Abstract

Device-to-Device (D2D) communication has become a promising candidate for future wireless communication systems to improve the system spectral efficiency, while reducing the latency and energy consumption of individual communication. With the assistance of cellular network, D2D communications can greatly reduce the transmit distance by utilizing the spatial dispersive nature of ever increasing user devices. Further, substantial spectrum reuse gain can be achieved due to the short transmit distance of D2D communication. It, however, significantly complicates the resource management and performance analysis of D2D communication underlaid cellular networks. Despite an increasing amount of academic attention and industrial interests, how to evaluate the system performance advantages of D2D communications with resource management remains largely unknown.

On account of the proximity requirement of D2D communication, the resource management of D2D communication generally consists of admission access control and resource allocation. Resource allocation of cellular assisted D2D communications is very challenging when frequency reuse is considered among multiple D2D pairs within a cell, as intense inter D2D interference is difficult to tackle and generally causes extremely large amount of signaling overheads for channel state information (CSI) acquisition. Hence, the first part of this thesis is devoted to the resource allocation of cellular assisted D2D communication and the performance analysis. A novel resource allocation scheme for cellular assisted D2D communication is developed with low signaling overhead, while maintaining high spectral efficiency. By utilizing the spatial dispersive nature of D2D pairs, a geography-based sub-cell division strategy is proposed to group the D2D pairs into multiple disjoint clus-

ters, and sub-cell resource allocation is performed independently for the D2D pairs within each sub-cell without the need of any prior knowledge of inter D2D interference. Under the proposed resource allocation scheme, tractable approximation for the inter D2D interference modeling is obtained and a computationally efficient expression for the average ergodic sum capacity of the cell is derived. The expression further allows us to obtain the optimal number of sub-cells that maximizes the average ergodic sum capacity of the cell. It is shown that with small CSI feedback, the system capacity/spectral efficiency can be improved significantly by adopting the proposed resource allocation scheme, especially in dense D2D deployment scenario.

The investigation of use cases for cellular assisted D2D communication is another important topic which has direct effect on the performance evaluation of D2D communication. Thanks to the spatial dispersive nature of devices, D2D communication can be utilized to harvest the vast amount of the idle computation power and storage space distributed at the devices, which yields sufficient capacities for performing computation-intensive and latency-critical tasks. Therefore, the second part of this thesis focuses on the D2D communication assisted Mobile Edge Computing (MEC) network. The admission access control of D2D communication is determined by both disciplines of mobile computing and wireless communications. Specifically, the energy minimization problem in D2D assisted MEC networks is addressed with the latency constraint of each individual task and the computing resource constraint of each computing entity. The energy minimization problem is formed as a two-stage optimization problem. At the first stage, an initial feasibility problem is formed to maximize the number of executed tasks, and the global energy minimization problem is tackled in the second stage while maintaining the maximum number of executed tasks. Both of the optimization problems in two

stages are NP-hard, therefore a low-complexity algorithm is developed for the initial feasibility problem with a supplementary algorithm further proposed for energy minimization. Simulation results demonstrate the near-optimal performance of the proposed algorithms and the fact that the number of executed tasks is greatly increased and the energy consumption per executed task is significantly reduced with the assistance of D2D communication in MEC networks, especially in dense user scenario.

Acknowledgements

I am indebted to many people who have given me guidance, support and encouragement during my Ph.D. candidature at the University of Kent.

First and foremost, I wish to express my deep gratitude to my supervisors, Dr. Huiling Zhu and Prof. Jiangzhou Wang, for the tireless guidance throughout my Ph.D study. I will always remember those meetings and discussions we had and the free atmosphere which led me to think independently. I feel very fortunate to work with both of them and benefit from both of their perspectives.

I would like to thank my Ph.D panel member, Prof. Nathan Gomes, for letting me be involved in many projects which broadens my vision beyond my research area. I would also like to thank Prof. Lin Dai from City University of Hong Kong for the instructive discussions which impressed me and inspired me a lot. Warmest thanks go to my colleagues and other technicians in our wireless communication group for their great support and friendly research environment.

Last but not least, deepest love and gratitude are sent to my parents and my fiancee, who always support me, care about me and encourage me. I love you all.

Contents

List of Figures

List of Algorithms

Abbreviations

1G	First Generation
4G	Fourth Generation
5G	Fifth Generation
AP	Access Point
AWGN	Additive White Gaussian Noise
BS	Base Station
BSCR	Best Subcarrier CSI
CSI	Channel State Information
D2D	Device-to-Device
i.i.d	Independent and Identically Distributed
IST	Iterative Switching Task
ITA	Iterative Task Allocation
LTE-A	Long Term Evolution Advanced
MCC	Mobile Cloud Computing
MEC	Mobile Edge Computing

MIMO	Multiple-Input-Multiple-Output
mmWave	Millimeter Wave
NOMA	Non-Orthogonal Multiple Access
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PA	Power Allocation
PL	Power Level
pdf	Probability Density Function
pmf	Probability Mass Function
PPP	Poisson Point Process
PHY	Physical layler
ProSe	Proximity Service
QoE	Quality of Experience
SAR	Subcarrier Achievable data-Rate
SINR	Signal-to-Interference-plus-Noise Ratio
SNR	Signal-to-Noise Ratio
UE	User Equipment

Chapter 1

Introduction

From the first generation (1G) mobile communication system launched in 1979 to the current fourth generation (4G) Long Term Evolution Advanced (LTE-A) system, the mobile communication network has been transformed from a pure voice communication system to a network that can support rich multimedia applications. Owing to the rapid development of smartphone applications and dramatic increase of smartphone users, the demand for high data rates has been growing exponentially, posing unprecedented challenges for designers of the next generation (i.e., the fifth generation (5G)) mobile communication systems.

To support a large number of cellular user equipments (UEs) with high-data-rate applications such as online gaming and streaming high-definition video, many new innovative solutions have been proposed (see [1, 2] for a comprehensive overview). In general, those solutions can be categorized into two methods:

1. Expanding the resources: The typical resource of wireless communication is critically limited spectrum. Extensive efforts have been made to utilize the millimeter wave (mmWave) spectrum, which was pervasive considered as unsuitable for wireless communications until recently [3].

Also, LTE-unlicensed is proposed to allow cellular UEs sharing the unlicensed spectrum with Wi-Fi [4].

2. Enhancing the resource efficiency: With regard to the characteristics of physical wireless communication channels, there exist various technologies to improve the spectral efficiency.

   As extra spatial dimensions of wireless channels can be used to improve the spectral efficiency according to the Multiple-Input-Multiple-Output (MIMO) theory [5], massive MIMO is developed by employing massive antennas at the base station (BS) [6].

   Another straightforward approach to improve the spectral efficiency is to reduce the communication cost by shortening the distance between transmitter and receiver. Since the UEs are geographically distributed in the cell, deploying either ultra-dense small cells [7] or massive distributed antennas [8] dispersively within the system can seize the UEs closely, hence significant gains can be achieved. Moreover, the device-to-device (D2D) communication is proposed by enabling two UEs in proximity to communicate with each other directly with the assistance of cellular networks [9, 10].

Among the new solutions, D2D communications is gaining increasing momentum, and emerging as a highly promising candidate for future wireless communication systems [11–13]. In D2D communication systems, exploiting direct communication between nearby UEs can improves the spectrum efficiency, while reducing the latency and energy consumption. In contrast to other technologies relying on the long-standing cellular architecture, i.e., centring around the base station, D2D communication has less requirement on the infrastructure deployment and achieves significant gains in direct proportion

to the ever increasing user density. The appealing features of D2D communication have attracted considerable attention from both industry and academia and been treated as one of the cutting-edge technologies leads the new trend of shifting from network-centric communication to user-centric communication. The following section provides an overview of the evolution and main challenges of D2D communications.

1.1 Overview of D2D communications

The concept of direct communication in local area has already been applied in several existing wireless technologies, which allow two UEs to communicate each other directly, such as Bluetooth, Zigebee and WiFi-direct [14, 15]. All of the aforementioned technologies are operated under pre-defined protocols in unlicensed spectrum bands, such self-configuring, dynamic decentralized networks are generally referred as wireless ad hoc networks. The capacity limits of ad hoc networks was well discussed in [16]. It is revealed that due to lack of centralized control, the throughput furnished to each UE diminishes to zero as the number of UEs is increased [16]. Moreover, the ad hoc networks are generally isolated from the cellular networks, thus manual settings are required to form the direct communication.

Since the cellular network is the most commonly and widely used central coordinator and service provider in wireless communication, it can provide real-time centralized control to the direct communication technology. Hence, D2D communication underlaying cellular networks was first introduced in [9]. Specifically, under the assistance of cellular base station, the cellular assisted D2D communication allows two nearby cellular UEs to form a D2D pair and communicate with each other directly without traversing the BS

or core network, thus improving the transmit quality dramatically due to short transmission distance. To be able to guarantee a planned (interference controlled) environment instead of an uncoordinated one, licensed spectrum can be utilized by the D2D communication underlaying cellular networks. In addition, the D2D discovery and establishment can be integrated with cellular signalling procedures to avoid the manual settings of direct communications [17, 18].

The first attempt to implement D2D communication in a cellular network was made by Qualcomms FlashLinQ [19], which is a synchronous PHY/MAC network architecture for distributed channel allocation of D2D communications underlaying cellular networks. Soon after, LTE-Direct was proposed by Qualcomm as a compelling solution for proximate discovery [20]. D2D communication was also investigated by 3GPP as Proximity Services (ProSe) and been included in 3GPP Release 12 with focus on one to many communications. An overview of standardization activities and the 3GPP ProSe is available in [11, 21].

