus, we assume that, at the end of each decision epoch, all UAVs will synchronise, which means that they will share information about CH queue lengths from previous decision epochs using UAV-to-UAV wireless links (such links can be low-bandwidth control links). Particularly, if a UAV has not served a certain CH previously, it will still be provided information about its queue length from previous decision epochs. Hence, at the beginning of each decision epoch, the state of the network is known to all UAVs.

Following Lyapunov optimisation tools, we define a quadratic Lyapunov function, $L_h = \frac{1}{2}\big(z_{1h}^2 + z_{2h}^2 + \sum_{g\in\mathcal{G}} q_{gh}^2\big)$ [202]. We also consider the total energy consumption of UAVs, $\sum_{k\in\mathcal{K}} \sum_{g\in\mathcal{S}_{kh}} \sum_{t\in\mathcal{T}_{gk,h}} p_{gh}\tau$, as an associated penalty, and the parameter v for tuning the penalty at the beginning of decision epoch h. Then, we can define the *one-epoch conditional Lyapunov drift plus penalty* at the decision epoch $h+1$ with penalty parameter v, as follows:

$$\Delta_{vh} = \mathbb{E}\left\{L_{h+1} - L_h | \boldsymbol{q}_h\right\} + v \sum_{k\in\mathcal{K}} \sum_{g\in\mathcal{S}_{kh}} \sum_{t\in\mathcal{T}_{gk,h}} p_{gh}\tau, \tag{6.22}$$

where $\mathbb{E}\left\{L_{h+1} - L_h | \boldsymbol{q}_h\right\}$ is a one-epoch conditional Lyapunov drift. Here, the expectation is taken over the randomness of arrival requests of event-driven traffic to the queues.

Lemma 6.1 Under any feasible scheduling algorithm Π_h that satisfies the ultra reliability constraint in (6.7) at decision epoch h, the one-epoch conditional Lyapunov drift-plus-penalty is bounded as follows:

$$\begin{aligned}\Delta_{vh} \leq & I + \sum_{g=1}^{G} q_{gh}(a_{gh} - d_{gh}) + z_{1h}(Q_h - M_{th}) + \\ & z_{2h}(Q_h^2 - B_{th}) + v \sum_{k\in\mathcal{K}} \sum_{g\in\mathcal{S}_{kh}} \sum_{t\in\mathcal{T}_{gk,h}} p_{gh}\tau,\end{aligned} \tag{6.23}$$

where $I = \frac{H^2 + (\max_g |\mathcal{G}_g|)^2 + M_{\text{th}}^2 + B_{\text{th}}^2}{2}$ is a constant value, and a_{gh} is the arrival of data packets from CH g during current decision epoch h and is not a random variable any more. *Proof:* See Appendix D.1. ■ Following [201, Theorem 4.2], all queues $\boldsymbol{q}_h$ are mean rate stable if for all decision epochs h and all possible values of $\boldsymbol{q}_h$, the one-epoch conditional Lyapunov drift-plus-penalty is bounded. Thus, the result in Lemma 1 suggests the following scheduling decision: at every decision epoch

h, observe the current state Θ_h and take a control action that greedily minimises the left-hand-side of the desired one-epoch conditional Lyapunov drift-plus-penalty inequality in (6.23). The proposed drift-plus-penalty minimisation problem aims to minimise the one-epoch conditional Lyapunov drift-plus-penalty, which approximately finds a solution for the joint scheduling and energy optimisation problem given in (6.8). Instead of minimising the one-epoch conditional Lyapunov drift-plus-penalty, one can minimise the maximum bound of the one-epoch conditional Lyapunov drift plus penalty, which is the right-hand side of the inequality in (6.23). Hence, the UAVs observe the current state Θ_h of all clusters, and make a scheduling decision Π_h during each decision epoch h. Meanwhile, the CHs determine the transmit power p_{gh} required to successfully transmit the packets to their serving UAV during decision epoch h. Thus, the following optimisation problem is solved for each decision epoch h:

$$\max_{\Pi_h, \boldsymbol{p}_h} \sum_{g=1}^{G} q_{gh} d_{gh} - z_{1h} Q_h - z_{2h} Q_h^2 - v \sum_{k \in \mathcal{K}} \sum_{g \in \mathcal{S}_{kh}} \sum_{t \in \mathcal{T}_{gk,h}} p_{gh} \tau, \tag{6.24}$$

s. t.:

$$R_{\text{th}} \leq \log_2 \left(1 + \frac{p_{gh} \beta \overline{\xi}_{gk,t}}{B \sigma^2} \right), \forall k \in \mathcal{K}, \forall g \in \mathcal{S}_{kh}, \forall t \in \mathcal{T}_{gk,h}, \tag{6.25}$$

$$\sum_{g \in \mathcal{S}_{kh}} \left(\sum_{t \in \mathcal{T}_{gk,h}} \tau \right) + \sum_{(i,j) \in \mathcal{P}_{kh}} \tau_{ij,k} \leq \tau H, \forall k \in \mathcal{K}, \tag{6.26}$$

$$\mathcal{T}_{gk,h} \subseteq \{t+1, ..., t+H\}, \forall g \in \mathcal{G}, \forall k \in \mathcal{K}, \tag{6.27}$$

$$\mathcal{S}_{kh} \cap \mathcal{S}_{wh} = \emptyset, \forall k, w \in \mathcal{K} \tag{6.28}$$

$$\mathcal{S}_{kh} \subseteq \mathcal{K}, \forall k \in \mathcal{K}, \tag{6.29}$$

$$|\mathcal{S}_{kh}| \leq \mu_k, \forall k \in \mathcal{K}, \tag{6.30}$$

$$\cup_{k \in \mathcal{K}} \mathcal{S}_{kh} = \mathcal{K}, \tag{6.31}$$

$$0 \leq p_{gh} \leq P_{\max}, \forall g \in \mathcal{G}. \tag{6.32}$$

At the end of each decision epoch h, the virtual queues z_{1h} and z_{2h} will be updated according to (6.20) and (6.21), respectively, and CH queue length q_{gh} will be updated according to (6.6). Optimisation problem (6.24) has two decision variables. The decision variables π_{kh} in (6.24) correspond to the scheduling optimisation problem of UAVs, whereas the decision variables p_{gh} in (6.24) corresponds to

the energy optimisation problem of CHs. In (6.24), if the tunable parameter $v = 0$, then the optimisation problem in (6.24) will focus more on the UAV scheduling optimisation problem, and the energy of CHs will no longer be a constraint. However, if $v = \infty$, the optimisation problem in (6.24) will focus more on the CH energy optimisation problem, that is sending the maximum number of packets from the CH to the UAV.

Lyapunov optimisation has been used, as the transmission probability of the CH queues is unknown. Due to using Lyapunov optimisation, the tunable drift-plus-penalty optimisation problem in (6.24) is now a linear weighted function, and the constraints are no longer a function of the random variable queue length, as in (6.10). However, any centralised solution solving the tunable drift-plus-penalty optimisation problem in (6.24) will be NP-hard and complex, due to the combinatorial nature of the problem [62]. Next, we propose a distributed energy-aware scheduling scheme for UAVs and CHs using *matching theory* [202], for the clustered IoT network. Hence, our goal is to obtain a scheduling strategy in which each UAV is associated with the most preferred CHs while taking into account the maximum queue length, and also obtain an energy strategy in which CHs can minimise their energy consumption during the dwelling time.

6.3 One-to-Many Matching for Ultra-Reliable Energy-Aware Scheduling

We propose a novel distributed energy-aware scheduling algorithm guaranteeing ultra reliability based on matching theory. Matching theory is a powerful mathematical tool that is decentralized [74]. Here, matching theory is used to formally model the tunable drift-plus-penalty optimisation problem in (6.24), as a *one-to-many matching game* using *costs*. Thus, the one-to-many matching game models the interactions between a number of CHs desiring to be served by one UAV. Hence, the game considers two disjoint sets of players, i.e., a CH set $\mathcal{A}$ and a UAV set $\mathcal{K}$, where players from each set have ranked preferences over players in the other set, and vice versa. In the context of the studied network, each UAV $k \in \mathcal{K}$ will be matched to a subset of CHs $\mathcal{S}_{kh} \subseteq \mathcal{A}_{kh}$, $\mathcal{A}_{kh} \subseteq \mathcal{A}$, while each CH $g \in \mathcal{A}_{kh}$ will be matched to a single UAV $k \in \mathcal{K}$. The proposed one-to-many matching game is defined as follows:

Definition 6.1 Given two disjoint finite sets of UAVs $\mathcal{K}$ and aggregators $\mathcal{A}$, the ultra reliability and energy-aware scheduling algorithm for UAVs in a clustered IoT networks, can be defined as a *one-to-many matching function* $\Phi : \mathcal{K} \cup \mathcal{A} \rightarrow \mathcal{K} \cup \mathcal{A}$ such that for all $k \in \mathcal{K}$ and $g \in \mathcal{A}$:

1. $\Phi(k) = \mathcal{S}_{kh} \subseteq \mathcal{A}$ and $|\Phi(k)| \leq \mu_k$,
2. $\Phi(g) \in \mathcal{K}$ and $|\Phi(g)| \in \{0, 1\}$,
3. $\Phi(g) = k$, if and only if $g \in \Phi(k)$.

From Definition 6.1, the first condition states that each UAV k can be matched to a subset of CHs $\mathcal{S}_{kh}$, with μ_k being the quota of UAV k, which is the maximum number of CHs that UAV k can serve, and thus, $|\Phi(\cdot)|$ denotes the set cardinality. Depending on the CH subset that UAV k will serve, the quota will not always be the same, and thus, is not fixed across all UAVs and decision epochs. The second condition states that each CH $g \in \mathcal{A}_{kh}$ can be matched to at most one UAV $k \in \mathcal{K}$, i.e., the quota for each CH is set to 1. The fourth condition states that if UAV k is matched to CH g, then CH g is also matched with UAV k. In the case that $\Phi(g) = g$, then CH g is matched to itself, which means that CH g is not assigned/matched to any UAV.

6.3.1 Utility Functions and Preference Relations of UAVs and CHs

The one-to-many matching game can be defined by $\mathfrak{G} = \{\mathcal{K}, \mathcal{A}, \succ_k, \succ_g\}$, where $\succ_k$ and $\succ_g$ refer to the *preference relations* for UAVs and CHs, respectively. A preference relation a defined as a complete, reflexive and transitive binary relation between the players in the opposite set. Thus, the preference relation for a UAV $\succ_k$ depends on the subset of CHs it wishes to serve, and the queue length of each CH to be served. UAV $k \in \mathcal{K}$ serving $\mathcal{S}$ is preferred to serving $\mathcal{S}'$, which is denoted as $\mathcal{S} \succ_k \mathcal{S}'$. On the other hand, a preference relation for a CH $\succ_g$ depends on the amount of transmission power required to send the data to a UAV, as well as the duration of dwelling time the UAV will collect the data. CH g prefers to be served by UAV k over UAV k', which is defined as $k \succ_g k'$. To derive the preference relations for the two disjoint sets of players, we propose individual utility functions for each set based on the optimisation problem in (6.24).

To enable ultra reliable communications in an IoT network, each UAV k aims to select a subset of CHs $\mathcal{S}_{kh} \subseteq \mathcal{A}_{kh}$ that will maximise the number of packets sent from each CH $g \in \mathcal{S}_{kh}$ to UAV k, while also guaranteeing that the maximum queue length of all CHs are finitely bounded. The proposed utility function for UAVs is a linear weighted function, where the weights are the queue lengths of the served CHs, and the virtual queue lengths. Given matching Φ and the transmit power vector $\boldsymbol{p}_h$ of all CHs, we define the utility function of UAV k serving a subset of CHs $\mathcal{S}_{kh}$ at decision epoch h, as follows:

$$U_k(\mathcal{S}_{kh}, \Phi) = \sum_{g \in \mathcal{S}_{kh}} q_{gh} d_{gk,h} - z_{1h} Q_h - z_{2h} Q_h^2, \tag{6.33}$$

where z_{1h} and z_{2h} are given by (6.20) and (6.21) respectively; and $d_{gk,h}$ is given by (6.2) which is a function of the dwelling time $\mathcal{T}_{gk,h}$ and the CH transmit power p_{gh}. In (6.33), as the queue length of CH g increases, the UAV serving this CH will allocate more dwelling time to this CH, in order to reduce its queue length and guarantee rate stability and ultra reliability.

Meanwhile, each CH g aims to maximise its transmission packet rate to UAV k, while minimising its total energy consumption. Given matching Φ we define the utility function of CH $g \in \mathcal{S}_{k,h}$ being served by UAV k at decision epoch h, as follows:

$$U_g(k, \Phi) = \sum_{t \in \mathcal{T}_{gk,h}} \left(\frac{\log_2 \left(1 + \frac{p_{gh} \beta \overline{\xi}_{gk,t}}{B\sigma^2}\right)}{D_g} - v p_{gh} \right) \tau. \tag{6.34}$$

Based on the utility function in (6.34), each CH g aims to minimise its transmission power based on the following optimisation problem:

$$\max_{p_{gh}} \Big(\sum_{t \in \mathcal{T}_{gk,h}} \left(\frac{\log_2 \left(1 + \frac{p_{gh} \beta \overline{\xi}_{gk,t}}{B\sigma^2}\right)}{D_g} - v p_{gh} \right) \tau \Big), \tag{6.35}$$

Subjected to:

$$R_{\text{th}} \leq \log_2 \left(1 + \frac{p_{gh} \beta \overline{\xi}_{gk,t}}{B\sigma^2}\right), t \in \mathcal{T}_{gk,h}, \tag{6.36}$$

$$0 \leq p_{gh} \leq P_{\max}. \tag{6.37}$$

Using the utility functions in (6.33) and (6.34), the preference relation of UAVs

and CHs at decision epoch h is given by:

$$\mathcal{S}_{kh} \succ_k \mathcal{S}'_{kh} \Leftrightarrow U_k(\mathcal{S}_{kh}, \Phi) > U_k(\mathcal{S}'_{kh}, \Phi), \tag{6.38}$$

$$k \succ_g k' \Leftrightarrow U_g(k, \Phi) \quad > U_g(k', \Phi), \tag{6.39}$$

for all $\mathcal{S}_{kh}, \mathcal{S}'_{kh} \subseteq \mathcal{A}$ and $k, k' \in \mathcal{K}$. The proposed matching game considered strict preference relations for each UAV and CH. Thus, from (6.38) each UAV ranks a subset of CHs by giving a preference to the set of CHs that provides the highest utility. Moreover, each CH can use (6.39) to rank the UAVs by giving a preference to a UAV that provides the highest utility.

6.3.2 Stable Matching Analysis of Distributed Ultra Reliable and Energy-Aware Scheduling Algorithm

To solve the proposed problem in (6.24) as a one-to-many matching game, we use the concept of *two-sided stable matching* to find a suitable matching between UAVs and CHs [203]. A two-sided stable matching for the proposed game is defined as [203]:

Definition 6.2 A UAV and CH pair $(k, g) \notin \Phi$ is said to be a *blocking pair* of the matching Φ, if and only if $k \succ_g \Phi(g)$ and $g \succ_k \Phi(k)$. A matching Φ is *stable*, if there is no blocking pair.

Hence, under a stable matching Φ, the set of UAVs can guarantee that the CH queue lengths will be bounded by a finite value, that is, satisfying the ultra-reliable and low-latency constraints for mission critical IoT communications. In order to solve a one-to-many matching problem, most of the prior works [75,204] have relied on the so-called deferred acceptance (DA) algorithm [203, 205]. However, in our proposed one-to-many matching game, we cannot directly apply the DA algorithm. This is due to the UAV utility being a function of the maximum queue length across all CHs in the network Q_h, i.e., Q_h is a function of all CH queues, and thus, is a dependant variable. Hence, each UAV not only considers which set of CHs to match with, but also that the set of CHs satisfy their ultra reliability constraint. In addition, the quota of each UAV changes in each decision epoch, where the DA algorithm assumes a fixed quota all the time. Therefore, our proposed one-to-many matching game can be defined as a *one-to-many matching game with*

externalities [153, 206–210]. Due to the externalities in the proposed matching game, the DA algorithm cannot be used to converge to a two-sided stable matching. Hence, we propose a new matching algorithm to solve the one-to-many matching game with externalities, and use the concept of *two-sided-stable matching* as defined in Definition 6.2. At the stable convergence of the one-to-many matching game with externalities, no UAV or CH will have incentive to change from their current matching, without causing a decrease in their utility function, respectively, i.e., a network-wide stability.