1.2 Challenges of D2D communications

Since the proximity between two UEs is essential to have direct communication, not every UE within the system is suitable for D2D communication. Therefore, it is vital to have admission access control for the D2D communication underlaying cellular networks. For conventional cellular communications, the admission access control of cellular communications are mainly determined by the physical channels between the UEs and BS. By contrast, multiple layers (such as physical layer, MAC layer, network layer, application layer, etc.) should be jointly considered for D2D communications under multifarious types

of applications and user scenarios. For those UEs can not be supported by D2D communication, cellular communication is the alternative choice. The unique heterogeneity and dynamism are the key features of D2D communications, which lead to the D2D communication as an appropriate technology to harvest significant gains by utilizing the spatial resources as well as the ever increasing user density. However, such heterogeneity and dynamism also bring several major challenges to the D2D communication underlaying cellular networks.

1.2.1 D2D Discovery and Establishment

D2D discovery and D2D link establishment are the first two challenges for D2D communications. Some works suggest autonomous D2D communications, where D2D UEs are allowed to broadcast messages for D2D discovery and set up the D2D communication by themselves [20, 22]. For public safety applications, autonomous D2D communications are necessary as a disaster relief solution when the cellular network is damaged [23, 24]. However, to be able to achieve significant performance gains, D2D communications need considerable assistance from cellular network. Specifically, extra control signalling is needed for D2D discovery and D2D link establishment. As D2D communications always exist in dense user scenario, the D2D discovery and establishment protocol only considering physical channel may result a large amount, yet inefficient signalling overheads [18, 25]. Fortunately, it was shown in [26] that the signalling overhead of D2D discovery in the 3GPP networks can be significantly reduced by acquiring the location of the UEs. In 3GPP, the proximity service are currently being studied in [21], the relevant evolved packet core (EPC) entities for the UE localization are the mobility manage-

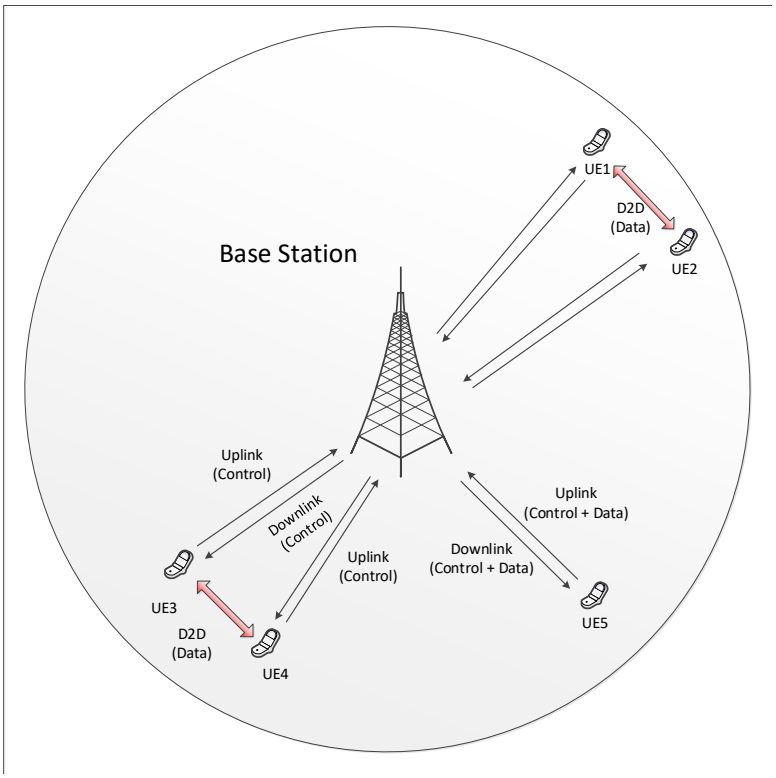


Figure 1.1: Graphic illustration of the general network architecture of D2D communication underlaying cellular networks. {UE1,UE2} , {UE3,UE4} are two D2D pairs, UE5 is a cellular UE.

ment entity (MME), the home subscriber server and the serving and packet data network gateway. Note that the location information of D2D UEs could be used to assist the resource allocation of D2D communication, which will be discussed in Chapter 3 . In addition, [17] shown that the accurate perception of social attributes of D2D UEs can further enhance the efficiency of the D2D discovery.

1.2.2 Resource Allocation of D2D communications

Since D2D communications are introduced to improve the spectral efficiency, radio resource allocation is no doubt to be the most important problem for D2D communication underlaying cellular networks. Due to the heterogeneity of D2D communication, multiple D2D UEs and cellular UEs may exist in the single cell simultaneously to share the licensed spectrum, as shown in Fig. 1.1. It is important to note that benefit from the short transmit distance of D2D communication, spectrum allocated for D2D communication can be reused to enhance the spectral efficiency [27, 28]. In [29], two resource allocation modes are proposed for the D2D UEs and cellular UEs to share the licensed spectrum, which are orthogonal mode and non-orthogonal mode. Specifically, for non-orthogonal sharing, the licensed spectrum can be shared among multiple D2D UEs and cellular UEs, accompanied with both inter D2D interference and cross D2D and cellular interference between D2D UEs and cellular UEs. For orthogonal sharing, D2D UEs are allocated with dedicated licensed spectrum. Hence only inter D2D interference exists among the D2D UEs within one cell.

To deal with the intra-cell interference caused by aggressive spectrum reuse, several efficient interference mitigation schemes are investigated under D2D communication underlaying cellular networks, such as beamforming [30, 31], interference alignment and cancellation [32, 33], cooperative communications [34–36]. However, these techniques have strong limitations due to specific system requirements. Therefore, a well-designed resource allocation with interference management scheme is critical for general D2D communication underlaying cellular networks. Generally, resource allocation of D2D communications is a complex and challenging problem and has attracted much attention in academia (see Section 2.1 for comprehensive review).

Moreover, although various resource allocation and interference manage-

ment schemes was proposed [27,28,37–40], they are generally too complicated which makes the performance analysis of D2D communications extremely difficult and challenging.

1.2.3 Cross-layer Admission Access Control of D2D communications

The D2D communication is one of the key components to enable some cutting edge user co-operation technologies, such as D2D assisted video transmission [41, 42], D2D assisted content caching [13, 43–45]. Note that for the aforementioned user co-operation technologies, the users within the system need to be paired or grouped based on the individual status of each user. That is, although D2D communications could enhance the users' quality of experience (QoE), the cross-layer admission access control is generally required to optimize the system performance.

As discussed in [44], thanks to the asynchronous content reuse property of content/video on-demand, a few popular files account for a large part of the traffic but are viewed by users at different times. Therefore, D2D communications can be utilized to deliver the requested content from nearby UEs' caches to reduce the backhaul data traffic. However, the performance gains of D2D assisted content caching is highly related to the popularity and distribution of each specific files, which are the information obtained by the network layer and application layer. That is, the link set up of D2D communications are determined not only by the physical channel conditions, but also by the matching in higher layers. Although the optimal content placement to maximize the density of successful receptions attracts plenty of research interests recently [46–50], the user scenario of D2D communications still generally lim-

ited. Therefore, how to expand the user scenarios for D2D communication underlaying cellular network is another key issue, which unfortunately still remains elusive.

1.2.4 Integration with Other Technologies

The wireless network is becoming more heterogeneous to meet the various of communication requirements. There are some promising technologies for the next generation of wireless communication networks, such as, high dense small cell networks, non-orthogonal multiple access (NOMA), mmWave.

In such multi-tier networks, jointly consideration of admission access control of D2D communication with other technologies are desired for different user scenarios, as discussed in [37, 51, 52]. Moreover, the intricately interference environment brings severe challenge to D2D communications [53].

1.3 Motivations and Thesis Outline

As D2D communications can significantly reduce the transmit distance of wireless communication, it have great potential to improve system spectral efficiency and become a basic technology for the future user centric communication networks. However, there exist several important challenges especially for the practical systems when the number of D2D pairs scaling up, as mentioned in previous section.

Among the fundamental challenges of D2D communications underlaying cellular networks, the resource management is the key technical drawback to realize the system performance advantage of D2D communications. Therefore,

this thesis is devoted to the designing of resource management for the cellular assisted D2D communications. Specifically, the resource management of D2D communication is addressed in two aspects. The first aspect is the resource allocation and performance analysis of D2D communication in orthogonal frequency division multiple access (OFDMA) system. The admission access control of a novel user scenario for D2D communication, D2D assisted mobile edge computing (MEC), is the second focused aspect.

The remainder of the thesis is organized as follows. The state-of-the-art of the resource management of D2D communication underlaying cellular networks and the contributions of this thesis are summarized in chapter 2. The resource allocation of cellular assisted D2D communication is discussed in chapter 3. Based on the interference analysis between multiple D2D pairs in a single cell, we show that D2D communication can effectively harvest the spatial resources from ever-increasing user devices, through proposed resource management scheme, to improve the system performance. Chapter 4 focus on the designing of admission access control scheme for D2D assisted MEC network. By considering both disciplines of mobile computing and wireless communications, it is shown that D2D communication can be utilized to harvest the vast amount of the idle computation power and storage space distributed at the edge devices in the system. Finally, concluding remarks and suggestions for future work are summarized in Chapter 5.

Chapter 2

Related Work and Thesis Contributions

Among the fundamental challenges of D2D communication underlaying cellular networks, this thesis focus on two important issues, which are the resource allocation and performance analysis of D2D communication in orthogonal frequency division multiple access system and the admission access control of D2D assisted MEC networks. The state-of-the-art for these two issues and the corresponding contributions are summarized in this chapter, respectively.