6.4 Proposed Algorithm Description

The proposed algorithm for solving the distributed ultra reliable and energy-aware scheduling problem during each decision epoch h, is shown in Algorithm 6. The algorithm is divided into three phases: a) initializing decision epoch h; b) matching evaluation; and c) CH power allocation. At the beginning of each decision epoch h, each UAV k will be located above a CH g, from the end of the previous decision epoch $h-1$, and thus, the initial location of UAV k is known. In Phase I, each UAV shares the queue lengths with all other UAVs in the set $\mathcal{K}$. Thus, based on this sharing of information, all UAVs can determine a consensus on the maximal queue length Q_h.

Phase II is the matching evaluation for decision epoch h. Each UAV determines its strategy set, that is, finding all possible combinations of CH subsets and dwelling times. Note that, since each UAV k has a limited coverage area, the subset of CHs that UAV k can propose to is smaller than the set of CHs, i.e., $\mathcal{A}_{kh} \subset \mathcal{A}$. Hence, the strategy set for UAV k is given by $\boldsymbol{s}_{kh} = [T_{gk,h}]_{G\times 1}$, where $T_{gk,h}$ is an element of $\boldsymbol{s}_{kh}$, such that $T_{gk,h}$ is the proposed dwelling time for CH $g \in \mathcal{A}_{kh}$ from UAV k. If $T_{gk,h} = 0$, this means UAV k does not stop over CH g to collect data from it. Thus, the set of all possible strategy set combinations for UAV k is denoted by Γ_{kh}, where $\boldsymbol{s}_{kh} \in \Gamma_{kh}$. Then, each UAV k calculates its utility for each possible strategy set within Γ_{kh}, using (6.33). Based on (6.38), UAV k finds the most preferred CH subset it wishes to serve, and proposes that CH subset while rejecting the other strategies. Then, the CHs determine whether to accept or reject the UAV proposal/s. The strategy for CH g is to maximise its packet rate, while minimising its transmission power based on the proposed payment from the currently matched

Algorithm 6 Matching Game for Ultra Reliable and Energy-Aware Scheduling in IoT Network

Input: Set of Aggregators: $\mathcal{A}$; Location of aggregators $\boldsymbol{w}_g$ $\forall g \in \mathcal{A}$; Set of UAVs: $\mathcal{K}$; UAV location $\boldsymbol{f}_{kt}$ at the beginning of each decision epoch; Queue values $q_{g(h-1)}$ for all $g \in \mathcal{A}$, and virtual queue values $z_{1(h-1)}$ and $z_{2(h-1)}$, from previous decision epoch, $h-1$;
Phase I - Initializing Decision Epoch h
Shares the queue lengths and also payments of the served CHs during the previous decision epoch to all other UAVs within $\mathcal{K}$;
Calculates the maximal queue length, Q_h;
Phase II - Matching Evaluation
for each $k \in \mathcal{K}$ **do**
    Determines all possible strategy set vectors, $\boldsymbol{s}_{kh}$;
    Determine the utility for each strategy set vector, using (6.33);
    Propose to the most preferred subset of CHs, based on $\succ_k$, i.e., the subset of CHs with the highest rank;
end for
if CH g receives a proposal/s **then**
    CH g calculates its utility for all proposals, using (6.34);
    Based on $\succ_g$, CH g accepts one UAV's proposal and rejects the rest;
end if
Phase III - CH Power Allocation
Given the matching Φ, each CH $g \in \mathcal{A}$ determines its transmission power p_{gh}, based on (6.35);
Output: Matching and payments at decision epoch h: $\Phi(h)$; Scheduling policy of all UAVs: Π_h; Allocated transmit power of all Aggregators: $\boldsymbol{p}_h$.

UAV. For each proposal, the CH calculates its utility, using (6.34), and accepts the most preferred proposal, based on (6.39).

Finally, in Phase III, given the matching Φ, each CH will determine its power allocation, given (6.35). These phases are repeated at each decision epoch, until we reach a stable matching, or stage in which no UAV wishes to propose to any other set of CHs, or no CH wishes to be served by any other UAV.

6.5 Stability, Convergence, Optimality, and Computational Cost of the Proposed Algorithm

In the following, we analyze the performance of our proposed one-to-many matching game and power allocation, in Algorithm 6. More specifically, we analyze the stability, convergence, computational cost, and optimally of the proposed algorithm.

6.5.1 Stability, Convergence, and Optimality

Here, we prove that the proposed Algorithm 6 converges to a two-sided stable matching, based on Definition 6.2.

Theorem 6.1 The proposed one-to-many matching game, from Algorithm 6, yields a two-sided stable matching between UAVs and CHs.

Proof: See Appendix D.2. ■

Proposition 6.1 The one-to-many matching algorithm reaches a global optimum of the total utility (sum utility across all UAVs and CHs), which corresponds to a two-sided stable matching.

Theorem 6.2 The proposed Algorithm 6, is guaranteed to converge to a stable matching and power allocation.

Proof: See Appendix D.3. ■

6.5.2 Computational Cost

To find the computational cost of the proposed Algorithm 6, we will focus on the matching evaluation phase (Phase II) and the CH power allocation phase (Phase III). In the matching evaluation phase, the complexity mainly lies in the process of finding all CH subsets for each UAV, as well as the corresponding dwelling times for each subset. However, since the UAVs have a limited coverage area, this will significantly reduce the computational cost in finding the CH subsets. Initially, each UAV is located above a CH, so regardless of all other strategy set combinations, there will be at least one strategy for each UAV. Thus, the computational cost for finding all possible strategy set vectors $\boldsymbol{s}_{kh}$ for all UAVs in each decision epoch h is given by $O(K(H+1)^A\mu)$. As the number of CHs per subset increases, this will cause the complexity to exponentially increase. In particular, in order to keep the number of proposals generally low, the quota for each UAV k μ_k needs to be restricted to a small finite number. In the matching evaluation phase, we also need to consider the proposal attempts for each CH during each decision epoch, which has a computational cost of $O(A)$. On the other hand, the computational cost of the CH power allocation phase, is given by $O(A)$.

Therefore, the overall computational cost of the proposal Algorithm 6, for the duration of the game T, is given by $O(\frac{H}{T}KA(H+1)^A)$.

Table 6.1: Simulation Parameters.

Parameter	**Value**
Carrier frequency (f_c)	2 GHz
Maximum transmit power of aggregators ($P_{\max}$)	30 dBm
Transmission power vector (P_{vec})	$[0-30]$ dBm
Minimum transmission rate of aggregators (R_{th})	3 dB
Transmission bandwidth per aggregator (B)	15 KHz
Duration of time slot t (τ)	1 second
Decision epoch length (H)	150 seconds
Length of each data packet (D_g)	50 bits
Maximum acceptable error probability (P_e)	10^{-8}
Noise power (σ^2)	-115 dBm
Path loss exponent for LoS links (ν)	2
Additional path loss to free space for LoS (ω_1)	3 dB
Additional path loss to free space for NLos (ω_2)	23 dB
UAV speed (v_k)	10 m/s
M_{th}	20
B_{th}	450

6.6 Simulation Results

For our simulations, we consider an IoT system, where MTDs are located in a geographical area with a 600 m radius. In this area, the set of MTDs form $|\mathcal{G}|$ disjoint clusters. In each cluster, an aggregator is designated as the cluster head. The CHs are uniformly distributed across all clusters. We consider a total number 4 aggregators within the considered geographical area. Moreover, we consider UAV-based communications, that act as low flying aerial BSs in an urban environment, with $\alpha = 11.95$ and $\varphi = 0.14$ for a 2 GHz carrier frequency [52]. Initially, the set of UAVs, $K = 2$, are randomly deployed above one CH within the network. Moreover, due to flight regulations and environmental obstacles, we assume that the altitude of the UAVs vary between 400 m and 600 m. Here, we model the MTD traffic based on a binomial distribution, where the probability of transmission is denoted as ρ_{tx}. Table 6.1 lists the rest of the simulation parameters. The matching game is evaluated using Monte Carlo simulations with varying transmission probabilities ρ_{tx}, queue length thresholds M_{th} and B_{th}, and the number of UAVs K. We compare our simulation results with pre-deployed stationary UAV BSs, which adopts the CH power control scheme from Algorithm 6. In the stationary case, all UAVs are assumed to be fixed above a CH.

Fig. 6.3 and 6.4 show different properties of the maximum queue length across

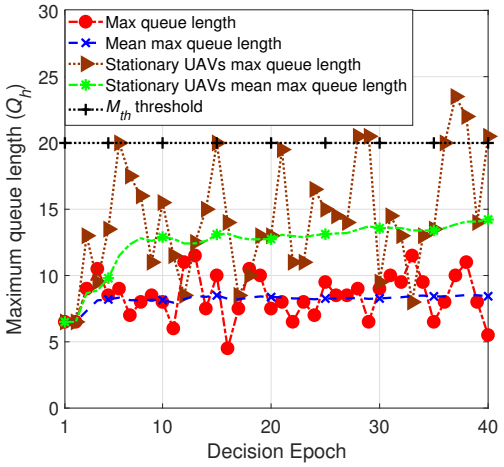


Figure 6.3: Maximum queue length, where $\rho_{\text{tx}} = 0.5$, $K = 2$ and $[M_{\text{th}}, B_{\text{th}}] = [20, 450]$.

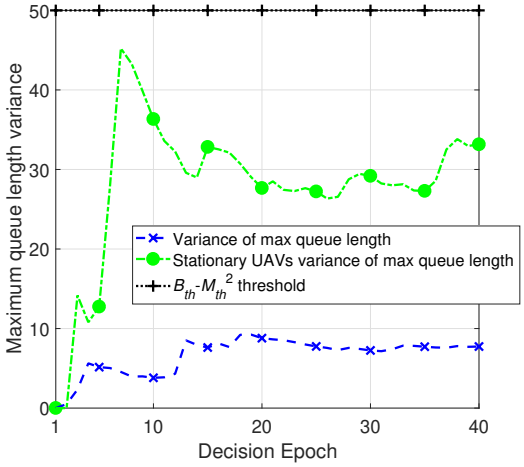


Figure 6.4: Maximum queue length variance, where $\rho_{\text{tx}} = 0.5$, $K = 2$ and $[M_{\text{th}}, B_{\text{th}}] = [20, 450]$.

all aggregators in the network, for a transmission probability of $\rho_{\text{tx}} = 0.5$, a number of UAVs $K = 2$, and a queue length threshold pair of $[M_{\text{th}}, B_{\text{th}}] = [20, 450]$. From these figures, we can see that the maximum queue length is bounded by a mean threshold M_{th}, as well as a variance threshold $B_{\text{th}} - M_{\text{th}}^2$, which means that the number of deployed UAVs is able to guarantee URLLC for the mission critical M2M traffic. In Fig. 6.3 and 6.4, we observe that, over the decision epochs, the mean and variance of the maximum queue length converge to stable outcomes, with a mean maximum queue length of 7 and a maximum queue length variance of 5 across all aggregators, for the proposed energy-aware scheduling algorithm. Also in Fig. 6.3 and 6.4, the proposed energy-aware scheduling algorithm reduced mean maximum queue length by around 37 % and maximum queue length variance by around 73 %, compared to the stationary UAV approach. Having a low variance for maximum queue length, indicates that the number of UAVs within the network is able to service all CHs sufficiently, without causing any one CH's queue length to increase significantly.

In Fig. 6.5, we show the minimum number of required UAVs to service the set of CHs within the network, under different transmission probabilities, ρ_{tx}, and queue length threshold pairs, $[M_{\text{th}}, B_{\text{th}}]$. We observe that, as the transmission probabilities increase, the minimum number of required UAVs to service the CHs also increases. This is due to the fact that increasing the transmission probabilities, will

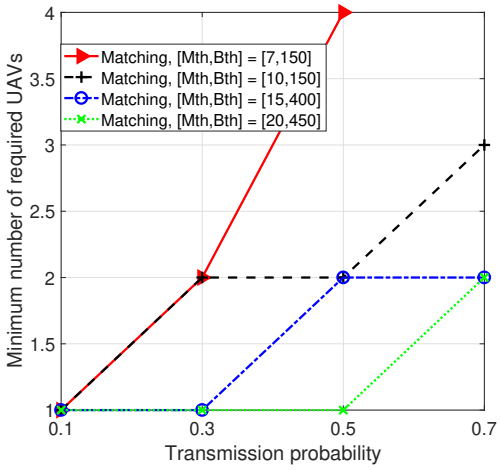


Figure 6.5: Minimum number of UAVs required to service the set $\mathcal{A}$ of aggregators as a function of ρ_{tx} and $[M_{\text{th}}, B_{\text{th}}]$.

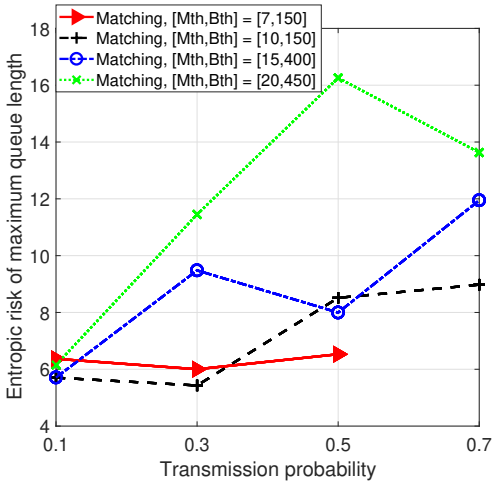


Figure 6.6: Average entropic risk for a minimum number of UAVs as a function of ρ_{tx} and $[M_{\text{th}}, B_{\text{th}}]$.

cause more MTDs within each cluster to transmit data packets to an aggregator. As illustrated in Fig. 6.5, as the mean and variance thresholds of the maximum queue length increases, fewer UAVs are required to service the set of aggregators in the network. In particular, for queue length thresholds $[M_{\text{th}}, B_{\text{th}}] = [7, 150]$ more than 4 UAVs will be needed to service the set of aggregators. This means that more UAVs need to be deployed to service all the MTDs and maintain the mission critical communication constraints, where the number of UAVs in the network will exceed the number of CHs, i.e., $K > A$.

Fig. 6.6 shows the average entropic risk of the maximum queue length, under different transmission probabilities, ρ_{tx}, and queue length threshold pairs, $[M_{\text{th}}, B_{\text{th}}]$. The average entropic risk of the maximum queue length is determined using: $\ln(\mathbb{E}\{e^{\delta Q_h}\}) = \mathbb{E}\{Q_h\} + \frac{\delta}{2}\text{Var}\{Q_h\} + \text{O}\{\delta^2\}$, for the minimum number of required UAVs to service the set of aggregators (outlined in Fig. 6.5). We observe that, as the mean and variance threshold of the maximum queue length increases, this causes the entropic risk to also increase, due to allowing larger maximum queue lengths across all aggregators. In fact, as the transmission probabilities increase, so does the average entropic risk, due to higher thresholds and less UAVs required to service all aggregators in the network. On average, in Fig. 6.6, the proposed energy-aware scheduling algorithm increases the average entropic risk by 65 %, when increasing the queue length thresholds from $[M_{\text{th}}, B_{\text{th}}] = [10, 150]$ to $[M_{\text{th}}, B_{\text{th}}] = [20, 450]$.

Fig. 6.7 shows the average transmit power per aggregator, for different tunable

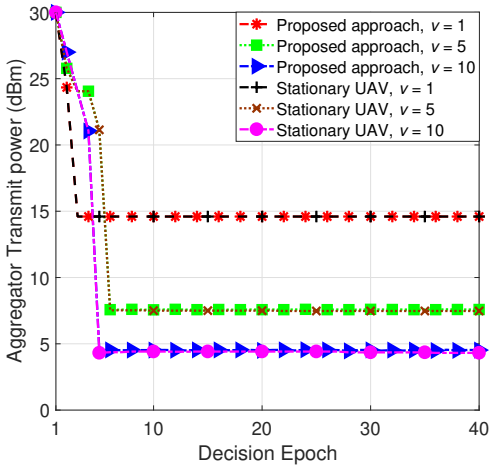


Figure 6.7: Average aggregator transmit power as a function of tunable parameter v, where $\rho_{\text{tx}} = 0.5$, $K = 3$ and $[M_{\text{th}}, B_{\text{th}}] = [10, 150]$.