2.1 Resource Allocation and Performance Analysis

In cellular systems, OFDMA has been widely adopted for multi-user transmissions. Based on the properties inherited from orthogonal frequency division multiplexing (OFDM) which can transform the frequency selective fad-

ing channel into multiple flat fading subcarriers, high system capacity can be achieved by exploring multi-user diversity when adaptively allocating the subcarriers to multiple UEs [54–57]. Since all the cellular UEs within one cell need to communicate with one central BS, subcarriers cannot be reused by multiple cellular UEs within one cell due to the formidable strong intra-cell interference. By contrast, for cellular assisted D2D communications, both transmitters and receivers of the D2D pairs are spatially distributed within the cell and the transmission distance between the transmitter and the receiver of a formed D2D pair, also referred to as the distance of the D2D pair, is very small compared to the cell radius. Thus reusing the same subcarriers among multiple D2D pairs and/or between D2D pairs and cellular UEs could improve the overall system capacity, if mutual interference among multiple D2D pairs and cellular UEs is managed carefully, as mentioned in 1.2.2.

In general, resource allocation for cellular OFDMA system is decoupled into *subcarrier assignment* to each UE and *power allocation* to each subcarrier [54]. For D2D communication underlaying cellular networks, as each subcarrier can be reused by multiple D2D pairs, *link selection* on each subcarrier is needed to determine the D2D pairs and the cellular UE using the same subcarrier. Since inter D2D interference is correlated with link selection, subcarrier assignment and power allocation, joint consideration of these three dimensions is required to achieve optimal system performance, e.g. system capacity.

In early research [58], simple network settings have been considered and the transmission power control was investigated when one D2D pair shares the same subcarrier with one cellular UE. For complex systems with multiple D2D pairs, even with fixed subcarrier assignment, the resource allocation with joint consideration of power allocation and link selection is strongly NP-hard [59].

Thus, various algorithms on the basis of advanced theories have been adopted to develop resource allocation for cellular D2D underlaid OFDMA system.

Due to the non-convexity of the received signal to interference plus noise ratio (SINR) at each UE and binary constraints of subcarrier allocation, the resource allocation of D2D communication is generally a non-convex mixed-integer problem, which is difficult to solve optimally. Most commonly used technique to tackle the non-convexity is the standard Lagrange dual relaxation [59,60]. It was shown in [60] that zero duality gap can be achieved when number of subcarriers grows to infinitely large. For resource limited system, either sequential convex programming or successive convex approximation were adopted in [61,62] to approximate the non-convex resource allocation problem.

[63] introduced a sub-optimal resource allocation scheme by modelling the interference among multiple D2D pairs and cellular UEs as a interference graph according to graph theory. As the interference was constructed independently in [63], hypergraph coloring was adopted in [64] to model the cumulative interference from multiple sources.

Game theory has also been employed from the economy perspective. In [65, 66], a combinatoric auction game based fictional pricing mechanism was designed for all D2D UEs to perform the distributed resource allocation. By setting the BS as leader and D2D pairs as followers, a two-stage Stackelberg game was formed in [67] as resource allocation for cellular D2D underlaid OFDMA system.

Although the cellular D2D underlaid OFDMA system was addressed in aforementioned research, the investigation of resource allocation was limited to two scenarios. In the first scenario, D2D pairs and cellular users share one resource unit, and therefore subcarrier assignment was not considered [58]. In the second scenario, transmission power on a subcarrier was fixed for each

D2D pair [63–65,67], implying that the total transmission power for the transmitter of a D2D pair increases with the number of subcarriers allocated to the device. However, in practice, the total transmission power of each mobile device is limited. Additionally, the perfect knowledge of channel state information (CSI) between any transmitter and receiver was generally assumed at the BS to make efficient resource allocation, which introduces a large amount of signalling overhead [68]. Therefore, it is necessary to develop a low complexity, low signalling overhead yet efficient scheme for resource allocation and interference management of D2D communication assisted cellular networks.

From existing researches, the system capacity may be increased by reusing the same spectrum among multiple D2D pairs and one cellular user within one cell. However, severe cross D2D and cellular interference would significantly degrade the performance of cellular users and lead to extra interference control overhead to cellular systems [69]. To reduce the impact of the D2D communication to existing cellular resource allocation, we focus on the cellular assisted D2D communication under orthogonal sharing mode as mentioned in [29], which is easy to be integrated with cellular systems. That is, D2D pairs within the cell are allocated with dedicated spectrum. Hence only inter D2D interference exists among D2D pairs within one cell. Furthermore, [70] proved that in a small area, spectrum should not be reused due to the strong interference. Since the inter D2D interference is largely determined by the distance-dependent large scale fading, the spatial dispersion property of D2D pairs, which means the distance between any two D2D pairs [1] could be relatively large compared to the distance of a D2D pair, can be applied to the link selection to avoid strong inter D2D interference. However, due to lack

[1]For the sake of simplicity, the distance between the transmitter of one D2D pair to the receiver of another D2D pair is referred to as the distance between the specific two D2D pairs in the rest of the thesis.

of theoretical analysis of the inter D2D interference, such spatial dispersion property of cellular assisted D2D communication has not been used for the resource allocation design.

2.2 D2D assisted Mobile Edge Computing

With the increasing popularity of UEs, such as smartphones and hand-held devices, more and more resource-hungry applications like high definition video gaming and virtual reality applications are developing rapidly and playing an important role in our end-UEs. However, vast computation-intensive and latency-critical applications bring severe challenges to the physical limitations of devices, such as CPU capability and battery life. In last decades, mobile cloud computing (MCC) has attracted much attention to deliver elastic computing power from the cloud centre to support resource-constrained end-UEs in the network.

In consideration of the extra energy and time consumed for the communication between end-UEs and the cloud centre, long transmission distance becomes a key drawback for MCC [71]. In recent years, a new paradigm called mobile edge computing is widely agreed to be a key technology for releasing the idle computation resource at the edge access point (AP) of the network instead of centralized cloud centre [72]. Although the energy and time consumption can be greatly reduced by eliminating the transmission distance from AP to cloud centre, the energy and time consumption for the wireless communication still remain, which need to be tackled carefully.

With respect to the wireless networks, most of the previous researches in MEC focused on the joint task offloading and resource allocation from the UEs to the infrastructure, i.e. edge AP, and the energy efficiency had been

studied comprehensively under different network structures [73–75]. However, a crucial issue of energy minimization problem is the feasibility problem for such MEC networks, which is to say, it may not be possible to meet the latency constraints of all the tasks, especially for the edge UEs due to the cell-edge effect. For example, it is expected that large amount of devices will be deployed in the future wireless systems, but the transmit and computation power of each device is strictly limited due to limited battery life. Hence, significant amount of the tasks would be offloaded to the cloud networks. However, to guarantee the quality of service, massive number of powerful APs need to be deployed in the network to guarantee the coverage and peak service demands, which is costly and inefficient.

From the aforementioned researches in Section 1.2.3, it can be concluded that cellular assisted D2D communications can greatly reduce the communication distance for the desired UEs with proper cross-layer admission access control, hence minimize the infeasible scenarios as well as the total energy consumption [76]. Therefore, D2D communication can be used to assisted the MEC networks by allowing the device to offload its task to the nearby devices. Unlike the D2D assisted content caching, which is difficult to achieve fully user spatial diversity, since every UE could be the potential beneficiary of MEC, higher diversity gains can be achieved in D2D assisted MEC networks. That is, the resources of any computing entities including both UEs and infrastructures can be utilized and due to the spatial dispersion nature of UEs, the offloading energy and time cost can be greatly reduced for all UEs including the cell edge UEs, and the number of infeasible tasks can be significantly reduced.

2.3 Thesis Contribution

As mentioned in 2.1, although a great deal of studies have been conducted on the resource allocation of D2D communication underlaying cellular networks, most of them focus on the complex algorithm design by assuming perfect CSI of the channels between any UEs within the cell. Hence, how to evaluate the performance gains of cellular assisted D2D communication unfortunately remains largely unknown.

In Chapter 3, a low complexity resource allocation scheme is proposed for D2D communications underlaid an OFDMA-based cellular system by taking advantage of the spatial dispersion property of D2D pairs. With the objective of achieving high spectrum efficiency with low complexity, the resource allocation is decoupled into two stages. First, a simple geography based D2D sub-cell division strategy is proposed to form the D2D pairs within one cell into multiple groups. Specifically, the coverage of one single cell is divided into multiple small areas under symmetric hexagonal topology, which we refer to as sub-cells, and the D2D pairs located within one sub-cell are set into one group. Then, to avoid the strong intra-cell interference, dedicated subcarriers for cellular D2D communications are orthogonally assigned to the D2D pairs within one sub-cell, whereas the subcarriers are reused between different sub-cells to enhance the overall spectral efficiency. In other words, the inter D2D interference only exists between different sub-cells. Moreover, to alleviate the heavy signalling overhead burden caused by CSI acquirement, the subcarrier assignment and power allocation for the D2D pairs in one sub-cell are performed without the need of the CSIs of the inter D2D interference channels.

Based on the proposed resource allocation scheme, the *average ergodic sum*

capacity[2] for cellular D2D underlaid OFDMA system is derived by thorough analysis of the inter D2D interference. The suitable number of sub-cells to maximize the average ergodic sum capacity is further obtained as an important system parameter to D2D sub-cell division strategy and hence substantial capacity/spectral efficiency gains is guaranteed by the proposed resource allocation scheme.

In Chapter 4, a novel user scenario for D2D communication, D2D assisted MEC networks are investigated. As discussed in 2.2, the MEC can be greatly enhanced with the assistance from cellular assisted D2D communication. To evaluate the amount of performance gain can D2D assisted MEC network achieve, we focus on the total energy consumption, which is constituted by the computing energy and wireless communication energy. Specifically, the total energy minimization problem and the feasibility problem of D2D assised MEC networks are formed in two stages by considering the admission access control of D2D communications, respectively.