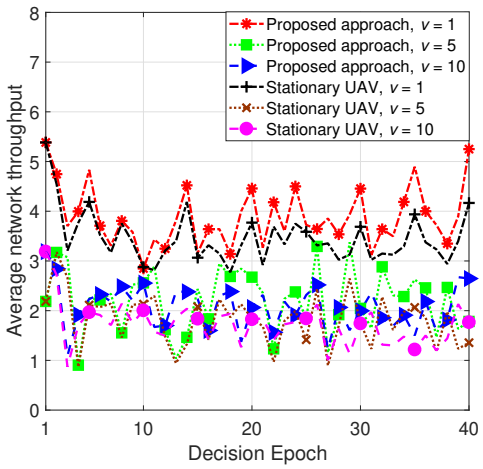


Figure 6.8: Average network throughput as a function of tunable parameter v, where $\rho_{\text{tx}} = 0.5$, $K = 3$ and $[M_{\text{th}}, B_{\text{th}}] = [10, 150]$.

parameters, v, where $\rho_{\text{tx}} = 0.5$, $K = 2$, and $[M_{\text{th}}, B_{\text{th}}] = [20, 450]$. We observe that, over the decision epochs, the average aggregator transmission power reduces and converges to a stable outcome. In addition, we can observe that any increase in the tunable parameter v will lead to a decrease in the average transmission power per aggregator, due to putting more emphasis on minimising transmission power as defined in the utility function, (6.34). The proposed energy-aware scheduling algorithm converges at decision epoch $h = 3$ for $v = 1$, $h = 6$ for $v = 5$, and $h = 5$ for $v = 10$. Over time, the proposed energy-aware scheduling algorithm converges to an average aggregator transmit power of 14.6 dBm for $v = 1$, 7.58 dBm for $v = 5$, and 4.52 dBm for $v = 10$. Also observed in Fig. 6.7, the stationary UAV approach also converges to a similar transmit power, due to using the same algorithm as the proposed approach.

In Fig. 6.8, we show the average network throughput, i.e., the amount of packets transmitted from an aggregator to a UAV during a dwelling time on average, under different tunable parameters, v, where $\rho_{\text{tx}} = 0.5$, $K = 2$, and $[M_{\text{th}}, B_{\text{th}}] = [20, 450]$. In this figure, we see that any increase in the tunable parameter v will lead to a decrease in network throughput, due to reducing transmission power per aggregator (as outlined in Fig. 6.7). Hence, there is a tradeoff between maximising network throughput and minimising transmission power, which can be observed in the aggregator utility function, (6.34). Moreover, we observe that there is a large oscillation in the average network throughput, which is due to the additional pack-

ets arriving at the CH at each decision epoch, and thus, causing the queue length of each aggregator in the network to vary. Over time, the proposed energy-aware scheduling algorithm increases average network throughput by around 14.14 % for $v = 1$, 26.78 % for $v = 5$, and 23.7 % for $v = 10$, compared to the stationary UAV approach.

6.7 Summary of Contributions

In this chapter, we have proposed a novel distributed energy-aware scheduling scheme for mission critical M2M communications, utilizing UAVs as flying BSs. In the proposed model, UAVs schedule themselves in order to collect data from each aggregator in a distributed manner, without knowing the a-priori knowledge of the probabilities associated with the event driven arrival requests from MTDs. Whereas, the aggregators determine suitable transmission power, in order to decrease energy consumption and maximise the number of transmitted packets. Hence, in this chapter we have answered question Q8, posed in Section 1.2.

We have modelled the problem using Lyapunov optimisation and a one-to-many matching game. Lyapunov optimisation has been used to transform the problem into a linear weighted function, where maximum queue length and variance across all aggregators is bounded within the network, in order to satisfy the ultra-reliable and low-latency constraints of mission critical M2M communications. Based on the one-to-many matching game, we have proposed a distributed energy-aware scheduling algorithm, in which UAVs seek to serve a subset of aggregators in order to satisfy the URLLC constraint. We have shown that the proposed distributed energy-aware algorithm converges to a stable matching that yields a two-sided stable matching, as well as that transmission power is minimised at a stable outcome. Simulations results have shown that the proposed distributed algorithm guarantees that the maximum queue length mean and variance are bounded by finite threshold values. The results also demonstrated that aggregator transmit power is decreased further when increasing the tunable parameter from the Lyapunov optimisation.

Chapter 7

Conclusions and Future Research Directions

This thesis has investigated how resource allocation in multimedia IoT applications can be optimised, such as for reliability, scalability, energy-efficiency, and quality-of-experience, in a distributed and autonomous manner. In this chapter, we draw some general conclusions from this thesis. We also outline some future research directions arising from this work.

The first half of this thesis studied interference management, resource allocation, and end-user satisfaction challenges that arise in underlaid D2D communications. Chapter 3 addressed the fundamental problem of flexibility for application-driven distributed resource allocation, using a dynamic multiple-objective (equivalently multi-criteria) Stackelberg game. The proposed non-scalarised solution reduced the real-time information exchange between D2D pairs and transmission power, and increased network throughput, while guaranteeing best social welfare and satisfaction across all D2D users. Moreover, such reduction in transmission power per D2D pair leads to reduced intra-cell interference, which, in turn, increases battery lifetime per D2D pair, and improves D2D user QoE and perceived benefit. Furthermore, prioritising communications for certain D2D applications will also lead to an increase in communication reliability and reduced latency.

In Chapter 4 we jointly optimised D2D mode selection, resource allocation, and power control in a distributed and autonomous manner, using non-cooperative and coalitional game theoretic approaches. Both approaches enabled D2D pairs to select optimal transmission mode, with the assistance of the BS, which essentially enables D2D pairs to operate more efficiently and better utilize existing cellular resources. In fact, the coalitional game enabled D2D pairs to cooperate and share

information, which in turn increased system capacity and throughput, while also reducing individual transmission power in order to reduce intra-cell interference and prolong battery lifetime. Moreover, importantly, as the network begins to evolve towards multimedia IoT, users will start to become more self-sufficient and intelligent. Thus, we further investigated a user-assisted D2D mode selection scheme as a viable solution, in which, the BS has partial knowledge of the network. In practice this approach is ideal, as it requires less signalling overhead and reduced network complexity, where extra resources within the network can be used to further enhance and improve existing communications within the network.

The second half of this thesis studied M2M communications, which is broken down into two categories, massive M2M communications and ultra-reliable low latency communications (URLLC). Chapter 5 addressed the problem of fully distributed, self-organising, correlation-aware clustering for a massive number of MTDs densely deployed in a cellular network. The problem was modelled using evolutionary game theory and stochastic geometry. We have shown that the proposed distributed algorithm converges to an ESS cluster formation, which is robust to a small portion of deviating MTDs due to stochastic changes in the M2M environment. Simulation results showed that the proposed evolutionary algorithm decreased transmission power per cluster, and increased the number of redundant bits that can be eliminated in a given cluster. Hence, by eliminating more redundant bits per cluster this can enable a reduction in the size of the transmitted packets to the BS, which in turn improves overall system efficiency.

Chapter 6 addressed the problem of distributed energy-aware scheduling for mission critical M2M communications, utilizing UAVs as flying BSs. The problem was modelled using a one-to-many matching game with externalities and Lyapunov optimisation. We have shown that the proposed distributed algorithm converges to a stable matching and power allocation, and found the minimum number of required UAVs to service a set of aggregators under different MTD traffic scenarios. Simulations results showed that the maximum queue length mean and variance were bounded by finite threshold values, which means that the number of UAVs is able to service the set of aggregators within a particular geographical area and satisfy the URLLC constraints. In addition, aggregator transmission power is reduced, which means that aggregators do not need to be recharged as frequently in remote locations.

7.1 Future Research

The research area of multimedia IoT is an important field with great potential. Following from the work presented in this thesis, there are a number of potential future research directions:

Multi-cell D2D Communications: In Chapters 3 and 4 of this thesis, we considered a single cell scenario when optimising mode selection, resource allocation, interference management, and end-user satisfaction. In this work, we only considered the interference that is generated within a single-cell. However, in practical systems, there are many macro cells (multi-cells) within the network, which can cause additional interference to D2D and cellular communications. An interesting research direction, is to consider a multi-cell scenario, with other game theoretic approaches which could be combined with the existing games in Chapters 3 and 4, in order to handle the additional interference and resource block allocation. See [211] for an example of this type of problem, where [211] does not consider using game theoretic approaches. In fact, modelling this work using game theory, would provide analysis of the dynamic nature of the network, particularly when D2D pairs change transmission modes, and also provide analysis of the effect on resource allocation and interference across the different cells. Moreover, in multi-cell scenarios a D2D pair transmitter could potentially be located in one cell, where the receiver is in another cell, and hence these users would experience interference from both cells.

Quality-of-Experience (QoE) in Future Wireless Networks: As discussed in this thesis, QoE is an emerging measure of quality of communications in future wireless networks. In Chapter 3, we investigated end-user satisfaction for D2D communications, where particular physical factors such as, reliability, priority, throughput, or transmit power, can be enhanced in order to increase D2D user's QoE. One possible extension of this work, would be to jointly consider other QoE approaches, such as mean opinion satisfaction as in [39] or [124], along with enhancing physical factors (as performed in Chapter 3), in order to capture a wider range of user QoE. Moreover, since Chapter 3 utilises PDR as a measure of end user satisfaction in the physical layer, a possible extension would be to emulate a protocol stack with at least the physical and medium access control (MAC) layers in the simulations. Further investigation into jointly optimising mode selection, power

allocation, resource block allocation and user QoE for D2D and cellular users, is also another interesting research direction to significantly enhance user satisfaction in future wireless networks. In fact, in multimedia IoT, there is a range of different devices, such as CTDs and MTDs, where considering QoE for all type of communications (i.e., cellular and M2M) could also be potential future work, where the heterogeneity of the different applications in wireless networks could be analysed [212].

Convergence Rate for Game Theoretic Approaches: Although game theoretic approaches are fully distributed and autonomous, one may ask whether the game converges in time? Or will the game tend to a stable state over time? Investigating the convergence rate of a game is a significant research direction when using game theory to solve particular problems for multimedia IoT. One suggestion is to consider the channel state information (CSI) as a cost in the utility function, where the rate of convergence will be considered in the optimisation problem of the game, in order to show how the game converges in time for real world applications.

Dynamic Games: In this thesis we mainly focused on static deterministic games (such as a non-cooperative game, Stackelberg game, coalitional game, and a matching game) and a dynamic game (such as an evolutionary game). There have been advances in dynamic Markovian games, where players do not have full information of the game state. Since IoT networks have a time varying nature, the Markovian games will enable players to adapt to changes in the network, as well as modelling the channel evolution and network traffic in the game.

UAV as Flying BSs: The work in Chapter 6 focussed on scheduling UAVs to cover aggregators on the ground and minimising the power consumption of the aggregators. Although scheduling UAVs ensures ground users are served, UAVs only have a finite battery lifetime. In particular, UAVs consume two forms of power, i.e., the power consumption of the drone itself (i.e., the mechanical power consumption), as well as the power consumption for relaying the collected data to a BS or server, and other signalling overhead tasks [213]. Hence, power allocation for UAVs as flying BSs is a future research direction, especially when UAVs are deployed in remote geographical regions, where power consumption needs to be limited in order to satisfy mission critical M2M applications. Moreover, there are still open research problems for UAVs in multimedia IoT, such as defining realistic UAV channel models, optimal UAV deployment, and UAV trajectory [27].

Distributed Machine Learning: A promising future research direction for designing future wireless networks, is the use of distributed machine learning. Distributed machine learning will enable users in the network to improve their performance by learning their environment and learning from their past experiences [27, 214]. In fact, distributed machine learning will enable users to predict future states or events with respect to their wireless usage and networks. Thus, adapting distributed machine learning to the existing problems presented in this thesis, could provide further performance benefits.

Appendix A

This appendix contains the proofs needed in Chapter 3.

A.1 Proof of Proposition 3.1

Transmit power p_{di} of D2D pair i belongs to a nonempty, convex, and compact subset of Euclidean space $\mathbb{R}^D$, where transmit power is bounded by $[p_{\min}, p_{\max}]$. To prove this, we will determine the first and second derivative of each utility function defined in (3.6), (3.7), and (3.8) [215]. Firstly, we prove that these three utility functions continuous within the range $[p_{\min}, p_{\max}]$, as follows:

$$\text{Casual class:} \quad \frac{\partial u_{\text{Ca},i}(\cdot)}{\partial p_{di}} = -\theta_i'; \tag{A.1}$$

$$\text{Interactive class:} \quad \frac{\partial u_{\text{Ia},i}(\cdot)}{\partial p_{di}} = -q + \frac{2|g_{di,di}|^2 k(\overline{\gamma}_d - \gamma_{di})}{I_{di} + N_0} - \theta_i'; \tag{A.2}$$

$$\text{Streaming class:} \quad \frac{\partial u_{\text{Sa},i}(\cdot)}{\partial p_{di}} = \frac{v|g_{di,di}|^2}{\ln(2)(I_{di} + N_0)(\gamma_{di} + 1)} - 2wp_{di} - \theta_i'; \tag{A.3}$$

where

$$\begin{aligned} \theta_i' &= \frac{\partial \theta_i}{\partial p_{di}}, \\ \therefore \theta_i' &= \delta\left(x + p_{di}x'\right)\left(0.5 + \alpha_{j,i}\widetilde{G}_{i,BS}\right) + p_{di}x\delta\alpha_{j,i}'\widetilde{G}_{i,BS}, \end{aligned} \tag{A.4}$$

$x' = \frac{\partial x}{\partial p_{di}}$, and $\alpha_{j,i}' = \frac{\partial \alpha_{j,i}}{\partial p_{di}}$. The intra-cell interference D2D pair i receives from cellular user j, and other D2D pairs reusing the same resource block as cellular user j, is given by, $I_{di} = p_{cj}|g_{cj,di}|^2 + \sum_{\substack{m \in \mathcal{M} \\ m \neq i}} p_{dm}|g_{dm,di}|^2$. Secondly, we prove that these three utility functions are also strictly concave with respect to transmit power,

as follows:

$$\text{Casual class:} \quad \frac{\partial^2 u_{\text{Ca},i}(\cdot)}{\partial p_{di}^2} = -\theta_i''; \tag{A.5}$$

$$\text{Interactive class:} \quad \frac{\partial^2 u_{\text{Ia},i}(\cdot)}{\partial p_{di}^2} = -2k\left(\frac{|g_{di,di}|^2}{I_{di}+N_0}\right)^2 - \theta_i''; \tag{A.6}$$

$$\text{Streaming class:} \quad \frac{\partial^2 u_{\text{Sa},i}(\cdot)}{\partial p_{di}^2} = -\frac{v}{\ln(2)(\gamma_{di}+1)^2} \times \left(\frac{|g_{di,di}|^2}{I_{di}+N_0}\right)^2 - 2w - \theta_i''; \tag{A.7}$$

where

$$\begin{aligned} \theta_i'' &= \frac{\partial^2 \theta_i}{\partial p_{di}^2} \\ \therefore \theta_i'' &= \delta\left(2x' + p_{di}x''\right)\left(0.5 + \alpha_{j,i}\widetilde{G}_{i,BS}\right) + \alpha_{j,i}'\delta\widetilde{G}_{i,BS}\left(x + 2p_{di}x'\right) + \alpha_{j,i}''p_{di}x\delta\widetilde{G}_{i,BS}, \end{aligned} \tag{A.8}$$

$x'' = \frac{\partial^2 x}{\partial p_{di}^2}$, and $\alpha_{j,i}'' = \frac{\partial^2 \alpha_{j,i}}{\partial p_{di}^2}$.

For (A.5), (A.6), and (A.7) to satisfy the strict concavity property, $\frac{\partial^2 u_{F,i}(\cdot)}{\partial p_{di}^2} < 0$, $u_{F,i} \in \{u_{\text{Ca}}, u_{\text{Ia}}, u_{\text{Sa}}\}$, the following constants k, v, and w must always be greater than zero.

In addition to the strict concavity analysis of the utility function, general supermodularity can also be considered. General supermodularity is a less restrictive analysis on the utility function, and is guaranteed if the strategy set is compact and each utility has increasing differences [216].