In the first stage, a low complexity algorithm is developed to maximize the number of executed tasks subject to the latency constraint of each individual task and the computing resource constraint of each computing entity, i.e. edge AP and UEs. A supplementary algorithm is then proposed in the second stage to minimize the total energy consumption while maintaining the maximum number of executed tasks. Simulation results show that the number of executed tasks and the corresponding energy consumption per executed task with our proposed algorithms are very close to those with the optimal exhaustive search. Furthermore, it is revealed that the number of executed tasks in MEC networks is significantly increased with the assistance of D2D

[2]Specifically, "average ergodic sum capacity" means that the ergodic sum capacity is averaged over the large-scale fading coefficients, e.g. the spatial location, of each D2D pairs.

communication, while the energy consumption per executed task is greatly reduced, especially in dense UE scenarios.

Chapter 3

Resource Allocation and Performance Analysis of D2D Communication

---

In this chapter, based on the spatial dispersive property of D2D pairs, a low complexity resource allocation scheme will be proposed for cellular assisted D2D communication. To eliminate the strong intra-cell interference, a simple geographic based sub-cell division strategy will be introduced to form the D2D pairs within one cell into multiple groups, such that subcarriers can be reused among the D2D pairs in different groups. The joint subcarrier assignment and power allocation scheme will be performed for the D2D pairs in the same sub-cell. The average ergodic sum capacity of proposed resource allocation scheme will be investigated by thorough analysis of the inter D2D interference and the optimal number of sub-cells will be further shown as an important parameter to the sub-cell division strategy.

This chapter is organized as follows. In Section 3.1, the system model for the cellular D2D underlaid OFDMA system is introduced along with the pro-

posed sub-cell division strategy. The sub-cell resource allocation algorithms and corresponding sum capacities under the single sub-cell structure are investigated in Section 3.2. The sum capacity under multiple sub-cell structure and the effect of the number of sub-cells are further investigated in Section 3.3. Finally, conclusions are summarized in Section 3.4.

3.1 System Model and Preliminary Analysis

Consider an OFDMA-based single cell system underlaid by D2D communications. As shown in Fig. 3.1(a), a set of K active D2D pairs, denoted as $\mathcal{K} = \{1, \cdots, K\}$ are located in the coverage area of one BS. The kth D2D pair, $D2D_k$, is formed by a transmitter device and a receiver device, called transmitter k and receiver k, respectively. In this model, the circular area with radius R is denoted as the coverage region of the BS centred at the origin. It is assumed that the D2D transmitters are distributed in the whole $\mathbb{R}^2$ plane following a homogeneous Poisson Point Process (PPP) Φ_A with density λ_A, and the receiver k associated with transmitter k is located uniformly on the circle centred by transmitter k with distance d_k. Then the average number of D2D pairs in the cell, denoted as $\bar{K}$, is given by $\bar{K} = \pi\lambda_A R^2$. The maximum transmit power of each D2D device is assumed to be limited to P_{max}.

To avoid the cross interference between D2D pairs and cellular users, a set of N dedicated subcarriers, denoted as $\mathcal{N} = \{1, \cdots, N\}$, are assigned for D2D communications. Each subcarrier can be reused among multiple D2D pairs within the cell in order to increase the system spectral efficiency. As a consequence, inter D2D interference exists among the D2D pairs sharing

the same subcarrier. Specifically, for the D2D pair $D2D_k$, the potential inter D2D interference on subcarrier n comes from the transmitter of any $D2D_{k'}$, $k' \in \mathcal{K} \setminus \{k\}$. Note that the interference only exists when the same subcarrier is reused by $D2D_k$ and $D2D_{k'}$. Hence, link selection is performed to determine the set of D2D pairs which may share the same subcarrier, or in other words, the set of D2D pairs which cannot share the same subcarrier.

Since the set of D2D pairs using the same subcarrier is largely affected by the link selection, it can be foreseen that the subcarrier assignment to D2D pairs is highly correlated with the link selection of each subcarrier. In addition, as each device has the maximum transmit power P_{max}, the power allocation of the assigned subcarriers to each D2D pair is inextricably interwoven with the subcarrier assignment and the link selection. Thus, joint consideration of link selection, subcarrier assignment and power allocation is required for resource allocation of the cellular D2D underlaid OFDMA system to achieve optimal spectral efficiency. However, the joint optimization problem is proven to be NP-hard [59], which requires unaffordable computing capability at the BS to trace the optimal resource allocation result. Moreover, it induces massive amount of signalling overhead for D2D channel measurement and feedback, especially when the number of D2D pairs K is large.

It has been shown in [70] that when a subcarrier is shared by two D2D pairs, which are spatially close to each other, strong inter D2D interference would degrade the performance of both D2D pairs significantly. In other words, inter D2D interference should be efficiently reduced by preventing the D2D pairs within a small area from reusing the same subcarrier. Therefore, in order to avoid the strong inter D2D interference, we propose to divide the cell into multiple disjoint small areas, which are referred to as sub-cells in this chapter, and orthogonally allocate the N subcarriers to the D2D pairs within

the same sub-cell. That is, N subcarriers can be reused among different sub-cells only, and the sub-cell division strategy is proposed as the substitution of the link selection to determine the set of D2D pairs which cannot share the same subcarrier.

3.1.1 Sub-cell Division Strategy

As the D2D transmitters are assumed to be uniformly distributed in the cell, without loss of generality, for any specific D2D pair $D2D_k$, the distance between any other D2D transmitter to the receiver of $D2D_k$ follows independent and identical distribution. Also, the channel gain of inter D2D interference between any two devices is mainly determined by the distance-dependent large scale fading. A simple geography-based sub-cell division strategy is proposed. Specifically, the single cell is considered as a J-tier symmetrical hexagonal sub-cell structure, as shown in Figs. 3.1 (b)-(d). By taking central sub-cell as the reference sub-cell, we denote its neighbouring sub-cells as the 1st tier sub-cells, which are crossed by the 1st tier circle, as shown in Fig. 3.1(b). Similarly, the jth tier is formed by the outer neighbouring sub-cells of the $(j-1)$th tier sub-cells, as illustrated in Figs. 3.1 (c)-(d). Based on the geographic characteristic of a J-tier hexagonal structure, the single cell is divided into a total of $L = 3 \cdot J \cdot (1 + J) + 1$ sub-cells. Each sub-cell can be then approximated as a circle area with radius R_L [77], and indexed from one to L. The sub-cell radius can be obtained by

$$R_L = \frac{R}{1 + 2 \cdot J} = \sqrt{\frac{3}{4L - 1}} \cdot R. \tag{3.1}$$

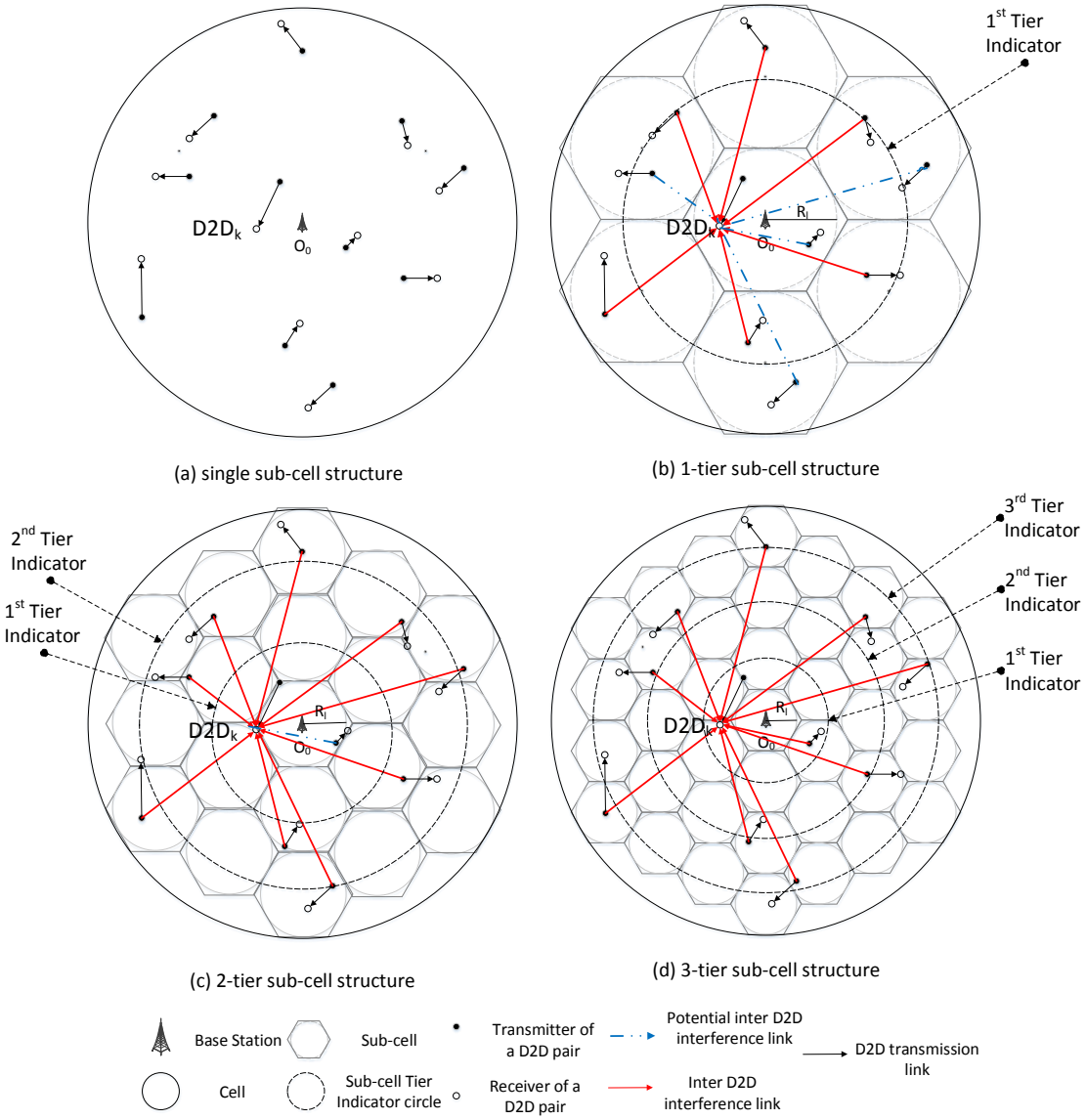


Figure 3.1: System model of cellular assisted D2D communications with hexagonal sub-cell division strategy.