A.2 Proof of Proposition 3.2

The BS satisfaction fee x belongs to a nonempty, convex, and compact subset of Euclidean space $\mathbb{R}^1$, where the BS satisfaction fee x is bounded by $(0, 1]$. The reuse fee $\alpha_{j,i}$ also belongs to a nonempty, convex, and compact subset of Euclidean space $\mathbb{R}^D$, where the reuse fee $\alpha_{j,i}$ is bounded by $\alpha_{j,i} > 0$ and $[\alpha_{j,\min}, \alpha_{j,\max}]$. Thus, we prove that the leader's utility function defined in (3.4) is continuous within the leader's strategy set $(0, 1]$ and the reuse fee range $[\alpha_{j,\min}, \alpha_{j,\max}]$ $\forall j \in \mathcal{C}$, as follows:

$$\frac{\partial u_L(\cdot)}{\partial x} = \beta\left(\overline{p} + \overline{I}_{i,BS}(\alpha_{j,i})\right) - 2x\lambda\overline{r}^2, \tag{A.9}$$

$$\frac{\partial u_L(\cdot)}{\partial \alpha_{j,i}} = \beta x p_{di} \widetilde{G}_{i,BS} - 2\alpha_{j,i}(p_{di}\widetilde{G}_{i,BS})^2. \tag{A.10}$$

Furthermore, we prove that the leader's utility function (3.4) is also strictly concave with respect to the BS satisfaction fee x across all users in the network, as well as the reuse fee $\alpha_{j,i}$ for each D2D pair i and the corresponding cellular user j, as follows:

$$\frac{\partial^2 u_L(\cdot)}{\partial x^2} = -2\lambda \overline{r}^2, \tag{A.11}$$

$$\frac{\partial^2 u_L(\cdot)}{\partial \alpha_{j,i}^2} = -2(p_{di}\widetilde{G}_{i,BS})^2. \tag{A.12}$$

Equation (3.4) satisfies the strict concavity property, as follows, $\frac{\partial^2 u_L(\cdot)}{\partial x^2} < 0$ and $\frac{\partial^2 u_L(\cdot)}{\partial \alpha_{j,i}^2} < 0$.

A.3 Proof of Theorem 3.1

We have shown that all three utility functions used by the followers are strictly concave, continuous, and compact, across the same strategy space, that is, $[p_{\min}, p_{\max}]$. Following from the proof in [171], if all followers utility functions, whatever they are, are strictly concave and continuous in the same strategy space that is compact, then a unique Stackelberg equilibrium follows.

The net utility across all followers is strictly concave, continuous, and compact, across their strategy space $[p_{\min}, p_{\max}]$, as follows:

$$U_F(x, \alpha_{j,i}, p_{di}) = \sum_{\substack{i \in \mathcal{B} \\ \mathcal{B} \subset \mathcal{D}}} u_{\mathrm{Ca},i} + \sum_{\substack{i \in \mathcal{Y} \\ \mathcal{Y} \subset \mathcal{D}}} u_{\mathrm{Ia},i} + \sum_{\substack{i \in \mathcal{Z} \\ \mathcal{Z} \subset \mathcal{D}}} u_{\mathrm{Sa},i}, \tag{A.13}$$

where D2D pairs remain in one application class during the game, which means that $\mathcal{B} \neq \mathcal{Y} \neq \mathcal{Z}$, and thus $\mathcal{D} = \mathcal{B} \cup \mathcal{Y} \cup \mathcal{Z}$.

Thus, maximising the net utility across all followers, ensures that the BS's perception and satisfaction is maximised, ensuring followers' maximum social welfare.

Appendix B

This appendix contains the proofs needed in Chapter 4.

B.1 Proof of Proposition 4.2

Firstly, we prove that the utility function for cellular users defined in (4.11) and the utility function for D2D pairs defined in (4.12) for the proposed non-cooperative power control game, are continuous within the range $[p_{\min}, p_{\max}]$. Transmit power p_i of user $i \in \mathcal{N}$ belongs to a nonempty, convex, and compact subset of Euclidean space $\mathbb{R}^{|\mathcal{N}|}$, where transmit power is bounded by $[p_{\min}, p_{\max}]$. In order to prove the utility is continuous, the first order derivative condition must be satisfied, i.e., $\frac{\partial u_{\mathrm{pc}_i}}{\partial p_i} = 0$. Note that, the proof for the cellular user utility is the same as the D2D pair utility. We outline the proof for continuity within the range $[p_{\min}, p_{\max}]$ for D2D pair i, as follows:

$$\frac{\partial u_{\mathrm{pc}_{di,\diamond}}(\cdot)}{\partial p_{di}} = -w_i + \frac{2|g_{di,di}|^2 v_i(\overline{\gamma}_d - \gamma_{di})}{I_{di} + N_0}, \tag{B.1}$$

where I_{di} is the intra-cell interference that D2D pair i receives from cellular user j, and other D2D pairs sharing the same resource block as cellular user j. Intra-cell interference will equal zero for D2D paris in cellular mode and for cellular users. Thus, the best response transmit power p^*_{di}, $\frac{\partial u_{\mathrm{pc}_i}}{\partial p_i} = 0$, is derived as follows:

$$p^*_{di} = \frac{I_{di} + N_0}{|g_{di,di}|^2} \left(\overline{\gamma}_d - \frac{w_i(I_{di} + N_0)}{2v_i} \right), \tag{B.2}$$

where the best response transmit power is searched within $\{p_{\min}, p^*_{di}, p_{\max}\}$. Hence, the utility function has a unique maximum, and thus, allows us to find a unique

solution for transmit power that is always maximum.

Secondly, we prove that the utility function for cellular users defined in (4.11) and the utility function for D2D pairs defined in (4.12) for the proposed non-cooperative power control game, are also strictly concave with respect to transmit power. In order to prove the utility is strictly concave, the second order derivative condition must be satisfied, i.e., $\frac{\partial^2 u_{\mathrm{pc}_i}}{\partial p_i^2} < 0$. Again, the proof for the cellular user utility is the same as the D2D pair utility. We outline the proof for strict concavity for D2D pair i, as follows:

$$\frac{\partial^2 u_{\mathrm{pc}_{i,\diamond}}(\cdot)}{\partial^2 p_{di}} = -2v_i \left(\frac{|g_{di,di}|^2}{I_{di} + N_0} \right)^2. \tag{B.3}$$

From (B.3) we see that the second derivative of the utility function with respect to transmit power satisfies the strictly concavity property, where the following constants w_i and v_i must be greater than zero.

B.2 Proof of Proposition 4.3

For the utility function for the two-armed bandit game, we prove that the expected reward is strictly concave and continuous. For simplicity reasons lets remove the expectation in (4.16) and drop the dependence on t for $\alpha_{di,RM}$, $u_{\mathrm{pc}_{di,\diamond}}$, and ρ. We outline the proof for continuity for D2D pair $i \in \mathcal{N}$, as follows:

$$u_{\mathrm{arm}_i}(\rho) = \int_{t=0}^{T} r \exp(-rt) \left((1 - \alpha_{di,RM}) u_{\mathrm{pc}_{di,CM}} + \rho \alpha_{di,RM} u_{\mathrm{pc}_{di,RM}} \right) dt | \rho(0), \tag{B.4}$$

$$\frac{\partial u_{\mathrm{arm}_i}(\rho)}{\partial \rho} = \left[- \exp(-rt) \left((1 - \alpha_{di,RM}) u_{\mathrm{pc}_{di,CM}} + \rho \alpha_{di,RM} u_{\mathrm{pc}_{di,RM}} \right) \Big|_{t=0}^{T} \right] \frac{\partial}{\partial \rho}, \tag{B.5}$$

$$= \left[(1 - \exp(-rT)) \left((1 - \alpha_{di,RM}) u_{\mathrm{pc}_{di,CM}} + \rho \alpha_{di,RM} u_{\mathrm{pc}_{di,RM}} \right) \right] \frac{\partial}{\partial \rho}. \tag{B.6}$$

From (4.14), we can rewrite (B.6), as follows:

$$\therefore \frac{\partial u_{\mathrm{arm}_i}(\rho)}{\partial \rho} = \alpha_{di,RM} u_{\mathrm{pc}_{di,RM}} \big(1 - \exp(-rT)\big) \times \left[\frac{\exp(-\alpha_{di,RM}\beta\delta)}{(1-\rho+\rho\exp(-\alpha_{di,RM}\beta\delta))^2}\right]. \tag{B.7}$$

Next, we prove that the utility function for the two-armed bandit game is strictly concave with respect to D2D pair i's belief ρ, as follows:

$$\frac{\partial^2 u_{\mathrm{arm}_i}(\rho)}{\partial^2 \rho} = \alpha_{di,RM} u_{\mathrm{pc}_{di,RM}} \big(1 - \exp(-rT)\big) \times \left[\frac{2\exp(-\alpha_{di,RM}\beta\delta) - 2\exp(-2\alpha_{di,RM}\beta\delta)}{(1-\rho+\rho\exp(-\alpha_{di,RM}\beta\delta))^3}\right]. \tag{B.8}$$

Equation (B.8) shows that the second derivative of the utility function with respect to D2D pair i's belief ρ satisfies the strict concavity property, due to the non-cooperative transmit power utility function, $u_{\mathrm{pc}_{di,RM}}$, always being negative ($u_{\mathrm{pc}_{di,RM}} < 0$).

B.3 Proof of Proposition 4.5

We observe that transmit power p_i of user i in coalition $\mathcal{F}_\pi \in \mathcal{F}$ belongs to a nonempty, convex, and compact subset of Euclidean space $\mathbb{R}^{|\mathcal{F}_\pi|}$, where transmit power is bounded by, $[p_{min}, p_{max}]$. Firstly we prove that the utility function defined in (4.18) is continuous by computing its first derivative with respect to transmit power, as follows:

$$\frac{\partial u_i(\mathcal{F}_\pi)}{\partial p_i} = \frac{\lambda_i \gamma_i}{\ln(2) p_i (1+\gamma_i)} - 2\omega p_i. \tag{B.9}$$

Thus, (4.18) is continuous within the range $[p_{min}, p_{max}]$. Furthermore, we prove that the utility function defined in (4.18) is also strictly concave, by computing the second order derivative of the utility with respect to transmit power, as follows:

$$\frac{\partial^2 u_i(\mathcal{F}_\pi)}{\partial p_i^2} = -\frac{\lambda_i \gamma_i^2}{\ln(2) p_i^2 (1+\gamma_i)^2} - 2\omega. \tag{B.10}$$

In order for (4.18) to satisfy the strict concavity property, $\frac{\partial^2 u_i(\mathcal{F}_\pi)}{\partial p_i^2} < 0$, the following terms λ_i, p_i, γ_i, and ω must always be greater than zero. Therefore, the utility function defined in (4.18) for each user in the CLC game is strictly concave

and continuous with respect to transmit power p_i.

B.4 Proof of Theorem 4.1

The coalition partition $\mathcal{F}^*$ is obtained for all cellular and D2D users from the proposed Algorithm 3 or Algorithm 4. Let $\mathcal{F}'$ be any other coalition partition which is different to $\mathcal{F}^*$. For all cellular and D2D users, with the following coalition partition:

$$\sum_{i:\exists \mathcal{F}_\pi \in \mathcal{F}^*, i \in \mathcal{F}_\pi} u_i(\mathcal{F}^*) \geq \sum_{i:\exists \mathcal{F}_\pi \in \mathcal{F}', i \in \mathcal{F}_\pi} u_i(\mathcal{F}') \tag{B.11}$$

and according to Definition 4.9, the achieved coalition partition $\mathcal{F}^*$ is Nash stable. Thus, the proposed CLC game achieves a Nash stable coalition partition that is socially efficient, where social efficiency maximizes social welfare across all coalitions (where the social welfare considered is the sum of all utility functions within the game) [72, 177]. We also see that $\mathcal{F}^*$ is individually stable and contractually individually stable, as outlined in Remark 4.7.

Appendix C

This appendix contains the proofs needed in Chapter 5.

C.1 Derivation of Equation (5.11): Closed-Form Expression for the Utility Function

The utility function $u_{\mathcal{K}_{i,j}}(t)$ of typical cluster $\mathcal{K}_{i,j}$ choosing type j, is expressed as follows:

$$u_{\mathcal{K}_{i,j}}(t) = -\frac{1}{j}\sum_{k=1}^{j-1}\int_0^{r_j}\left(\frac{\mathbb{E}(I_j)}{g_{ki}}\right)\left(2^{\frac{H_k}{TB}}-1\right)f_{\mathcal{D}_i}^{(k)}(r)dr - \frac{1}{j}\left(\frac{BN_0}{g_i}\right)\left(2^{\frac{H_{\mathcal{K}_{i,j}}}{TB}}-1\right)-\delta j. \tag{C.1}$$

In (C.1), there is a tradeoff between average cluster size and average transmission power for MTDs. Thus, as the average cluster size increases, the average transmit power per cluster also increases, due to the increasing number of MTDs on average. The proposed utility function in (C.1) captures the average transmission power for all MTDs across the M2M and cellular links within the cluster, sending all the gathered data of the cluster. From (5.3), we can rewrite (C.1) as follows:

$$\begin{aligned} u_{\mathcal{K}_{i,j}}(t) = &-\frac{1}{j}\sum_{k=1}^{j-1}\int_0^{r_j}\left(\mathbb{E}(I_j)\left(\frac{4\pi r}{\mu}\right)^{\nu}\right)\left(2^{\frac{H_k}{TB}}-1\right)f_{\mathcal{D}_i}^{(k)}(r)dr- \\ &\frac{1}{j}\left(\frac{BN_0}{g_i}\right)\left(2^{\frac{H_m+H_m(j-1)\left(1-\frac{\alpha}{\frac{r_j}{c}+1}\right)}{TB}}-1\right)-\delta j. \end{aligned} \tag{C.2}$$

The utility function in (C.2) for a given cluster $\mathcal{K}_{i,j}$, is a function of the cluster radius r_j and the cluster type j, where j is equal to the average cluster size,

$j = K_{i,j}$. Furthermore, from (5.7) and (5.10), we can rewrite the utility function (C.2) of cluster $\mathcal{K}_{i,j}$ choosing type j as a closed-form expression, as follows:

$$u_{\mathcal{K}_{i,j}}(t) = -\frac{1}{j}\left(\frac{4\pi}{\mu}\right)^{\nu}\mathbb{E}(I_j)\left(2^{\frac{H_m}{TB}}-1\right)\sum_{k=1}^{j-1}\frac{(j-1)!}{(k-1)!(j-1-k)!}\int_0^{r_j}\frac{2r^{\nu+1}}{r_j^2}\left(\frac{r^2}{r_j^2}\right)^{k-1}\times \left(1-\frac{r^2}{r_j^2}\right)^{j-1-k}dr - \frac{1}{j}\left(\frac{BN_0}{g_i}\right)\left(2^{\frac{H_m+H_m(j-1)\left(1-\frac{\alpha}{\frac{r_j}{c}+1}\right)}{TB}}-1\right)-\delta j. \tag{C.3}$$

To further simplify (C.3), if $\frac{r^2}{r_j^2} = a$, then $dr = \frac{r_j da}{2a^{\frac{1}{2}}}$ and $r = r_j a^{\frac{1}{2}}$. Thus:

$$\int_0^{r_j}\frac{2r^{\nu+1}}{r_j^2}\left(\frac{r^2}{r_j^2}\right)^{k-1}\left(1-\frac{r^2}{r_j^2}\right)^{j-1-k}dr = \int_0^1 2ar_j^{\nu-1}a^{\frac{\nu-1}{2}}a^{k-1}(1-a)^{j-1-k}\frac{r_j da}{2a^{\frac{1}{2}}}$$
$$\therefore = r_j^{\nu}\int_0^1 a^{k+\frac{\nu-2}{2}}(1-a)^{j-1-k}da. \tag{C.4}$$

Since $\int_0^{r_j}\frac{2r^{\nu+1}}{r_j^2}\left(\frac{r^2}{r_j^2}\right)^{k-1}\left(1-\frac{r^2}{r_j^2}\right)^{j-1-k}dr = \frac{r_j^{\nu}\Gamma(j-k)\Gamma(k+\frac{v}{2})}{\Gamma(j+\frac{v}{2})}$, where $\Gamma(\cdot)$ is the Gamma function, and j and k are integers, such that $j > k$, and $2k+\nu > 0$ where $\nu > 2$ and ν is not an integer. Thus, (C.3) can be rewritten as:

$$u_{\mathcal{K}_{i,j}}(t) = -\frac{1}{j}\left(\frac{4\pi}{\mu}\right)^{\nu}\mathbb{E}(I_j)\left(2^{\frac{H_m}{TB}}-1\right)\sum_{k=1}^{j-1}\frac{(j-1)!}{(k-1)!(j-1-k)!}\frac{r_j^{\nu}\Gamma(j-k)\Gamma(k+\frac{\nu}{2})}{\Gamma(j+\frac{\nu}{2})} - \frac{1}{j}\left(\frac{BN_0}{g_i}\right)\left(2^{\frac{H_m+H_m(j-1)\left(1-\frac{\alpha}{\frac{r_j}{c}+1}\right)}{TB}}-1\right)-\delta j. \tag{C.5}$$

Since path loss, ν, is not always an integer we cannot express (C.5) as a function of factorials, thus we can further simplify (C.5) in terms of the Γ function, given that $\Gamma(n) = (n-1)!$, $\Gamma(n+1) = n\Gamma(n)$, and $(n-1-m)! = \Gamma(n-m)$, where n and m are positive integers. (C.5) can be rewritten as:

$$u_{\mathcal{K}_{i,j}}(t) = -\frac{1}{j}\left(\frac{4\pi r_j}{\mu}\right)^{\nu}\mathbb{E}(I_j)\left(2^{\frac{H_m}{TB}}-1\right)\sum_{k=1}^{j-1}\frac{\Gamma(j)\Gamma(k+\frac{\nu}{2})}{\Gamma(k)\Gamma(j+\frac{\nu}{2})} - \frac{1}{j}\left(\frac{BN_0}{g_i}\right)\left(2^{\frac{H_m+H_m(j-1)\left(1-\frac{\alpha}{\frac{r_j}{c}+1}\right)}{TB}}-1\right)-\delta j. \tag{C.6}$$

Therefore, by deriving the closed-form expression of the summation in (C.6), $\sum_{k=1}^{j-1} \frac{\Gamma(j)\Gamma(k+\frac{\nu}{2})}{\Gamma(k)\Gamma(j+\frac{\nu}{2})} = \frac{2(j-1)}{2+\nu}$, we can find the closed-form expression of (C.1), as defined in (5.11).