With the sub-cell division, K D2D pairs are divided into L groups, where group l includes a set of D2D pairs, $\mathcal{K}_l$, which are located within the lth sub-cell. Each group contains $K_l = |\mathcal{K}_l|$ D2D pairs and $\sum_1^L K_l = K$.

Note that when $L = 1$, N subcarriers are orthogonally assigned to K D2D pairs within the cell and there is no inter D2D interference, while the number of interference links increases with the number of sub-cells for any specific D2D pair $D2D_k$, as shown in Fig. 3.1. Hence, it can be seen intuitively that the cumulative inter D2D interference received at $D2D_k$ depends on the number of sub-cells L.

3.1.2 Signal Model

In this chapter, the channel of each subcarrier is assumed to be unchanged during a resource allocation period. If subcarrier n is allocated to $D2D_{k_l}$ in sub-cell l, the received signal at the receiver of $D2D_{k_l}$ can be written as

$$y_{k_l,n} = \sqrt{P_{k_l,n}} \cdot g_{k_l,n} \cdot s_{k_l,n} + \underbrace{\sum_{\substack{l'\neq l \\ l'=1}}^{L} \sum_{k_{l'}\in\mathcal{K}_{l'}} \sqrt{P_{k_{l'},n}} \cdot g_{k_{l'},k_l,n} \cdot s_{k_{l'},n}}_{Interference} + z_{k_l,n}, \quad (3.2)$$

where $s_{k_l,n}$ is a complex scalar representing the information signal from $D2D_{k_l}$ on subcarrier n with $\mathbb{E}(|s_{k_l,n}|^2) = 1$, and $P_{k_l,n}$ denotes the transmit power of $D2D_{k_l}$ on subcarrier n. $z_{k_l,n} \sim CN(0, \sigma^2)$ is the additive white Gaussian noise (AWGN) at the receiver of $D2D_{k_l}$. The channel gain from the transmitter

$D2D_{k_{l'}}$ in sub-cell l' to receiver $D2D_{k_l}$ is denoted as $g_{k_{l'},k_l,n}$[1], given by

$$g_{k_{l'},k_l,n} = \iota_{k_{l'},k_l} \cdot h_{k_{l'},k_l,n}, \tag{3.3}$$

where $h_{k_{l'},k_l,n}$ is the small-scale fading and assumed to be an independent and identically distributed (i.i.d.) Rayleigh random variable with zero mean and unit variance. $\iota_{k_{l'},k_l}$ represents the path loss coefficient, which is given by

$$\iota_{k_{l'},k_l} = {d_{k_{l'},k_l}}^{-\frac{\alpha}{2}}, \tag{3.4}$$

where $d_{k_{l'},k_l}$ is the distance from $D2D_{k_{l'}}$ to $D2D_{k_l}$, and α is the path loss exponent.

The achievable data rate of $D2D_{k_l}$ on subcarrier n can then be expressed as

$$C_{k_l,n} = B \cdot \log_2 \left(1 + \frac{P_{k_l,n} \cdot |g_{k_l,n}|^2}{I^c_{k_l,n} + \sigma^2} \right), \tag{3.5}$$

where B denotes the bandwidth of one subcarrier. $I^c_{k_l,n}$ is the cumulative inter D2D interference power received at $D2D_{k_l}$ on subcarrier n, which can be obtained according to (3.2) as

$$I^c_{k_l,n} = \sum_{\substack{l'=1 \\ l' \neq l}}^{L} \sum_{k_{l'} \in \mathcal{K}_{l'}} P_{k_{l'},n} \cdot \left| g_{k_{l'},k_l,n} \right|^2. \tag{3.6}$$

[1]For simple notation, the channel gain on subcarrier n between the transmitter and receiver of $D2D_{k_l}$, $g_{k_l,k_l,n}$ is denoted as $g_{k_l,n}$. In the rest of the chapter, the subscript of any random variable between the transmitter and receiver of one D2D pair is simplified in the same manner.

3.1.3 Problem Formulation

With the geography-based sub-cell division strategy proposed in Section 3.1.1, the sum capacity of K D2D pairs within the cell can be seen as the sum capacity of L sub-cells, each contains $\mathcal{K}_l$ D2D pairs. For the sake of clarity, we introduce the subcarrier assignment indicator denoted as $\omega_{k_l,n}$, where $\omega_{k_l,n} = 1$ indicates subcarrier n is allocated to $D2D_{k_l}$, otherwise, $\omega_{k_l,n} = 0$, for all $k_l \in \mathcal{K}_l$ and $n \in \mathcal{N}$. The sum capacity of the cell under L sub-cell structure can then be written as

$$C_{TOT} = \sum_{l=1}^{L} \sum_{k_l \in \mathcal{K}_l} \sum_{n \in \mathcal{N}} \omega_{k_l,n} \cdot B \cdot \log_2 \left(1 + \frac{P_{k_l,n} \cdot |g_{k_l,n}|^2}{\sum_{\substack{l'=1 \\ l' \neq l}}^{L} \sum_{k_{l'} \in \mathcal{K}_{l'}} \omega_{k_{l'},n} \cdot P_{k_{l'},n} \cdot \left| g_{k_{l'},k_l,n} \right|^2 + \sigma^2} \right). \quad (3.7)$$

The joint sub-cell division, subcarrier assignment and power allocation resource allocation problem can be formulated as

$$\max_{L,\ \{\omega_{k_l,n}\},\ \{P_{k_l,n}\}} C_{TOT} \quad (3.8)$$

$$s.t. \qquad \omega_{k_l,n} \in \{0,1\}, \quad \forall\ k_l \in \mathcal{K}_l,\ n \in \mathcal{N}, l = 1, \cdots, L; \quad (3.8a)$$

$$\sum_{k_l \in \mathcal{K}_l} \omega_{k_l,n} \leqslant 1, \quad \forall\ n \in \mathcal{N}, l = 1, \cdots, L; \quad (3.8b)$$

$$\sum_{n \in \mathcal{N}} P_{k_l,n} \leqslant P_{max}, \quad \forall\ k_l \in \mathcal{K}_l, l = 1, \cdots, L; \quad (3.8c)$$

where constraint (3.8a) indicates the interger property of the subcarrier assignment indicator and constraint (3.8b) indicates that the subcarrier n is

only allocated to one D2D pair within each sub-cell, which implies that a subcarrier cannot be reused within one sub-cell. Constraint (3.8c) shows the total transmit power of a D2D pair is limited by the maximum transmit power P_{max}.

It can be clearly seen from (3.8) that based on the sub-cell division strategy, there are three parameters to be optimized for the resource allocation to achieve the maximum sum capacity, i.e., the optimal number of sub-cells L_{opt} and the resource allocation results $\{\omega_{k_l,n}^{opt}\}_{k_l\in\mathcal{K}_l,n\in\mathcal{N}}$, $\{P_{k_l,n}^{opt}\}_{k_l\in\mathcal{K}_l,n\in\mathcal{N}}$ for each sub-cell $l = 1, \cdots, L_{opt}$. The resource allocation of L sub-cells are correlated with each other due to the inter D2D interference. Therefore, to achieve the optimal resource allocation, the CSI between any D2D pairs in the cell, $\{g_{k',k,n}\}_{k'\in\mathcal{K},k\in\mathcal{K},n\in\mathcal{N}}$ are required at BS, leading to $\mathcal{O}(K^2N)$ signalling overhead for the CSI acquirement.

In order to alleviate the high signalling overhead, in this chapter, sub-cell based resource allocation is proposed to perform subcarrier assignment and power allocation for the D2D pairs within one sub-cell regardless of the inter D2D interference. As a result, the signalling overhead for the CSI acquirement is significantly reduced from $\mathcal{O}(K^2N)$ to $\mathcal{O}(KN)$, since only the CSI of each D2D pair, $\{g_{k,n}\}_{k\in\mathcal{K},n\in\mathcal{N}}$, is needed. On the other hand, it is noted that the sum capacity is decisively affected by the inter D2D interference. Under the proposed sub-cell division strategy, for any specific D2D pair, the cumulative interference depends on the number of sub-cells. Therefore, we are specifically interested in the optimal number of sub-cells, $\hat{L}_{opt}$, that maximizes the average ergodic sum capacity of the system over the locations of all D2D pairs.

3.2 Sum Capacity under Single Sub-cell Structure

As discussed in Section 3.1.1, the subcarriers are orthogonally assigned to the D2D pairs within the same sub-cell regardless of interference. In this section, a special case, where the single sub-cell structure (i.e. $L = 1$, as shown in Fig. 1(a)) is considered, thus no inter D2D interference exists. The sub-cell resource allocation, including subcarrier assignment and power allocation, is investigated first and the ergodic sum capacity is then studied in section 3.2.2.