C.2 Proof of Theorem 5.1

Since the set of cluster types in game $\mathcal{G}_E$ is finite, the replicator dynamic will converge to the steady state which is the (symmetric) mixed strategy Nash equilibrium of $\mathcal{G}_E$ [217]. Let $\boldsymbol{x}^*$ be the converged population state of the replicator dynamic from proposed Algorithm 5. Following (5.12), at the converged population state of proposed Algorithm 5, we have:

$$\frac{\partial \dot{x}_j^*(t)}{\partial x_m^*(t)} = \begin{cases} \overline{u}_{\mathcal{K}_{i,j}}(\boldsymbol{x}^*,t) - U_{\mathcal{K}_i}(\boldsymbol{x}^*,t) + \left(\frac{\partial \overline{u}_{\mathcal{K}_{i,j}}(\boldsymbol{x}^*,t)}{\partial x_j^*(t)} - \frac{\partial U_{\mathcal{K}_i}(\boldsymbol{x}^*,t)}{\partial x_j^*(t)}\right) x_j^*, & \text{for } j = m \\ \left(\frac{\partial \overline{u}_{\mathcal{K}_{i,j}}(\boldsymbol{x}^*,t)}{\partial x_m^*(t)} - \frac{\partial U_{\mathcal{K}_i}(\boldsymbol{x}^*,t)}{\partial x_m^*(t)}\right) x_m^*, & \text{for } j \neq m \end{cases} \tag{C.7}$$

Based on (5.12), at the convergence, for any $0 < x_j^*(t)$, $\overline{u}_{\mathcal{K}_{i,j}}(\boldsymbol{x}^*,t) - U_{\mathcal{K}_i}(\boldsymbol{x}^*,t) = 0$. Moreover, $\frac{\partial \overline{u}_{\mathcal{K}_{i,j}}(\boldsymbol{x}^*,t)}{\partial x_j^*(t)} = u_{\mathcal{K}_{i,j}}$, $\frac{\partial U_{\mathcal{K}_i}(\boldsymbol{x}^*,t)}{\partial x_j^*(t)} = \overline{u}_{\mathcal{K}_{i,j}}(\boldsymbol{x},t) + u_{\mathcal{K}_{i,j}}(\boldsymbol{x},t)x_j^* + \sum_{n\in\mathcal{S}\setminus\{j\}} u_{\mathcal{K}_{i,nj}}(\boldsymbol{x},t)x_n^*$, $\frac{\partial \overline{u}_{\mathcal{K}_{i,j}}(\boldsymbol{x}^*,t)}{\partial x_m^*(t)} = u_{\mathcal{K}_{i,jm}}$, and $\frac{\partial U_{\mathcal{K}_i}(\boldsymbol{x}^*,t)}{\partial x_m^*(t)} = \overline{u}_{\mathcal{K}_{i,m}}(\boldsymbol{x},t) + u_{\mathcal{K}_{i,jm}}(\boldsymbol{x},t)x_m^* + \sum_{n\in\mathcal{S}\setminus\{m\}} u_{\mathcal{K}_{i,jn}}(\boldsymbol{x},t)x_n^*$. Therefore, (C.7) can be rewritten as follows:

$$\frac{\partial \dot{x}_j^*(t)}{\partial x_m^*(t)} = \big(u_{\mathcal{K}_{i,jm}}(\boldsymbol{x}^*,t) - 2\overline{u}_{\mathcal{K}_{i,m}}(\boldsymbol{x}^*,t)\big)x_m^*. \tag{C.8}$$

Since $\boldsymbol{x}^*$ is mixed strategy Nash equilibrium, i.e., $u_{\mathcal{K}_{i,m}}(\boldsymbol{x}^*,t) \leq \sum_{j'\in\mathcal{S}} u_{\mathcal{K}_{i,mj'}}(t)x_{j'}^*$, and based on (5.13) $\overline{u}_{\mathcal{K}_{i,m}}(\boldsymbol{x}^*,t) = \sum_{j'\in\mathcal{S}} u_{\mathcal{K}_{i,mj'}}(t)x_{j'}^*$, and thus, $\frac{\partial \dot{x}_j^*(t)}{\partial x_m^*(t)} < 0$. Consequently, the convergence state of the proposed algorithm, and specifically the steady state of the replicator dynamic in (5.12) is locally asymptotically stable. Since steady state of the replicator dynamic is a locally asymptotically stable state in our problem, then the proposed Algorithm 5 will converge to ESS [217, 218]. This completes the proof.

We assume $\boldsymbol{x}^\epsilon$ is a preference profile vector, where ϵ potion of MTDs deviate while the remaining $1-\epsilon$ potion of MTDs choose the converged population state $\boldsymbol{x}^*$. Let $x_j^\epsilon = x_j^* + \epsilon_j \frac{\partial \dot{x}_j}{\partial x_j}(\boldsymbol{x}^*)$. Since, $0 \leq x_j^\epsilon$, we have:

$$|\epsilon_j| \leq 1, \forall j \in \mathcal{S}, \tag{C.9}$$

$$\sum_{j \in \mathcal{S}} \epsilon_j = 0, \tag{C.10}$$

$$\therefore |\epsilon_j| \leq \frac{1}{2\overline{u}_{\mathcal{K}_{i,j}}(\boldsymbol{x}^*, t) - u_{\mathcal{K}_{i,j}}(\boldsymbol{x}^*, t)} \forall j \in \mathcal{S}. \tag{C.11}$$

To hold the above inequalities, the maximum value of ϵ is given by (5.16).

Appendix D

This appendix contains the proofs needed in Chapter 6.

D.1 Proof of Lemma 6.1

Considering the fact that for any $\forall x \in \mathbb{R}$, $(\max\{x, 0\})^2 \leq x^2$, first by squaring both sides of (6.6):

$$q_{g(h+1)}^2 - q_{gh}^2 \leq (q_{gh} - d_{gk,h})^2 + a_{gh}^2 + 2a_{gh}q_{gh} - q_{gh}^2. \tag{D.1}$$

By simplification, we have:

$$\frac{q_{g(h+1)}^2 - q_{gh}^2}{2} \leq \frac{d_{gk,h}^2 + a_{gh}^2}{2} - q_{gh}(d_{gk,h} - a_{gh}). \tag{D.2}$$

Similarly,

$$\frac{z_{1(h+1)}^2 - z_{1h}^2}{2} \leq \frac{(Q_h - M_{\text{th}})^2}{2} + z_{1h}(Q_h - M_{\text{th}}), \tag{D.3}$$

$$\frac{z_{2(h+1)}^2 - z_{2h}^2}{2} \leq \frac{(Q_h^2 - B_{\text{th}})^2}{2} + z_{2h}(Q_h^2 - B_{\text{th}}). \tag{D.4}$$

By taking conditional expectations of (D.2)-(D.4) and summing them, and then adding v, we have:

$$L_{h+1} - L_h \leq I + \sum_{g=1}^{G} q_{gh}(a_{gh} - d_{gk,h}) + z_{1h}(Q_h - M_{\text{th}}) + \tag{D.5}$$

$$z_{2h}(Q_h^2 - B_{\text{th}}) \tag{D.6}$$

where $I = \frac{H^2+(\max_g |\mathcal{G}_g|)^2+M_{\text{th}}^2+B_{\text{th}}^2}{2}$ because $d_{gk,h} \leq H$ and $a_{gh} \leq \max_g |\mathcal{G}_g|$. Thus, we see that the lemma follows.

D.2 Proof of Theorem 6.1

At each decision epoch h, all UAVs $k \in \mathcal{K}$ are involved in the one-to-many matching game. The convergence of Algorithm 6 at each decision epoch h is guaranteed, since a UAV never proposes to a CH twice. In the worst case, all UAVs will propose to all CHs once, in decision epoch h.

Once the algorithm converges, the resulting matching between UAVs and CHs is two-sided stable. Let's assume that there exists a UAV and CH pair, $(k, g) \notin \Phi$, that blocks the matching Φ. Firstly, since the algorithm has converged, this means that UAV k does not need to change the subset of CHs it is serving. Thus, UAV k would not replace any CH subset $\mathcal{S}_{kh} \in \Phi(k)$ with $\mathcal{S}'_{kh}$, as $\mathcal{S}_{kh} \succ_k \mathcal{S}'_{kh}$. Secondly, if UAV k wanted to now propose to CH subset $\mathcal{S}'_{kh}$, then this would have occurred earlier prior to convergence. On the other hand, if UAV k did propose to CH subset $\mathcal{S}'_{kh}$ prior to convergence and the proposal was rejected, this means that $\Phi(g) \succ_g k$, where $g \in \mathcal{S}'_{kh}$, which contradicts that (g, k) is a blocking pair.

D.3 Proof of Theorem 6.2

In Theorem 6.1, we have proved that the one-to-many matching game will converge to a two-sided stable matching. Therefore, we need to prove that the convex optimization problem in (6.35) for CH power allocation will also converge to a stable outcome. We have considered a power control game to solve the optimization problem in (6.35), and observe that transmit power p_{gh} for CH g belongs to a nonempty, convex, and compact subspace of Euclidean space $\mathbb{R}^{|\mathcal{A}|}$. Note that transmit power is bounded by, $[0, P_{\max}]$. Firstly, we prove that the utility function (6.34) defined in the optimization problem (6.35) is continuous, by computing its first derivative with respect to transmit power, as follows:

$$\frac{\partial U_g(\cdot)}{\partial p_{gh}} = \frac{\beta \overline{\xi}_{gk,t}}{D_g B \sigma^2 \left(1 + \frac{p_{gh} \beta \overline{\xi}_{gk,t}}{B\sigma^2}\right) \ln(2)} - v. \tag{D.7}$$

Thus, the utility in (6.34) is continuous within the range $[0, P_{\max}]$. Next, we

prove that the utility (6.34) is also strictly concave, by computing the second order derivative with respect to transmit power, as follows:

$$\frac{\partial^2 U_g(\cdot)}{\partial p_{gh}^2} = -\frac{(\beta\overline{\xi}_{gk,t})^2}{D_g(B\sigma^2)^2\left(1+\frac{p_{gh}\beta\overline{\xi}_{gk,t}}{B\sigma^2}\right)^2\ln(2)} < 0. \tag{D.8}$$

Therefore, the utility function (6.34) defined in the optimization problem (6.35) for each CH is strictly concave and continuous with respect to transmit power, and thus, will guarantee convergence to a stable transmit power outcome.

Overall, in each decision epoch h, the matching and power allocation guarantees that the total sum utility, i.e., $U_{\text{sum},h} = \sum_{k\in\mathcal{K}} U_k(\mathcal{S}_{kh}, \Phi, \Psi_1, \Psi_2) + \sum_{g\in\mathcal{A}} U_g(k, \Phi, \Psi_1, \Psi_2)$, will grow, as follows:

$$U_{\text{sum},h} \geq U_{\text{sum},(h-1)}. \tag{D.9}$$

In particular, since the number of matches is finite, there must exist at least one optimal matching which leads to the maximum sum utility, i.e., maximum social welfare.

Bibliography

[1] ITU-R, "Minimum technical performance requirements for IMT-2020 radio interface(s)," *ITU - International Telecommunication Union, Tech. Rep. Workshop*, 2016.

[2] Sergey Andreev, Dmitri Moltchanov, Olga Galinina, Alexander Pyattaev, Aleksandr Ometov, and Yevgeni Koucheryavy, "Network-assisted device-to-device connectivity: Contemporary vision and open challenges," in *Proc. of 21th European Wireless Conference*, Budapest, Hungary, May 2015, pp. 1–8.

[3] Qin Wang, Yanxiao Zhao, Wei Wang, Daniel Minoli, Kazem Sohraby, Hongbo Zhu, and Ben Occhiogrosso, "Multimedia IoT systems and applications," in *Proc. of Global Internet of Things Summit (GIoTS)*, Geneva, Switzerland, Jun. 2017, IEEE, pp. 1–6.

[4] Nisha Panwar, Shantanu Sharma, and Awadhesh Kumar Singh, "A survey on 5G: The next generation of mobile communication," *Physical Communication*, vol. 18, pp. 64–84, Mar. 2016.

[5] Ala Al-Fuqaha, Mohsen Guizani, Mehdi Mohammadi, Mohammed Aledhari, and Moussa Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015.

[6] Alessandro Floris and Luigi Atzori, "Quality of experience in the multimedia internet of things: Definition and practical use-cases," in *Proc. of IEEE International Conference on Communication Workshop (ICCW)*, London, UK, Jun. 2015, pp. 1747–1752.

[7] Cisco, "Cisco visual networking index: Forecast and trends, 2017–2022," *White Paper*, Nov. 2018.

[8] Jeffrey G Andrews, Stefano Buzzi, Wan Choi, Stephen V Hanly, Angel Lozano, Anthony CK Soong, and Jianzhong Charlie Zhang, "What will 5G be?," *IEEE Journal on selected areas in communications*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[9] Jiajia Liu, Nei Kato, Jianfeng Ma, and Naoto Kadowaki, "Device-to-device communication in LTE-Advanced networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 1923–1940, Fourthquarter 2015.

[10] Shancang Li, Li Da Xu, and Shanshan Zhao, "5G internet of things: A survey," *Journal of Industrial Information Integration*, Jun. 2018.

[11] Thomas Fehrenbach, Rohit Datta, Barış Göktepe, Thomas Wirth, and Cornelius Helge, "URLLC services in 5G-low latency enhancements for LTE," *arXiv preprint arXiv:1808.07034*, 2018.

[12] Cisco, "Cisco visual networking index: Forecast and methodology, 2016–2021," *White Paper Cisco public*, Jun. 2017.

[13] 3GPP, "Study on scenarios and requirements for next generation access technologies (release 15)," *3GPP TR 38.913*, 2016.

[14] Mohammed H Alsharif and Rosdiadee Nordin, "Evolution towards fifth generation (5G) wireless networks: Current trends and challenges in the deployment of millimetre wave, massive mimo, and small cells," *Telecommunication Systems*, vol. 64, no. 4, pp. 617–637, Jul. 2017.

[15] Ian F Akyildiz, Shuai Nie, Shih-Chun Lin, and Manoj Chandrasekaran, "5G roadmap: 10 key enabling technologies," *Computer Networks*, vol. 106, pp. 17–48, Sep. 2016.

[16] Pavel Mach, Zdenek Becvar, and Tomas Vanek, "In-band device-to-device communication in OFDMA cellular networks: A survey and challenges," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 1885–1922, Fourthquarter 2015.

[17] Kwang-Cheng Chen and Shao-Yu Lien, "Machine-to-machine communications: Technologies and challenges," *Ad Hoc Networks*, vol. 18, pp. 3–23, Jul. 2014.

[18] Kan Zheng, Fanglong Hu, Wenbo Wang, Wei Xiang, and Mischa Dohler, "Radio resource allocation in LTE-Advanced cellular networks with M2M communications," *IEEE Communications Magazine*, vol. 50, no. 7, pp. 184–192, Jul. 2012.

[19] Yi Liu, Chau Yuen, Xianghui Cao, Naveed Ul Hassan, and Jiming Chen, "Design of a scalable hybrid MAC protocol for heterogeneous M2M networks," *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 99–111, Feb. 2014.

[20] Mohammad Mozaffari, Walid Saad, Mehdi Bennis, and Mérouane Debbah, "Unmanned aerial vehicle with underlaid device-to-device communications: Performance and tradeoffs," *IEEE Trans. on Wireless Communications*, vol. 15, no. 6, pp. 3949–3963, June. 2016.

[21] Mingzhe Chen, Mohammad Mozaffari, Walid Saad, Changchuan Yin, Mérouane Debbah, and Choong Seon Hong, "Caching in the sky: Proactive deployment of cache-enabled unmanned aerial vehicles for optimized quality-of-experience," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 35, no. 5, pp. 1046–1061, May 2017.

[22] Taehyeun Park, Nof Abuzainab, and Walid Saad, "Learning how to communicate in the Internet of Things: Finite resources and heterogeneity," *IEEE Access, Special Issue on Optimization for Emerging Wireless Networks: IoT, 5G and Smart Grid Communication Networks*, vol. 4, pp. 7063–7073, Nov. 2016.

[23] 3GPP TR 36.746 v15.1.1, "Technical specification group radio access network: Study on further enhancements to lte device to device (D2D), user equipment (UE) to network relays for internet of things (IoT) and wearables," Apr. 2018.

[24] Marko Höyhtyä, Olli Apilo, and Mika Lasanen, "Review of latest advances in 3gpp standardization: D2d communication in 5g systems and its energy consumption models," *Future Internet*, vol. 10, no. 1, pp. 3, Jan. 2018.

[25] Pawan Kumar Verma, Rajesh Verma, Arun Prakash, Ashish Agrawal, Kshirasagar Naik, Rajeev Tripathi, Maazen Alsabaan, Tarek Khalifa, Tamer Abdelkader, and Abdulhakim Abogharaf, "Machine-to-machine (M2M) com-

munications: A survey," *Journal of Network and Computer Applications*, vol. 66, pp. 83–105, May 2016.

[26] Shao-Yu Lien, Kwang-Cheng Chen, and Yonghua Lin, "Toward ubiquitous massive accesses in 3GPP machine-to-machine communications," *IEEE Communications Magazine*, vol. 49, no. 4, Apr. 2011.

[27] Mohammad Mozaffari, Walid Saad, Mehdi Bennis, Young-Han Nam, and Mérouane Debbah, "A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," *arXiv preprint arXiv:1803.00680*, 2018.

[28] Yong Zeng, Rui Zhang, and Teng Joon Lim, "Wireless communications with unmanned aerial vehicles: opportunities and challenges," *arXiv preprint arXiv:1602.03602*, 2016.

[29] Mohammad Mozaffari, Walid Saad, Mehdi Bennis, and Mérouane Debbah, "Efficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage.," *IEEE Communications Letters*, vol. 20, no. 8, pp. 1647–1650, Jun. 2016.

[30] R Irem Bor-Yaliniz, Amr El-Keyi, and Halim Yanikomeroglu, "Efficient 3-d placement of an aerial base station in next generation cellular networks," in *Proc. of IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–5.

[31] Lei Lei, Zhangdui Zhong, Chuang Lin, and Xuemin Shen, "Operator controlled device-to-device communications in LTE-Advanced networks," *IEEE Wireless Communications*, vol. 19, no. 3, pp. 96, Jun. 2012.

[32] Gábor Fodor, Erik Dahlman, Gunnar Mildh, Stefan Parkvall, Norbert Reider, György Miklós, and Zoltán Turányi, "Design aspects of network assisted device-to-device communications," *IEEE Communications Magazine*, vol. 50, no. 3, pp. 170–177, Mar. 2012.

[33] Marco Belleschi, Gabor Fodor, and Andrea Abrardo, "Performance analysis of a distributed resource allocation scheme for D2D communications," in *Proc. of IEEE Global Telecommunications (GLOBECOM) Workshops*, Houston, TX, USA, Dec. 2011, pp. 358–362.

[34] Feiran Wang, Chen Xu, Lingyang Song, and Zhu Han, "Energy-efficient resource allocation for device-to-device underlay communication," *IEEE Trans. on Wireless Communications*, vol. 14, no. 4, pp. 2082–2092, Apr. 2015.

[35] Zhenyu Zhou, Kaoru Ota, Mianxiong Dong, and Chen Xu, "Energy-efficient matching for resource allocation in D2D enabled cellular networks," *IEEE Trans. on Vehicular Technology*, vol. 66, no. 6, pp. 5256–5268, Jun. 2017.

[36] Mohammad G Khoshkholgh, Yan Zhang, Kwang-Cheng Chen, Kang G Shin, and Stein Gjessing, "Connectivity of cognitive device-to-device communications underlying cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 1, pp. 81–99, Jan. 2015.

[37] Furqan Jameel, Zara Hamid, Farhana Jabeen, Sherali Zeadally, and Muhammad Awais Javed, "A survey of device-to-device communications: Research issues and challenges," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 2133–2168, Thirdquarter 2018.

[38] Jiangchuan Liu, Nei Kato, Jiaxin Ma, and Naoto Kadowaki, "Device-to-device communication in LTE-Advanced networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 1923 – 1940, Fourthquarter 2015.

[39] Melkamu Deressa, Min Sheng, Martin Wimmers, Junyu Liu, and Muluneh Mekonnen, "Maximizing quality of experience in device-to-device communication using an evolutionary algorithm based on users' behavior," *IEEE Access*, vol. 5, pp. 3878–3888, Mar. 2017.

[40] Ducheng Wu, Qihui Wu, Yuhua Xu, Jianjun Jing, and Zhiqiang Qin, "QoE-based distributed multichannel allocation in 5G heterogeneous cellular networks: A matching-coalitional game solution," *IEEE Access*, vol. 5, pp. 61–71, Sep. 2016.

[41] Shuan He and Wei Wang, "Context-aware QoE-price equilibrium for wireless multimedia relay communications using stackelberg game," in *Proc. of IEEE Conference on Computer Communications Workshops (INFOCOM)*, Atlanta, GA, USA, May 2017, pp. 506–511.

[42] Mehdi Bennis, Mérouane Debbah, and H Vincent Poor, "Ultra-reliable and low-latency wireless communication: Tail, risk and scale," *arXiv preprint arXiv:1801.01270*, 2018.

[43] Philipp Schulz, Maximilian Matthe, Henrik Klessig, Meryem Simsek, Gerhard Fettweis, Junaid Ansari, Shehzad Ali Ashraf, Bjoern Almeroth, Jens Voigt, Ines Riedel, et al., "Latency critical IoT applications in 5G: Perspective on the design of radio interface and network architecture," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 70–78, Feb. 2017.

[44] Hung-Yun Hsieh, Chih-Hua Chang, and Wei-Chih Liao, "Not every bit counts: Data-centric resource allocation for correlated data gathering in machine-to-machine wireless networks," *ACM Trans. on Sensor Networks (TOSN)*, vol. 11, no. 2, pp. 38:1–33, Feb. 2015.

[45] Zaher Dawy, Walid Saad, Arunabha Ghosh, Jeffrey G Andrews, and Elias Yaacoub, "Toward massive machine type cellular communications," *IEEE Wireless Communications*, vol. 24, no. 1, pp. 120–128, Feb. 2017.

[46] Tzu-Chuan Juan, Shih-En Wei, and Hung-Yun Hsieh, "Data-centric clustering for data gathering in machine-to-machine wireless networks," in *Proc. of IEEE International Conference on Communications (ICC) Workshops*, Budapest, Hungary, Jun. 2013, pp. 89–94.

[47] Hojin Song, Hung-Yun Hsieh, Yun-Da Tsai, and Wan Choi, "Correlation-aware machine selection for M2M data gathering in cellular networks," in *Proc. of IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Hong Kong, China, Sep. 2015, pp. 1184–1189.

[48] Mohammad Mozaffari, Walid Saad, Mehdi Bennis, and Mérouane Debbah, "Mobile Internet of Things: Can UAVs provide an energy-efficient mobile architecture?," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*, Washington, DC, USA, Dec. 2016, pp. 1–6.

[49] Yawei Pang, Yanru Zhang, Yunan Gu, Miao Pan, Zhu Han, and Pan Li, "Efficient data collection for wireless rechargeable sensor clusters in harsh

terrains using UAVs," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*, Austin, TX, USA, Dec. 2014, pp. 234–239.

[50] Mohammad Mozaffari, Walid Saad, Mehdi Bennis, and Mérouane Debbah, "Mobile unmanned aerial vehicles (UAVs) for energy-efficient internet of things communications," *IEEE Trans. on Wireless Communications*, vol. 16, no. 11, pp. 7574–7589, Nov. 2017.

[51] Cheng Zhan, Yong Zeng, and Rui Zhang, "Energy-efficient data collection in UAV enabled wireless sensor network," *IEEE Wireless Communications Letters*, vol. 7, no. 3, pp. 328–331, Jun. 2018.

[52] Akram Al-Hourani, Sithamparanathan Kandeepan, and Simon Lardner, "Optimal LAP altitude for maximum coverage," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 569–572, Dec. 2014.

[53] Giuseppe Durisi, Tobias Koch, and Petar Popovski, "Toward massive, ultra-reliable, and low-latency wireless communication with short packets," *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1711–1726, Sep. 2016.

[54] Changyang She, Chenyang Yang, and Tony QS Quek, "Radio resource management for ultra-reliable and low-latency communications," *IEEE Communications Magazine*, vol. 55, no. 6, pp. 72–78, Jun. 2017.

[55] Bikramjit Singh, Olav Tirkkonen, Zexian Li, and Mikko A Uusitalo, "Contention-based access for ultra-reliable low latency uplink transmissions," *IEEE Wireless Communications Letters*, vol. 7, no. 2, pp. 182–185, Jun. 2018.

[56] Hamidreza Shariatmadari, Sassan Iraji, Zexian Li, Mikko A Uusitalo, and Riku Jäntti, "Optimized transmission and resource allocation strategies for ultra-reliable communications," in *Proc. of IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Valencia, Spain, Sep. 2016, pp. 1–6.

[57] Chen-Feng Liu and Mehdi Bennis, "Ultra-reliable and low-latency vehicular transmission: An extreme value theory approach," *IEEE Communications Letters*, vol. 22, no. 6, pp. 1292–1295, Jun. 2018.

[58] Chen-Feng Liu, Mehdi Bennis, and H Vincent Poor, "Latency and reliability-aware task offloading and resource allocation for mobile edge computing," in *Proc. of IEEE Globecom Workshops (GC Wkshps)*, Singapore, Dec. 2017, pp. 1–7.

[59] Md Mostofa Kamal Tareq, Omid Semiari, Mohsen Amini Salehi, and Walid Saad, "Ultra reliable, low latency vehicle-to-infrastructure wireless communications with edge computing," *arXiv preprint arXiv:1808.06015*, 2018.

[60] David B Smith, Marius Portmann, Wee Lum Tan, and Wayes Tushar, "Multi-source–destination distributed wireless networks: Pareto-efficient dynamic power control game with rapid convergence," *IEEE Trans. on Vehicular Technology*, vol. 63, no. 6, pp. 2744–2754, Jul. 2014.

[61] Petyon Young and Shmuel Zamir, *Handbook of game theory*, vol. 4, Elsevier, 2014.

[62] Edoardo Amaldi and Viggo Kann, "On the approximability of minimizing nonzero variables or unsatisfied relations in linear systems," *Theoretical Computer Science*, vol. 209, pp. 237–260, Dec. 1998.

[63] Kevin Leyton-Brown and Yoav Shoham, "Essentials of game theory: A concise multidisciplinary introduction," *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 2, no. 1, pp. 1–88, 2008.

[64] Mark Felegyhazi and Jean-Pierre Hubaux, "Game theory in wireless networks: A tutorial," ., 2006.

[65] Lingyang Song, Dusit Niyato, Zhu Han, and Ekram Hossain, "Game-theoretic resource allocation methods for device-to-device communication," *IEEE Wireless Communications*, vol. 21, no. 3, pp. 136–144, Jun. 2014.

[66] Drew Fudenberg and Jean Tirole, *Game Theory*, MIT press, 1991.

[67] Walid Saad, Zhu Han, Mérouane Debbah, Are Hjørungnes, and Tamer Başar, "Coalitional game theory for communication networks," *IEEE Signal Processing Magazine*, vol. 26, no. 5, pp. 77–97, Sep. 2009.

[68] Tamer Basar and Geert Jan Olsder, *Dynamic noncooperative game theory*, vol. 23, Siam, 1999.

[69] Abay Molla Kassa and Semu Mitiku Kassa, "Deterministic solution approach for some classes of nonlinear multilevel programs with multiple followers," *Journal of Global Optimization*, vol. 68, no. 4, pp. 729–747, Aug. 2017.

[70] Baoding Liu, "Stackelberg-nash equilibrium for multilevel programming with multiple followers using genetic algorithms," *Computers & Mathematics with Applications*, vol. 36, no. 7, pp. 79–89, Feb. 1998.

[71] Hamidou Tembine, Eitan Altman, Rachid El-Azouzi, and Yezekael Hayel, "Evolutionary games in wireless networks," *IEEE Trans. on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 40, no. 3, pp. 634–646, Jun. 2010.

[72] Anna Bogomolnaia and Matthew O Jackson, "The stability of hedonic coalition structures," *Games and Economic Behavior*, vol. 38, no. 2, pp. 201–230, 2002.

[73] Jacques H Dreze and Joseph Greenberg, "Hedonic coalitions: Optimality and stability," *Econometrica: Journal of the Econometric Society*, vol. 48, no. 4, pp. 987–1003, May 1980.

[74] Siavash Bayat, Yonghui Li, Lingyang Song, and Zhu Han, "Matching theory: Applications in wireless communications," *IEEE Signal Processing Magazine*, vol. 33, no. 6, pp. 103–122, 2016.

[75] Yunan Gu, Walid Saad, Mehdi Bennis, Merouane Debbah, and Zhu Han, "Matching theory for future wireless networks: fundamentals and applications," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 52–59, May 2015.

[76] Nicole Sawyer and David B Smith, "Flexible resource allocation in device-to-device communications using stackelberg game theory," *IEEE Trans. on Communications*, vol. 67, no. 1, pp. 653–667, Jan. 2019.

[77] Nicole Sawyer and David B Smith, "Pareto-efficient cross-layer repeated game for device-to-device (D2D) communications," in *Proc. of IEEE International Conference on Communications (ICC)*, Kuala Lumpur, Malaysia, May 2016, pp. 1–6.

[78] Nicole Sawyer and David B Smith, "A Nash stable cross-layer coalition formation game for device-to-device communications," in *Proc. of IEEE International Conference on Communications (ICC)*, Paris, France, May 2017, pp. 1–6.

[79] Nicole Sawyer and David B Smith, "A nash stable cross-layer coalitional game for resource utilization in device-to-device communications," *IEEE Trans. on Vehicular Technology*, vol. 67, no. 9, pp. 8608–8622, Sep. 2018.

[80] Nicole Sawyer, Mehdi Naderi Soorki, Walid Saad, and David B. Smith, "Evolutionary coalitional game for correlation-aware clustering in machine-to-machine communications," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*, Singapore, Dec. 2017, pp. 1–6.

[81] Nicole Sawyer, Mehdi Naderi Soorki, Walid Saad, and David B Smith, "Evolutionary games for correlation-aware clustering in massive machine-to-machine networks," *IEEE Trans. on Communications*, May 2019.

[82] Nicole Sawyer, Mehdi Naderi Soorki, Walid Saad, Ni Ding, David B Smith, Mohammad Hossein Manshaei, and Mohammad Mozaffari, "Energy-aware scheduling for ultra-reliable internet of things communications using unmanned aerial vehicles," *under preparation to be submitted to IEEE Internet of Things Journal*, 2019.

[83] Phond Phunchongharn, Ekram Hossain, and Dong In Kim, "Resource allocation for device-to-device communications underlaying LTE-Advanced networks," *IEEE Wireless Communications*, vol. 20, no. 4, pp. 91–100, Aug. 2013.

[84] Xingqin Lin, Jeffrey G Andrews, and Amitava Ghosh, "Spectrum sharing for device-to-device communication in cellular networks," *IEEE Trans. on Wireless Communications*, vol. 13, no. 12, pp. 6727–6740, Dec. 2014.

[85] M Belleschi, G Fodor, DD Penda, M Johansson, and A Abrardo, "A joint power control and resource allocation algorithm for D2D communications," *KTH, School of Electrical Engineering (EES), Automatic Control, Tech. Rep*, 2012.

[86] Hesham ElSawy, Ekram Hossain, and Mohamed-Slim Alouini, "Analytical modeling of mode selection and power control for underlay D2D communication in cellular networks," *IEEE Trans. on Communications*, vol. 62, no. 11, pp. 4147–4161, Nov. 2014.