3.2.1 Single Sub-cell Resource Allocation

With the objective of maximizing the sum capacity of the K D2D pairs, the resource allocation problem of single sub-cell under individual power constraint of each D2D pair can be formulated as:

$$\max_{\{\omega_{k,n}^{opt}\},\ \{P_{k,n}^{opt}\}} C = \sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{N}} B\cdot\omega_{k,n}\cdot\log_2\left(1+\frac{P_{k,n}\cdot|g_{k,n}|^2}{\sigma^2}\right) \tag{3.9}$$

$$\text{s.t.} \qquad \omega_{k,n}\in\{0,1\}, \forall\ k\in\mathcal{K},\ n\in\mathcal{N}; \tag{3.9a}$$

$$\sum_{k\in\mathcal{K}}\omega_{k,n}\leqslant 1,\ \forall\ n\in\mathcal{N}; \tag{3.9b}$$

$$\sum_{n\in\mathcal{N}}P_{k,n}\leqslant P_{max},\ \forall\ k\in\mathcal{K}; \tag{3.9c}$$

Due to the integer constraint $\omega_{k,n}$, (3.9) is neither convex nor concave and thus hard to tackle. Note that in most cases, equality in (3.9b) and (3.9c) needs to be satisfied in order to obtain the optimal solution. That is, (3.9b)

and (3.9c) can be replaced by

$$\sum_{k\in\mathcal{K}} \omega_{k,n} = 1,\ \forall\ n \in \mathcal{N}, \tag{3.10}$$

and

$$\sum_{n\in\mathcal{N}} P_{k,n} = P_{max},\ \forall\ k \in \mathcal{K}, \tag{3.11}$$

respectively. Based on constraints (3.10) and (3.11), the optimal solution to the optimization problem (3.9) can be found by using the global searching method, which will be discussed in next sub-section.

3.2.1.1 Global searching method

To achieve the optimal system throughput, all the possible subcarrier combinations which satisfy constraint (3.10) should be considered. Specifically, since each subcarrier can only be allocated to one D2D pair at one time, there exists $S = K^N$ number of different subcarrier combinations, where each of the subcarrier combinations is denoted as $\Omega_s = \{\omega_{k,n}\}_{k\in\mathcal{K},n\in\mathcal{N}}$. For the subcarrier combination s, the power for the subcarriers $\mathcal{N}_k^s$ assigned to $D2D_k$ is allocated according to the water-filling algorithm [78]. The transmit power in subcarrier $n_k \in \mathcal{N}_k^s$ can be shown as

$$\hat{P}_{k,n_k}^s = \sigma^2 \left(\frac{1}{\varsigma} - \frac{1}{|g_{k,n_k}|^2} \right)^+, \tag{3.12}$$

where $(x)^+ \triangleq \max(x, 0)$, and the threshold ς is selected to meet the power constraint (3.11). After power allocation, the achievable system throughput of each subcarrier combinations, denoted as $\hat{C}^s_{TOT}$, will be calculated according to (3.9). Then, the optimal subcarrier and power allocation result of the exhaustive searching method can be obtained as $\left(\Omega_{s^*}, \hat{P}^{s^*}_{k,n_k}\right) = \arg\max\{\hat{C}^s_{TOT}\}_{K^N}$.

Even though the exhaustive searching method can achieve optimal result, but its computational complexity is K^N, as each subcarrier $n \in \mathcal{N}$ can be allocated to any D2D pairs $k \in \mathcal{K}$. Note that even for a system with a small number of D2D pairs, the computational complexity would be extremely high, since the number of subcarriers N could be very large. For instance, transmission bandwidth of 10 MHz are divided into 1024 subcarriers in LTE systems. Thus, simplified algorithms are discussed in the following subsections.

3.2.1.2 Baseline resource allocation algorithm

Greedy resource allocation algorithm is commonly used to maximize the total system throughput in OFDMA downlink resource allocation [78]. In the greedy downlink resource allocation, each subcarrier is allocated to the user with best CSI on the subcarrier. As the BS is the transmitter for the downlink signals, thus only one total power constraint needs to be satisfied [78], water-filling algorithm called global water-filling will be performed to assign the power to all the subcarriers at the BS.

When adopting the greedy resource allocation to the multiple pairs of D2D communication, as each device will have a unique maximum power constraint as shown in (3.11), global water-filling power allocation algorithm is no longer suitable. Instead, local water-filling power allocation is performed for each D2D pair to maximize the achievable data-rate for individual D2D link. This

method can be summarized as *Best subcarrier CSI (BSCR)-based resource allocation scheme*. Firstly, each subcarrier is allocated to the D2D pair which has maximum achievable data rate (denote as $\Delta r_{k,n}$). Then, local water-filling power allocation is carried out for the power allocation of each D2D pair among the subcarriers allocated to the pair.

3.2.1.3 Sub-optimal resource allocation algorithm

It is note that, for one D2D pair $D2D_k$, the number of subcarriers allocated, N_k, is a variable. Please note, if the total transmit power is given for $D2D_k$, the power level after performing water-filling for $D2D_k$ will decrease with increasing N_k. Fig. A.1 illustrates how the power allocation (PA) and power level (PL) are changed when $N_k = 9$ and $N_k = 10$ subcarriers are allocated to $D2D_k$. When the total power level for all the subcarriers allocated to $D2D_k$ decreases, the total data-rate for $D2D_k$ may decreases. For the BSCR-based resource allocation scheme, the number of subcarriers allocated to one D2D pair is ignored, which may cause the problem of allocating quite large number of subcarriers to one D2D pair, thus degrades the system throughput. Therefore, in this section, with the consideration of simplifying the global searching method and addressing the power level changes caused by the change of number of subcarriers allocated to one D2D pair, a two step sub-optimal resource allocation algorithm is proposed.

In the first step, it is assumed that equal power allocation is applied to all subcarriers allocated to one D2D link. The transmit power for $D2D_k$ on each subcarrier is given by

$$P_{k,n_k} = \frac{P_{max}}{N_k}, n_k \in \mathcal{N}_k \tag{3.13}$$

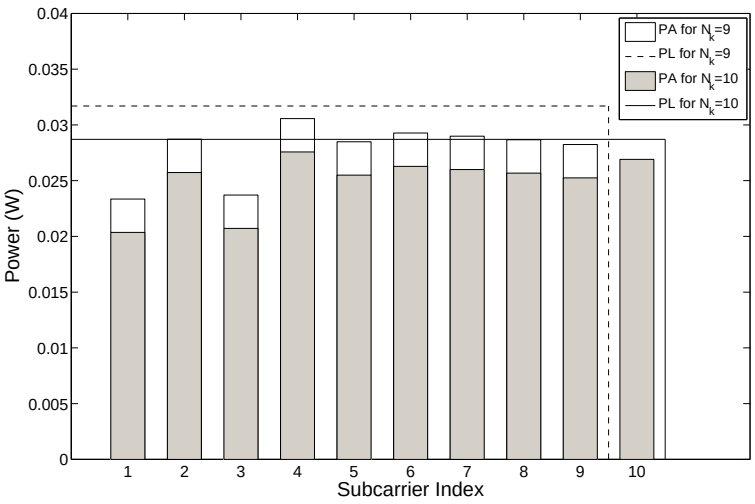


Figure 3.2: Power level and power allocation for $D2D_k$ when $N_k = 9$ and 10.

where $\mathcal{N}_k$ represents the set of subcarriers allocated to $D2D_k$ and N_k is the number of subcarriers allocated to $D2D_k$. The instantaneous achievable system throughput can be shown as

$$R = \sum_{k=1}^{K} \sum_{n_k \in \mathcal{N}_k} C_{k,n_k}, n_k \in \mathcal{N}_k \tag{3.14}$$

For a given set $\mathcal{N}_k$ for $D2D_k$, if one more subcarrier n_k^* is allocated to $D2D_k$, i.e. $\mathcal{N}_k$ replaced by $\mathcal{N}_k \bigcup \{n_k^*\}$, the change of achievable data-rate for $D2D_k$,$\Delta r_{k,n_k^*}$, can be given by

$$\Delta r_{k,n_k^*}(\mathcal{N}_k) = \sum_{n_k \in \mathcal{N}_k \bigcup \{n_k^*\}} C_{k,n_k}\left(\frac{P_{max}}{N_k+1}, d_k\right) - \sum_{n_k \in \mathcal{N}_k} C_{k,n_k}\left(\frac{P_{max}}{N_k}, d_k\right). \tag{3.15}$$

In our proposed algorithm, assuming $\mathcal{N}_k$ has been allocated to $D2D_k$, whether an unallocated subcarrier n can be allocated to $D2D_k$ depends on

whether its data rate increase $\Delta r_{k,n}(\mathcal{N}_k)$ is the maximum among all the D2D pairs. That is, subcarrier n will be allocated to $k^* = \arg\max_{k,n} \Delta r_{k,n}(\mathcal{N}_k)$. Therefore, our proposed scheme is called *Subcarrier achievable data-rate (SAR)-based algorithm.*

Based on (3.15), the subcarrier allocation is performed as the first step for suboptimal resource allocation algorithm, which is present as Algorithm 1.

Algorithm 1 Subcarrier allocation of SAR resource allocation scheme

1: **Initialization:** $\mathcal{N}_k = \Phi$,$N_k = 0$, $R_k = 0$,$\Delta r_{k,n}(N_k) = R_k(P_{max}, d_k)$ for all $k \in \{1, \cdots, K\}$ and $n \in \{1, \cdots, N\}$.
2: **for** $n \in (1, \cdots, N)$ **do**
3: **for** $k \in (1, \cdots, K)$ **do**
4: Calculate $\Delta r_{k,n}(\mathcal{N}_k)$ by using (3.15); $\{k^*, n^*\} = \arg\max_{k,n} \Delta r_{k,n}(\mathcal{N}_k)$; $\mathcal{N}_{k^*} = \mathcal{N}_k \bigcup \{n\}$; $N_{k^*} = N_{k^*} + 1$; $\omega_{k^*,n^*} = 1$
5: **end for**
6: **end for**

After subcarrier allocation, in the second step of SAR-based resource allocation, water-filling is carried out for each $D2D_k$ on its allocated subcarriers, $\mathcal{N}_k$, given in (3.12).