[87] Yi Li, M Cenk Gursoy, Senem Velipasalar, and Jian Tang, "Joint mode selection and resource allocation for D2D communications via vertex coloring," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*, Singapore, Dec. 2017, pp. 1–6.

[88] Klaus Doppler, Chia-Hao Yu, Cassio B Ribeiro, and Pekka Janis, "Mode selection for device-to-device communication underlaying an LTE-advanced network," in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2010, pp. 1–6.

[89] Chia-Hao Yu, Klaus Doppler, Cassio B Ribeiro, and Olav Tirkkonen, "Resource sharing optimization for device-to-device communication underlaying cellular networks," *IEEE Trans. on Wireless communications*, vol. 10, no. 8, pp. 2752–2763, Aug. 2011.

[90] Setareh Maghsudi and Slawomir Stanczak, "Transmission mode selection for network-assisted device to device communication: A levy-bandit approach," in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 7009–7013.

[91] Yueming Cai, Hualiang Chen, Dan Wu, Wendong Yang, and Liang Zhou, "A distributed resource management scheme for D2D communications based on coalition formation game," in *Proc. of IEEE International Conference on Communications (ICC) Workshops*, Sydney, NSW, Australia, Jun. 2014, pp. 355–359.

[92] Hualiang Chen, Dan Wu, and Yueming Cai, "Coalition formation game for green resource management in D2D communications," *IEEE Communications Letters*, vol. 18, no. 8, pp. 1395–1398, Aug. 2014.

[93] Khajonpong Akkarajitsakul, Phond Phunchongharn, Ekram Hossain, and Vijay K Bhargava, "Mode selection for energy-efficient D2D communications in

LTE-Advanced networks: A coalitional game approach," in *Proc. of IEEE International Conference on Communication Systems (ICCS)*, Singapore, Nov. 2012, pp. 488–492.

[94] Dan Wu, Yueming Cai, Rose Qingyang Hu, and Yi Qian, "Dynamic distributed resource sharing for mobile D2D communications," *IEEE Trans. on Wireless Communications*, vol. 14, no. 10, pp. 5417–5429, Oct. 2015.

[95] Dan Wu, Jinlong Wang, Rose Qingyang Hu, Yueming Cai, and Liang Zhou, "Energy-efficient resource sharing for mobile device-to-device multimedia communications," *IEEE Trans. on Vehicular Technology*, vol. 63, no. 5, pp. 2093–2103, Jun. 2014.

[96] Peng Cheng, Lei Deng, Hui Yu, Youyun Xu, and Hailong Wang, "Resource allocation for cognitive networks with D2D communication: An evolutionary approach," in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*, Shanghai, China, Apr. 2012, pp. 2671–2676.

[97] Qiaoyang Ye, Mazin Al-Shalash, Constantine Caramanis, and Jeffrey G Andrews, "Distributed resource allocation in device-to-device enhanced cellular networks," *IEEE Trans. on Communications*, vol. 63, no. 2, pp. 441–454, Feb. 2015.

[98] Rongqing Zhang, Lingyang Song, Zhu Han, Xiang Cheng, and Bingli Jiao, "Distributed resource allocation for device-to-device communications underlaying cellular networks," in *Proc. of IEEE International Conference on Communications (ICC)*, Budapest, Hungary, Jun. 2013, pp. 1889–1893.

[99] Yong Li, Depeng Jin, Jian Yuan, and Zhu Han, "Coalitional games for resource allocation in the device-to-device uplink underlaying cellular networks," *IEEE Trans. on Wireless Communications*, vol. 13, no. 7, pp. 3965–3977, Jul. 2014.

[100] Saikiran Bulusu, Neelesh B Mehta, and Suresh Kalyanasundaram, "Rate adaptation, scheduling, and mode selection in D2D systems with partial channel knowledge," *IEEE Trans. on Wireless Communications*, vol. 17, no. 2, pp. 1053–1065, Feb. 2018.

[101] Haijun Zhang, Chunxiao Jiang, Norman C Beaulieu, Xiaoli Chu, Xianbin Wang, and Tony QS Quek, "Resource allocation for cognitive small cell networks: A cooperative bargaining game theoretic approach," *IEEE Trans. on Wireless Communications*, vol. 14, no. 6, pp. 3481–3493, Jun. 2015.

[102] Ajmery Sultana, Lian Zhao, and Xavier Fernando, "Efficient resource allocation in device-to-device communication using cognitive radio technology," *IEEE Trans. on Vehicular Technology*, vol. 66, no. 11, pp. 10024–10034, Nov. 2017.

[103] Si Wen, Xiaoyue Zhu, Zhesheng Lin, Xin Zhang, and Dacheng Yang, "Energy efficient power allocation schemes for device-to-device (D2D) communication," in *Proc. of IEEE 78th Vehicular Technology Conference (VTC Fall)*, Las Vegas, NV, USA, Sep. 2013, pp. 1–5.

[104] Rui Yin, Caijun Zhong, Guanding Yu, Zhaoyang Zhang, Kai Kit Wong, and Xiaoming Chen, "Joint spectrum and power allocation for D2D communications underlaying cellular networks," *IEEE Trans. on Vehicular Technology*, vol. 65, no. 4, pp. 2182–2195, Apr. 2016.

[105] Li Ping Qian, Shengli Zhang, Wei Zhang, and Ying Jun Angela Zhang, "Interference-aware system utility maximization for cognitive radio networks," in *Proc. of IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Atlanta, GA, USA, Dec. 2014, pp. 1160–1164.

[106] Lu Yang and Wei Zhang, "Interference coordination in device-to-device communication," in *Interference Coordination for 5G Cellular Networks*, pp. 31–51. Springer, 2015.

[107] Michael Maskery, Vikram Krishnamurthy, and Qing Zhao, "Decentralized dynamic spectrum access for cognitive radios: cooperative design of a non-cooperative game," *IEEE Trans. on Communications*, vol. 57, no. 2, pp. 459–469, Feb. 2009.

[108] Chunmei Xia, Shaoyi Xu, and Kyung Sup Kwak, "Resource allocation for device-to-device communication in LTE-A network: A Stackelberg game approach," in *Proc. of IEEE 80th Vehicular Technology Conference (VTC Fall)*, Vancouver, BC, Canada, Sep. 2014, pp. 1–5.

[109] Guodong Zhang, Jinming Hu, Wei Heng, Xiang Li, and Gang Wang, "Distributed power control for D2D communications underlaying cellular network using stackelberg game," in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*, San Francisco, CA, USA, Mar. 2017, pp. 1–6.

[110] Yan Han, Xiaofeng Tao, and Xuefei Zhang, "Power allocation for device-to-device underlay communication with femtocell using stackelberg game," in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*, Barcelona, Spain, Apr. 2018, pp. 1–6.

[111] Yanxiang Jiang, Hui Ge, Mehdi Bennis, Fu-Chun Zheng, and Xiaohu You, "Power control via stackelberg game for small-cell networks," *arXiv preprint arXiv:1802.04775*, 2018.

[112] Feiran Wang, Lingyang Song, Zhu Han, Qun Zhao, and Xiaoli Wang, "Joint scheduling and resource allocation for device-to-device underlay communication," in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*, Shanghai, China, Apr. 2013, pp. 134–139.

[113] Gábor Fodor and Norbert Reider, "A distributed power control scheme for cellular network assisted D2D communications," in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM)*, Kathmandu, Nepal, Dec. 2011, pp. 1–6.

[114] Gabor Fodor, Demia Della Penda, Marco Belleschi, Mikael Johansson, and Andrea Abrardo, "A comparative study of power control approaches for device-to-device communications," in *Proc. of IEEE International Conference on Communications (ICC)*, Budapest, Hungary, Jun. 2013, pp. 6008–6013.

[115] Qijie Wang, Wei Wang, Seongwook Jin, Hengliang Zhu, and Nai Tong Zhang, "Quality-optimized joint source selection and power control for wireless multimedia D2D communication using Stackelberg game," *IEEE Trans. on Vehicular Technology*, vol. 64, no. 8, pp. 3755–3769, Aug. 2015.

[116] Zheng Chang, Yang Hu, Yan Chen, and Bing Zeng, "Cluster-oriented device-to-device multimedia communications: Joint power, bandwidth, and link se-

lection optimization," *IEEE Trans. on Vehicular Technology*, vol. 67, no. 2, pp. 1570–1581, Feb. 2018.

[117] Foad Hajiaghajani, Ramtin Davoudi, and Mehdi Rasti, "A joint channel and power allocation scheme for device-to-device communications underlaying uplink cellular networks," in *Proc. of IEEE Conference on Computer Communications Workshops (INFOCOM)*, San Francisco, CA, USA, Apr. 2016, pp. 768–773.

[118] Yuan Liu, Rui Wang, and Zhu Han, "Interference-constrained pricing for D2D networks," *IEEE Trans. on Wireless Communications*, vol. 16, no. 1, pp. 475–486, Jan. 2017.

[119] Susan Dominic and Lillykutty Jacob, "Distributed resource allocation for D2D communications underlaying cellular networks in time-varying environment," *IEEE Communications Letters*, vol. 22, no. 2, pp. 388–391, Feb. 2018.

[120] Shama Noreen, Navrati Saxena, and Abhishek Roy, "Discount interference pricing mechanism for data offloading in D2D communications.," *IEEE Communications Letters (Early Access)*, Jun. 2018.

[121] Namyoon Lee, Xingqin Lin, Jeffrey G Andrews, and Robert W Heath, "Power control for D2D underlaid cellular networks: Modeling, algorithms, and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 1, pp. 1–13, Jan. 2015.

[122] Mahmoud A Abo-Sinna and Ibrahim A Baky, "Interactive balance space approach for solving multi-level multi-objective programming problems," *Information sciences*, vol. 177, no. 16, pp. 3397–3410, Aug. 2007.

[123] Ming Han, Ping Wang, Hui Fang Pang, Fu Qiang Liu, Xin Hong Wang, and Ngoc Van Nguyen, "User satisfaction based resource allocation using artificial fish swarm algorithm for D2D communication," in *Sensors, Measurement and Intelligent Materials II*, Mar. 2014, vol. 475-476 of *Applied Mechanics and Materials*, pp. 885–892.

[124] Mingkai Chen, Lei Wang, Jianxin Chen, and Xin Wei, "QoE-driven D2D media services distribution scheme in cellular networks," *Wireless Communications and Mobile Computing*, vol. 2017, Jul. 2017.

[125] Xinzhou Wu, Saurabha Tavildar, Sanjay Shakkottai, Tom Richardson, Junyi Li, Rajiv Laroia, and Aleksandar Jovicic, "FlashLinQ: A synchronous distributed scheduler for peer-to-peer ad hoc networks," *IEEE/ACM Trans. on Networking (TON)*, vol. 21, no. 4, pp. 1215–1228, Aug. 2013.

[126] Navid Naderializadeh and Amir Salman Avestimehr, "ITLinQ: A new approach for spectrum sharing in device-to-device communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1139–1151, Jun. 2014.

[127] Sueng Jae Bae, Dong Hyun Kim, Bum-Gon Choi, and Min Young Chung, "Transmission power control for FlashLinQ device-to-device communication system," in *Proc. of IEEE 77th Vehicular Technology Conference (VTC Spring)*, Dresden, Germany, Jun. 2013, pp. 1–5.

[128] Xinping Yi and Giuseppe Caire, "ITLinQ+: An improved spectrum sharing mechanism for device-to-device communications," in *Proc. of 49th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, Nov. 2015, pp. 1310–1314.

[129] S Ali Hesammohseni and Mohamed Oussama Damen, "Provably near optimal link scheduling and power control for wireless device-to-device networks," in *Proc. of Australian Communications Theory Workshop (AusCTW)*, Melbourne, VIC, Australia, Jan. 2016, pp. 47–52.

[130] Kaiming Shen and Wei Yu, "FPLinQ: A cooperative spectrum sharing strategy for device-to-device communications," in *Proc. of IEEE International Symposium on Information Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 2323–2327.

[131] Mehrdad Kiamari, Chenwei Wang, A Salman Avestimehr, and Haralabos Papadopoulos, "SINR-threshold scheduling with binary power control for D2D networks," *arXiv preprint arXiv:1708.09596*, 2017.

[132] Si Wen, Xiaoyue Zhu, Xin Zhang, and Dacheng Yang, "QoS-aware mode selection and resource allocation scheme for device-to-device (D2D) communication in cellular networks," in *Proc. of IEEE International Conference on Communications Workshops (ICC)*, Budapest, Hungary, Jun. 2013, pp. 101–105.

[133] Zhiyong Du, Dong Liu, and Lijie Yin, "User in the loop: QoE-oriented optimization in communication and networks," in *Proc. of 6th International Conference on Computer Science and Network Technology (ICCSNT)*, Dalian, China, Oct. 2017, pp. 420–424.

[134] Rainer Schoenen, Gurhan Bulu, Amir Mirtaheri, Tamer Beitelmal, and Halim Yanikomeroglu, "Quantified user behavior in user-in-the-loop spatially and demand controlled cellular systems," in *Proc. of 18th European Wireless Conference*, Poznan, Poland, Apr. 2012, pp. 1–8.

[135] Rainer Schoenen and Halim Yanikomeroglu, "User-in-the-loop: Spatial and temporal demand shaping for sustainable wireless networks," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 196–203, Feb. 2014.

[136] Weicheng Zhang, Yaodong Li, Hai Lu, Xuemin Hong, and Jianghong Shi, "User-in-the-loop content delivery in cellular communication networks with heterogeneous user behaviors.," *Applied Sciences (2076-3417)*, vol. 8, no. 5, pp. 1–18, May 2018.

[137] Ziyang Wang, Rainer Schoenen, Halim Yanikomeroglu, and Marc St-Hilaire, "Load balancing in cellular networks with user-in-the-loop: A spatial traffic shaping approach," in *Proc. of IEEE International Conference on Communications (ICC)*, London, UK, Jun. 2015, pp. 2638–2643.

[138] Rainer Schoenen and Halim Yanikomeroglu, "Dynamic demand control with differentiated QoS in user-in-the-loop controlled cellular networks," in *Proc. of IEEE 77th Vehicular Technology Conference (VTC Spring)*, Dresden, Germany, Jun. 2013, pp. 1–6.

[139] Mehdi Neshat, Ghodrat Sepidnam, Mehdi Sargolzaei, and Adel Najaran Toosi, "Artificial fish swarm algorithm: A survey of the state-of-the-art,

hybridization, combinatorial and indicative applications," *Artificial Intelligence Review*, vol. 42, no. 4, pp. 965–997, Dec. 2014.

[140] Guowang Miao, Amin Azari, and Taewon Hwang, "E^2-MAC: Energy efficient medium access for massive M2M communications," *IEEE Trans. on Communications*, vol. 64, no. 11, pp. 4720–4735, Nov. 2016.

[141] Hyun-Kwan Lee, Dong Min Kim, Youngju Hwang, Seung Min Yu, and Seong-Lyun Kim, "Feasibility of cognitive machine-to-machine communication using cellular bands," *IEEE Wireless Communications*, vol. 20, no. 2, pp. 97–103, Apr. 2013.

[142] Mehdi Naderi Soorki, Walid Saad, Mohammad Hossein Manshaei, and Hossein Saidi, "Stochastic coalitional games for cooperative random access in M2M communications," *IEEE Trans. on Wireless Communications*, vol. 16, no. 9, pp. 6179–6192, Sep. 2017.

[143] Chieh Yuan Ho and Ching-Yao Huang, "Energy-saving massive access control and resource allocation schemes for M2M communications in OFDMA cellular networks," *IEEE Wireless Communications Letters*, vol. 1, no. 3, pp. 209–212, Jun. 2012.

[144] Bilal R Al-Kaseem, Andrews O Nyanteh, and Hamed S Al-Raweshidy, "Self-organized clustering technique based on sink mobility in heterogeneous M2M sensor networks," in *Proc. of International Conference for Students on Applied Engineering (ICSAE)*, Newcastle upon Tyne, UK, Oct. 2016, pp. 431–436.

[145] Shih-En Wei, Hung-Yun Hsieh, and Hsuan-Jung Su, "Joint optimization of cluster formation and power control for interference-limited machine-to-machine communications," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*, Anaheim, CA, USA, Dec. 2012, pp. 5512–5518.

[146] Fatima Hussain, Alagan Anpalagan, and Muhammad Naeem, "Multi-objective MTC device controller resource optimization in M2M communication," in *Proc. of 27th Biennial Symposium on Communications (QBSC)*, Kingston, ON, Canada, Jun. 2014, pp. 184–188.

[147] Hashim Safdar, Norsheila Fisal, Rahat Ullah, Wajahat Maqbool, Faiz Asraf, Zubair Khalid, and AS Khan, "Resource allocation for uplink M2M communication: A game theory approach," in *Proc. of IEEE Symposium on Wireless Technology and Applications (ISWTA)*, Kuching, Malaysia, Sep. 2013, pp. 48–52.

[148] Siavash Bayat, Yonghui Li, Zhu Han, Mischa Dohler, and Branka Vucetic, "Distributed data aggregation in machine-to-machine communication networks based on coalitional game," in *Proc. of Wireless Communications and Networking Conference (WCNC)*, Istanbul, Turkey, Apr. 2014, pp. 2026–2031.

[149] Nof Abuzainab, Walid Saad, Choong-Seon Hong, and H Vincent Poor, "Cognitive hierarchy theory for distributed resource allocation in the Internet of Things," *IEEE Trans. on Wireless Communications*, vol. 16, no. 12, pp. 7687–7702, Dec. 2017.

[150] Hung-Yun Hsieh, Tzu-Chuan Juan, Yun-Da Tsai, and Hong-Chen Huang, "Minimizing radio resource usage for machine-to-machine communications through data-centric clustering," *IEEE Trans. on Mobile Computing*, vol. 15, no. 12, pp. 3072–3086, Dec. 2016.

[151] Walid Saad, Zhu Han, Mérouane Debbah, and Are Hjorungnes, "A distributed coalition formation framework for fair user cooperation in wireless networks," *IEEE Trans. on wireless communications*, vol. 8, no. 9, pp. 4580–4593, Sep. 2009.

[152] Chih-Hua Chang, Wei-Chih Liao, Hung-Yun Hsieh, and Hsuan-Jung Su, "Not every bit counts: Shifting the focus from machine to data for machine-to-machine communications," in *Proc. of Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, Nov. 2012, pp. 581–585.

[153] Kenza Hamidouche, Walid Saad, and Mérouane Debbah, "Popular matching games for correlation-aware resource allocation in the internet of things," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*, Singapore, Dec. 2017, pp. 1–8.

[154] Elham Kalantari, Halim Yanikomeroglu, and Abbas Yongacoglu, "On the number and 3D placement of drone base stations in wireless cellular networks," in *Proc. of IEEE 84th Vehicular Technology Conference (VTC-Fall)*, Montreal, QC, Canada, Sep. 2016, pp. 1–6.

[155] Arman Azizi, Nader Mokari, and Mohamad Reza Javan, "Joint radio resource allocation, 3D placement and user association of aerial base stations in IoT networks," *arXiv preprint arXiv:1710.05315*, 2017.

[156] Qingqing Wu, Yong Zeng, and Rui Zhang, "Joint trajectory and communication design for multi-UAV enabled wireless networks," *IEEE Trans. on Wireless Communications*, vol. 17, no. 3, pp. 2109–2121, Mar. 2018.

[157] Yong Zeng, Rui Zhang, and Teng Joon Lim, "Throughput maximization for UAV-enabled mobile relaying systems," *IEEE Trans. on Communications*, vol. 64, no. 12, pp. 4983–4996, Dec. 2016.

[158] Yong Zeng and Rui Zhang, "Energy-efficient UAV communication with trajectory optimization," *IEEE Trans. Wireless Commun*, vol. 16, no. 6, pp. 3747–3760, Jun. 2017.

[159] Sotheara Say, Hikari Inata, Jiang Liu, and Shigeru Shimamoto, "Priority-based data gathering framework in UAV-assisted wireless sensor networks," *IEEE Sensors Journal*, vol. 16, no. 14, pp. 5785–5794, Jul. 2016.

[160] Changyang She, Chenxi Liu, Tony QS Quek, Chenyang Yang, and Yonghui Li, "UAV-assisted uplink transmission for ultra-reliable and low-latency communications," in *Proc. of IEEE International Conference on Communications Workshops (ICC Workshops)*, Kansas City, MO, USA, May 2018, pp. 1–6.

[161] Yue Li and Lin Cai, "UAV-assisted dynamic coverage in a heterogeneous cellular system," *IEEE Network*, vol. 31, no. 4, pp. 56–61, Jul. 2017.

[162] Zeng-you Sun and Dong-na Yang, "A D2D wireless resource allocation scheme based on overall fairness," *3D Research*, vol. 10, no. 2, pp. 1–10, Jun. 2019.

[163] Xuejia Cai, Jun Zheng, and Yuan Zhang, "A graph-coloring based resource allocation algorithm for D2D communication in cellular networks," in *Proc. of IEEE International Conference on Communications (ICC)*, London, UK, Jun. 2015, pp. 5429–5434.

[164] Haesik Kim, "Coding and modulation techniques for high spectral efficiency transmission in 5G and satcom," in *Proc. of 23rd European Signal Processing Conference (EUSIPCO)*, Nice, France, Sep. 2015, pp. 2746–2750.

[165] HM Savitha and Muralidhar Kulkarni, "Performance evaluation of turbo coded OFDM systems and application of turbo decoding for impulsive channel," *ICTACT International Journal on Communication Technology*, vol. 1, no. 3, pp. 175–183, Sep. 2010.

[166] Jeongseok Ha, Apurva N Mody, Joon H Sung, John R Barry, Steven W McLaughlin, and Gordon L Stüber, "LDPC coded OFDM with alamouti/SVD diversity technique," *Wireless Personal Communications*, vol. 23, no. 1, pp. 183–194, Oct. 2002.

[167] Sarah Koskie and Zoran Gajic, "A Nash game algorithm for SIR-based power control in 3G wireless CDMA networks," *ACM Trans. on Networking (TON)*, vol. 13, no. 5, pp. 1017–1026, Oct. 2005.

[168] Yichao Chen, Shibo He, Fen Hou, Zhiguo Shi, and Xu Chen, "Optimal user-centric relay assisted device-to-device communications: An auction approach," *Institution of Engineering and Technology (IET) Communications*, vol. 9, no. 3, pp. 386–395, Feb. 2015.

[169] Peter Marbach and Randall Berry, "Downlink resource allocation and pricing for wireless networks," in *Proc. of IEEE International Conference on Computer Communications (INFOCOM)*, New York, NY, USA, Jun. 2002, pp. 1470–1479.

[170] Cem U Saraydar, Narayan B Mandayam, and David J Goodman, "Efficient power control via pricing in wireless data networks," *IEEE Trans. on Communications*, vol. 50, no. 2, pp. 291–303, Feb. 2002.

[171] J Ben Rosen, "Existence and uniqueness of equilibrium points for concave N-person games," *Econometrica: Journal of the Econometric Society*, vol. 33, no. 3, pp. 520–534, Jul. 1965.

[172] Godfrey Keller, Sven Rady, and Martin Cripps, "Strategic experimentation with exponential bandits," *Econometrica*, vol. 73, no. 1, pp. 39–68, 2005.

[173] Sébastien Bubeck, Nicolo Cesa-Bianchi, et al., "Regret analysis of stochastic and nonstochastic multi-armed bandit problems," *Foundations and Trends® in Machine Learning*, vol. 5, no. 1, pp. 1–122, 2012.

[174] Aleksandrs Slivkins, "Introduction to multi-armed bandits," *arXiv preprint arXiv:1904.07272*, 2019.

[175] Suryapratim Banerjee, Hideo Konishi, and Tayfun Sönmez, "Core in a simple coalition formation game," *Social Choice and Welfare*, vol. 18, no. 1, pp. 135–153, Jan. 2001.

[176] Gönül Doğan, Marcel ALM Van Assen, Arnout Van de Rijt, and Vincent Buskens, "The stability of exchange networks," *Social Networks*, vol. 31, no. 2, pp. 118–125, May 2009.

[177] Haris Aziz and Bart De Keijzer, "Complexity of coalition structure generation," in *Proc. of The 10th International Conference on Autonomous Agents and Multiagent Systems - Volume 1*, Taipei, Taiwan, May 2011, pp. 191–198.

[178] Martin Haenggi, *Stochastic geometry for wireless networks*, Cambridge University Press, Cambridge, UK, 2012.

[179] Thomas M Cover and Joy A Thomas, *Elements of information theory*, John Wiley & Sons, 2012.

[180] Sung Nok Chiu, Dietrich Stoyan, Wilfrid S Kendall, and Joseph Mecke, *Stochastic geometry and its applications*, John Wiley & Sons, Hoboken, NJ, USA, 2013.

[181] Mehrnaz Afshang and Harpreet S Dhillon, "Fundamentals of modeling finite wireless networks using binomial point process.," *IEEE Trans. Wireless Communications*, vol. 16, no. 5, pp. 3355–3370, May 2017.

[182] Shin-Yeh Tsai, Sok-Ian Sou, and Meng-Hsun Tsai, "Reducing energy consumption by data aggregation in M2M networks," *Wireless personal communications*, vol. 74, no. 4, pp. 1231–1244, Feb. 2014.

[183] Halil Yetgin, Kent Tsz Kan Cheung, Mohammed El-Hajjar, and Lajos Hanzo Hanzo, "A survey of network lifetime maximization techniques in wireless sensor networks," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 828–854, Secondquarter 2017.

[184] Ossama Younis and Sonia Fahmy, "Heed: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Trans. on mobile computing*, vol. 3, no. 4, pp. 366–379, 2004.

[185] Li Qing, Qingxin Zhu, and Mingwen Wang, "Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks," *Computer communications*, vol. 29, no. 12, pp. 2230–2237, 2006.

[186] Samad Ali, Nandana Rajatheva, and Walid Saad, "Fast uplink grant for machine type communications: Challenges and opportunities," *arXiv preprint arXiv:1801.04953*, 2018.

[187] Christian Hoymann, David Astely, Magnus Stattin, Gustav Wikstrom, Jung-Fu Cheng, Andreas Hoglund, Mattias Frenne, Ricardo Blasco, Joerg Huschke, and Fredrik Gunnarsson, "LTE release 14 outlook," *IEEE Communications Magazine*, vol. 54, no. 6, pp. 44–49, 2016.

[188] 3rd Generation Partnership Project, "Evolved universal terrestrial radio access (e-utra); study on latency reduction techniques for LTE (release 14)," *3GPP TR 36.881*, 2016.

[189] Sundeep Pattem, Bhaskar Krishnamachari, and Ramesh Govindan, "The impact of spatial correlation on routing with compression in wireless sensor networks," *ACM Trans. on Sensor Networks (TOSN)*, vol. 4, no. 4, pp. 24:1–33, Aug. 2008.

[190] Sunil Srinivasa and Martin Haenggi, "Distance distributions in finite uniformly random networks: Theory and applications," *IEEE Trans. on Vehicular Technology*, vol. 59, no. 2, pp. 940–949, Feb. 2010.

[191] Manzoor Ahmed Khan and Hamidou Tembine, "Evolutionary coalitional games in network selection," in *Proc. of Wireless Advanced (WiAd)*, London, United Kingdom, Jun. 2011, pp. 185–194.

[192] Xin Luo and Hamidou Tembine, "Evolutionary coalitional games for random access control," in *Proc. of IEEE International Conference on Computer Communications (INFOCOM)*, Turin, Itally, Apr. 2013, pp. 535–539.

[193] Tadeusz Płatkowski, "Evolutionary coalitional games," *Dynamic Games and Applications*, vol. 6, no. 3, pp. 396–408, Sep. 2016.

[194] Tingsong Jiang, Xu Tan, Xi Luan, Xiaoning Zhang, and Jianjun Wu, "Evolutionary game based access class barring for machine-to-machine communications," in *Proc. of 16th International Conference on Advanced Communication Technology (ICACT)*, Pyeongchang, South Korea, Feb. 2014, pp. 832–835.

[195] Javad Salimi Sartakhti, Mohammad Hossein Manshaei, and Mehdi Sadeghi, "MMP–TIMP interactions in cancer invasion: An evolutionary game-theoretical framework," *Journal of theoretical biology*, vol. 412, pp. 17–26, Jan. 2017.

[196] Dusit Niyato and Ekram Hossain, "Dynamics of network selection in heterogeneous wireless networks: An evolutionary game approach," *IEEE Trans. on Vehicular Technology*, vol. 58, no. 4, pp. 2008–2017, May 2009.

[197] J Maynard Smith, "The theory of games and the evolution of animal conflicts," *Journal of theoretical biology*, vol. 47, no. 1, pp. 209–221, Sep. 1974.

[198] J Maynard Smith and George R Price, "The logic of animal conflict," *Nature*, vol. 246, no. 5427, pp. 15, Nov. 1973.

[199] Zhu-Jun Yang, Jie-Cheng Huang, Chun-Ting Chou, Hung-Yun Hsieh, Chin-Wei Hsu, Ping-Cheng Yeh, and Chia-Chun Alex Hsu, "Peer discovery for device-to-device (D2D) communication in LTE-A networks," in *Proc. of IEEE Globecom Workshops (GC Wkshps)*, Atlanta, GA, USA, Dec. 2013, pp. 665–670.

[200] Andrea Goldsmith, *Wireless communications*, Cambridge university press, 2005.

[201] Michael J Neely, "Stochastic network optimization with application to communication and queueing systems," *Synthesis Lectures on Communication Networks*, vol. 3, no. 1, pp. 1–211, 2010.

[202] Longbo Huang and Michael J Neely, "Utility optimal scheduling in energy-harvesting networks," *IEEE/ACM Trans. on Networking (TON)*, vol. 21, no. 4, pp. 1117–1130, 2013.

[203] Alvin E Roth et al., "Two-sided matching with incomplete information about others preferences," *Games and Economic Behavior*, vol. 1, no. 2, pp. 191–209, 1989.

[204] David Gale and Lloyd S Shapley, "College admissions and the stability of marriage," *The American Mathematical Monthly*, vol. 69, no. 1, pp. 9–15, Jan. 1962.

[205] Eduard A Jorswieck, "Stable matchings for resource allocation in wireless networks," in *Proc. of 17th International Conference on Digital Signal Processing (DSP)*, Corfu, Greece, Jul. 2011, pp. 1–8.

[206] Omid Semiari, Walid Saad, and Mehdi Bennis, "Joint millimeter wave and microwave resources allocation in cellular networks with dual-mode base stations," *IEEE Trans. on Wireless Communications*, vol. 16, no. 7, pp. 4802–4816, Jul. 2017.

[207] Jingjing Zhao, Yuanwei Liu, Kok Keong Chai, Maged Elkashlan, and Yue Chen, "Matching with peer effects for context-aware resource allocation in d2d communications," *IEEE Communications Letters*, vol. 21, no. 4, pp. 837–840, Apr. 2017.

[208] Boya Di, Lingyang Song, and Yonghui Li, "Sub-channel assignment, power allocation, and user scheduling for non-orthogonal multiple access networks," *IEEE Trans. on Wireless Communications*, vol. 15, no. 11, pp. 7686–7698, Nov. 2016.

[209] Tachporn Sanguanpuak, Sudarshan Guruacharya, Nandana Rajatheva, Mehdi Bennis, and Matti Latva-Aho, "Multi-operator spectrum sharing for small cell networks: A matching game perspective," *IEEE Trans. on Wireless Communications*, vol. 16, no. 6, pp. 3761–3774, Jun. 2017.

[210] Sarder Fakhrul Abedin, Md Golam Rabiul Alam, SM Ahsan Kazmi, Nguyen H Tran, Dusit Niyato, and Choong Seon Hong, "Resource allocation for ultra-reliable and enhanced mobile broadband IoT applications in fog network," *IEEE Trans. on Communications*, vol. 67, no. 1, pp. 489–502, Jan. 2019.

[211] Daniel Verenzuela and Guowang Miao, "Scalable d2d communications for frequency reuse>> 1 in 5G," *IEEE Trans. on Wireless Communications*, vol. 16, no. 6, pp. 3435–3447, Jun. 2017.

[212] Mamta Agiwal, Abhishek Roy, and Navrati Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1617–1655, Thirdquarter 2016.

[213] Haichao Wang, Guoru Ding, Feifei Gao, Jin Chen, Jinlong Wang, and Le Wang, "Power control in UAV-supported ultra dense networks: Communications, caching, and energy transfer," *IEEE Communications Magazine*, vol. 56, no. 6, pp. 28–34, Jun. 2018.

[214] Chunxiao Jiang, Haijun Zhang, Yong Ren, Zhu Han, Kwang-Cheng Chen, and Lajos Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, Apr. 2017.

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

[216] Donald M Topkis, *Supermodularity and complementarity*, Princeton university press, 1998.

[217] Yoav Shoham and Kevin Leyton-Brown, *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*, Cambridge University Press, Cambridge, UK, 2008.

[218] Mathukumalli Vidyasagar, *Nonlinear systems analysis*, vol. 42, Siam, 2002.