3.2.2 Ergodic Sum Capacity with Single Sub-cell Structure

In this section, we investigate the ergodic sum capacity of both BSCR and SAR based sub-cell resource allocation and show how much performance gain can be achieved via central resource allocation scheme.

Generally, since both BSCR and SAR based sub-cell resource allocation are performed based on the CSI of K D2D pairs, which are i.i.d., the ergodic capacity of any D2D pair is identical with each other. Therefore, the ergodic

sum capacity of K D2D pairs can be obtained from (3.3), (3.4) and (3.9), as

$$\begin{aligned}\bar{C}_{TOT} &= \mathbb{E}_{\{\mathcal{N}_k\}_{k\in\mathcal{K}},\{h_{k,n_k}\}_{k\in\mathcal{K},n_k\in\mathcal{N}_k}}\left\{\sum_{k\in\mathcal{K}} C_k\right\} \\ &= K\cdot\mathbb{E}_{\mathcal{N}_k,\{h_{k,n_k}\}_{n_k\in\mathcal{N}_k}}\left\{\sum_{n_k\in\mathcal{N}_k} B\cdot\log_2\left(1+\frac{P_{k,n_k}\cdot {d_k}^{-\alpha}\cdot|h_{k,n_k}|^2}{\sigma^2}\right)\right\},\end{aligned} \tag{3.16}$$

where the transmit power of assigned subcarrier n_k to $D2D_k$, P_{k,n_k}, is determined by the resource allocation. Specifically, water-filling algorithm is used to allocate the maximum transmit power P_{max} of $D2D_k$ to the set of assigned subcarriers $\mathcal{N}_k$ according to $\{h_{k,n_k}\}_{n_k\in\mathcal{N}_k}$. However, as the $\{h_{k,n_k}\}_{n_k\in\mathcal{N}_k}$ is determined by subcarrier assignment for a large K, thanks to the multi-user diversity, the power allocated to subcarrier n_k can be approximated with the power equally allocated to $\mathcal{N}_k$ [79, 80],

$$P_{k,n_k}\approx\frac{P_{max}}{N_k},\forall k\in\mathcal{K},\ n_k\in\mathcal{N}_k. \tag{3.17}$$

Also, due to the limited transmit power, the transmit distance of any D2D pair is generally limited and similar to each other. Therefore, it is assumed that the distance of any D2D pair is fixed to one, hence (3.16) can be rewritten as

$$\bar{C}_{TOT} = K\cdot B\cdot\mathbb{E}_{\mathcal{N}_k,\{h_{k,n_k}\}_{n_k\in\mathcal{N}_k}}\left\{\sum_{n_k\in\mathcal{N}_k}\log_2\left(1+\frac{\xi\cdot|h_{k,n_k}|^2}{N_k}\right)\right\}, \tag{3.18}$$

where $\xi=\frac{P_{max}}{\sigma^2}$.

For BSCR scheme, each subcarrier is allocated to the D2D pair with the best CSI, therefore, fully multi-user diversity of K D2D pairs is obtained at each subcarrier assignment. As the square of magnitude of Rayleigh fading channel follows exponential distribution, according to order statistics of exponential distribution [81], the pdf of $|h_{k,n_k}|^2, \forall n_k \in \mathcal{N}_k$, can be obtained as

$$f_{|h_{k,n_k}|^2}(x) = K \cdot (1 - e^{-x})^{K-1} \cdot e^{-x}. \tag{3.19}$$

Also, since the subcarrier assignment is determined by $\{h_{k,n}\}_{k\in\mathcal{K},n\in\mathcal{N}}$, which is i.i.d. for all the K D2D pairs and N subcarriers, each subcarrier is allocated to $D2D_k$ with equal probability $\frac{1}{K}$. Therefore, the number of allocated subcarriers to $D2D_k$, N_k follows binomial distribution with the probability mass function (pmf) given by

$$\mathbb{P}_{BSCR}(N_k = n) = \binom{N}{n} \cdot \left(\frac{1}{K}\right)^n \cdot \left(1 - \frac{1}{K}\right)^{N-n}, \tag{3.20}$$

where $\binom{N}{n} = \frac{N!}{n!\cdot(N-n)!}$ represents the binomial coefficient.

By taking (3.19) and (3.20) into (3.18), the ergodic sum capacity for BSCR scheme could be derived as

$$\bar{C}_{BSCR} = K \cdot B \cdot \sum_{n=1}^{N} n \cdot \int_0^\infty \log_2\left(1 + \frac{\xi \cdot x}{n}\right) \cdot f_{|h_{k,n_k}|^2}(x)\, dx \cdot \mathbb{P}_{BSCR}(N_k = n). \tag{3.21}$$

For SAR scheme, the explicit distribution of $|h_{k,n_k}|^2$ and N_k are difficult to obtain directly, since the subcarriers are assigned iteratively and each iteration

is correlated with the preceding results. Although each subcarrier cannot be guaranteed to be assigned to the D2D pair with the best channel condition, (3.19) can be used to obtain the upper bound by assuming full multi-user diversity of K D2D pairs. Moreover, appendix A shows that when $\frac{N}{K}$ is large enough, the pdf of the number of allocated subcarriers per D2D pair for SAR scheme is fitted by

$$f_{\tilde{N}_k}^{SAR}(x; N, \frac{1}{K}) = \frac{x^{N-1} \cdot e^{-K \cdot x} \cdot K^N}{\Gamma(N)}, \tag{3.22}$$

where $\Gamma(N) = (N-1)!$, as N is a positive integer.

By taking (3.22), (3.19) into (3.18), an upper bound of the average ergodic sum capacity for SAR scheme can be obtained as

$$\tilde{C}_{SAR} \leq \bar{C}_{SAR}^{u} = K \cdot B \cdot \int_0^\infty \int_0^\infty y \cdot \log_2 \left(1 + \frac{\xi \cdot x}{y}\right) \cdot f_{|h_{k,n_k}|^2}(x)\, dx \cdot f_{\tilde{N}_k}^{SAR}(y)\, dy, \tag{3.23}$$

where the equality holds when $\frac{N}{K} \to \infty$.

Figs. 3.3 and 3.4 show the numerical and simulation results of sum capacity for both BSCR scheme and SAR scheme when the number of subcarriers N and the number of D2D pairs K varies respectively. It can be clearly seen that the numerical results of (3.21) matches the simulation results for BSCR scheme. For SAR scheme, the simulation result approaches the numerical results of (3.23) with negligible difference by increasing N or reducing K. This indicates that the upper bound derived in (3.23) will be tightened when $\frac{N}{K}$ tends to be infinite.

The optimal ergodic sum capacity is shown in Fig. 3.3, which is obtained

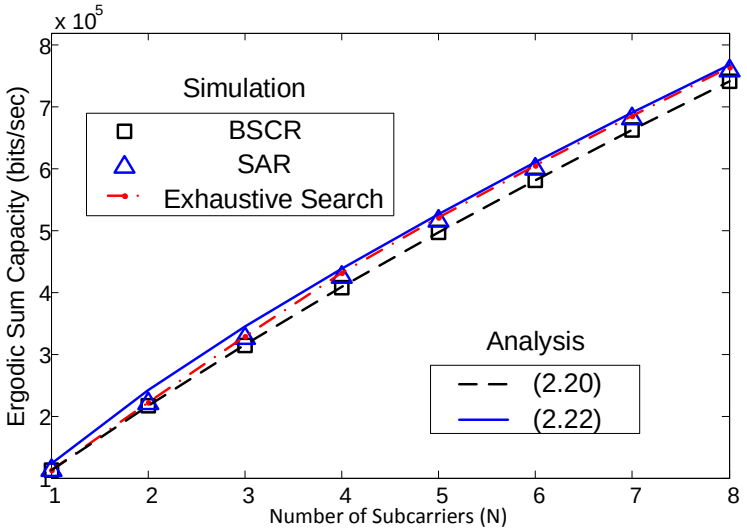


Figure 3.3: Ergodic capacities for different resource allocation schemes versus the number of subcarriers N, when $K = 4$, $\xi = 20$dB, $B = 15$KHz.

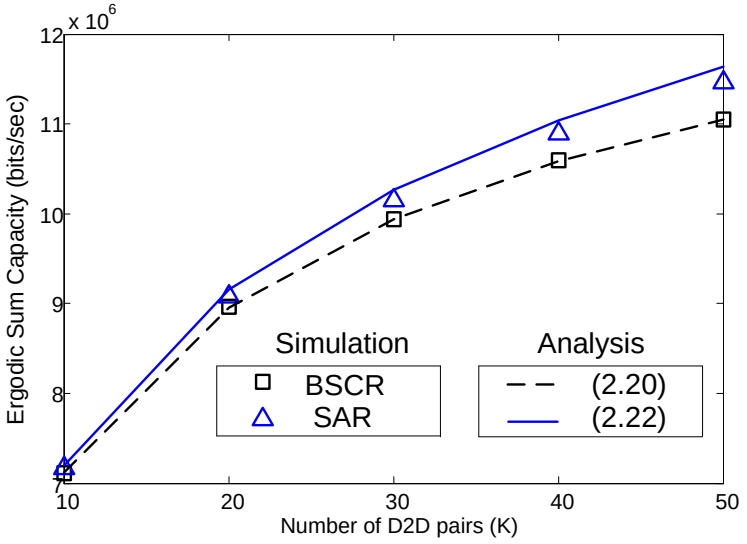


Figure 3.4: Ergodic capacities for different resource allocation schemes versus the number of D2D pairs K, when $N = 100$, $\xi = 20$dB, $B = 15$KHz.

by exhaustive search among all the possible subcarrier assignments for all K D2D pairs and N subcarriers. It is clearly shown that SAR scheme can always achieve near optimal ergodic sum capacity while the performance gap between the BSCR scheme and optimal results increases as N increases.

The closed form expression of sum capacities of both SAR and BSCR schemes are further derived and summarized in Theorem 1, to show the difference of sum capacities, which can be achieved from SAR and BSCR schemes.

Specifically, the closed form of (3.21), (3.23) is estimated by using second order of Taylor expansion to the channel gain, and number of subcarriers per D2D pair respectively.

Theorem 1. *The ergodic sum capacity of one cell for N subcarriers orthogonal assigned to K D2D pairs with BSCR-based resource allocation and SAR-based resource allocation can be approximated to*

$$\bar{C}_{BSCR} \approx B \cdot N \cdot \log_2\left(1 + \frac{\xi \cdot K \cdot \mathbb{H}(K)}{N}\right) - \frac{B \cdot \pi^2 \cdot \xi^2}{12 \cdot \log(2) \cdot (\xi \cdot K \cdot \mathbb{H}(K) + N)^2} - B \cdot (K-1)\,, \tag{3.24}$$

and

$$\bar{C}^u_{SAR} \approx \Omega(K, N, \xi), \tag{3.25}$$

where

$$\Omega(K,N,\xi)=\frac{B\cdot(N-1)^2}{N-2}\cdot\log_2\left(1+\frac{\xi\cdot K\cdot\mathbb{H}(K)}{N-1}\right)-\frac{B\cdot\pi^2\cdot(N-1)\cdot\xi^2\cdot K^2}{12\cdot\log 2\cdot(\xi\cdot K\cdot\mathbb{H}(K)+N-1)^2},\tag{3.26}$$

respectively.

Proof. See Appendix B □

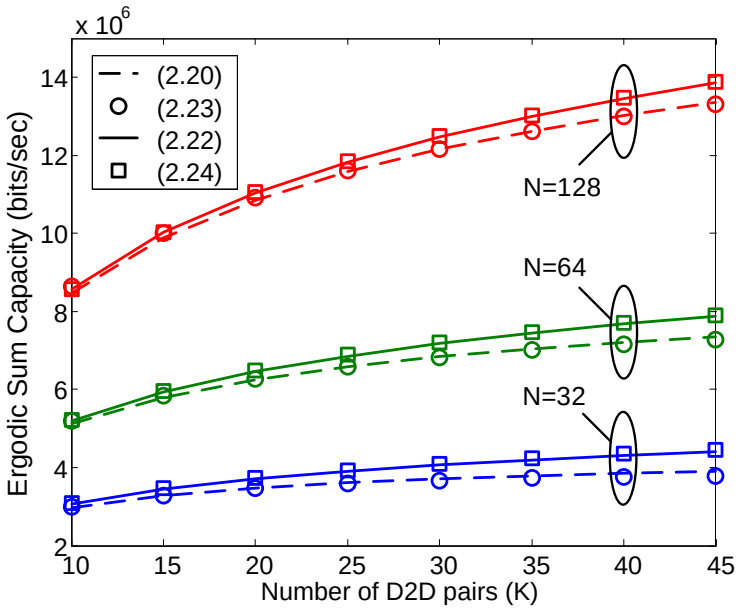


Figure 3.5: Average ergodic capacities for different resource allocation schemes versus number of D2D pairs within the cell when $N = 32, 64, 128$, $\xi = 20$dB, $B = 15$KHz

The numerical results of (3.21), (3.23), (3.24) and (3.25), are presented in Fig. 3.5, regarding to different K and N, respectively. As we can clearly see from the figure, (3.24), (3.25) serves as good approximation to (3.21),

(3.23) when both K and N scales. From (3.24) and (3.25), $\bar{C}_{TOT}^{BSCR}$ scales in the orders of $\Theta\left(N\log(\frac{K\log K}{N}) - K - \frac{1}{(K\log K+N)^2}\right)$ and $\tilde{C}_{TOT}^{SAR}$ scales in the orders of $\Theta\left(N\log(\frac{K\log K}{N}) - \frac{NK^2}{(K\log K+N)^2}\right)$. Therefore, it can be concluded that SAR scheme always achieves higher ergodic sum capacity than BSCR scheme, especially when K is large.

3.3 Sum Capacity under Multiple Sub-cell Structure

In contrast to the single sub-cell structure, the sum capacity of L sub-cells C_{TOT} is crucially affected by the inter D2D interference, as N subcarriers are universally reused among D2D pairs within different sub-cells. Since the sub-cell resource allocation is performed in each sub-cell without any prior knowledge of inter D2D interference, the optimal number of sub-cells L_{opt} becomes a vital tuning parameter to adjust the inter D2D interference between different sub-cells. Particularly, as the inter D2D interference is determined by the positions of all the D2D pairs within the cell, in this section, the *average ergodic sum capacity* of L sub-cell structure over all possible D2D pair locations is investigated by delving into the formation of the inter D2D interference among different sub-cells, and the optimal number of sub-cells $\hat{L}_{opt}$ that can maximize the average ergodic sum capacity is discussed.

3.3.1 Performance Analysis

As mentioned in section 3.1, the K D2D pairs are assumed to be distributed in the whole cell coverage according to a homogeneous PPP Φ_A with density

λ_A. With the proposed sub-cell division strategy, the D2D pairs within each sub-cell K_l is then distributed as a PPP $\Phi_A^{(l)}$ with density $\lambda_A^{(l)} = \lambda_A$ and the number of D2D pairs in the sub-cell l is Poisson distributed with probability mass function (pmf) shown as

$$\mathbb{P}(K_l = k) = \frac{\left(\pi \cdot {R_l}^2 \cdot \lambda_A^{(l)}\right)^k}{k!} e^{-\pi \cdot {R_l}^2 \cdot \lambda_A^{(l)}}. \tag{3.27}$$

For simplicity, let A_{k_l} be the location of $D2D_{k_l}$ with polar coordinates $(\gamma_{k_l}, \phi_{k_l})$, the average ergodic sum capacity $\bar{C}_{TOT}$ can be obtained from (3.7), as

$$\bar{C}_{TOT} = \mathbb{E}_{\Phi_A, \{h_{k,k',n}\}_{k \in \mathcal{K}, k' \in \mathcal{K}, n \in \mathcal{N}}} \left(\sum_{l=1}^{L} C_l \right), \tag{3.28}$$

where C_l is the sum capacity of sub-cell l, given by

$$C_l = \sum_{A_{k_l} \in \Phi_A^{(l)}} \sum_{n \in \mathcal{N}_{k_l}} B \cdot \log_2 \left(1 + \frac{P_{k_l,n} \cdot |g_{k_l,n}|^2}{I_{k_l,n}^c + \sigma^2} \right). \tag{3.29}$$

where $g_{k_l,n} = h_{k_l,n}$, as the transmit distance of $D2D_k$ is considered to be one.

Note that under PPP assumption, there are occasions when a sub-cell has no active D2D pair. Here, we denote active as the event of one sub-cell having at least one D2D pair within its coverage, which has the probability

$$\mathbb{P}(\text{active}) = 1 - \mathbb{P}(K_l = 0) = 1 - e^{-\pi \cdot {R_l}^2 \cdot \lambda_A^{(l)}}. \tag{3.30}$$

Since the sub-cells are geographic symmetrically divided within the cell, the active sub-cell set, which contains sub-cells having at least one D2D pair,

is denoted as $\mathcal{L}_{active}$, with $|\mathcal{L}_{active}| = L_{active}$. The pmf of L_{active} can be obtained as

$$\mathbb{P}(L_{active} = l) = \binom{L}{l} \cdot \left(1 - e^{-\pi \cdot {R_l}^2 \cdot \lambda_A^{(l)}}\right)^l \cdot \left(e^{-\pi \cdot {R_l}^2 \cdot \lambda_A^{(l)}}\right)^{L-l}. \tag{3.31}$$

For the purpose of illustration, the central sub-cell is assumed as the reference sub-cell l_r without loss of generality, and one typical D2D pair randomly located at $A_{k_r}(\gamma_{k_r}, \phi_{k_r})$ within sub-cell l_r is assumed as the reference D2D pair $D2D_{k_r}$. By combining (3.31) with (3.28), the average ergodic sum capacity can be rewritten as

$$\bar{C}_{TOT} = \sum_{l=1}^{L} l \cdot \mathbb{P}(L_{active} = l) \cdot \bar{C}_r, \tag{3.32}$$

where $\bar{C}_r$ is the average sum capacity of active sub-cell l_r, given by

$$\bar{C}_r = \mathbb{E}_{\{\Phi_A, \boldsymbol{H}\}} \left[\sum_{A_{k_r} \in \Phi_A^{(r)}} \sum_{n \in \mathcal{N}_{k_r}} B \cdot \log_2 \left(1 + \frac{P_{k_r,n} \cdot |h_{k_r,n}|^2}{I^c_{k_r,n} + \sigma^2}\right)\right]. \tag{3.33}$$

According to the sub-cell resource allocation, if one subcarrier is allocated to $D2D_{k_r}$, there must be one D2D pair $D2D_{k_{l'}}$ from each active sub-cell in $\mathcal{L}_{active}$ except l_r that reuses the same subcarrier, and thus introduces inter D2D interference to $D2D_{k_r}$. That is, there exist $L_{active} - 1$ inter D2D interference signals on each subcarrier that allocated to $D2D_{k_r}$.

Since D2D pairs are independent and identical distributed, the inter D2D interference at each subcarrier has the same distribution. Therefore, the inter D2D interference in this section is specified for one subcarrier $n_r \in \mathcal{N}_{k_r}$. It is

shown in [82] that in a dense D2D deployed system, the cumulative inter D2D interference signal can be modelled as a complex Gaussian random variable with zero mean and variance $I^v_{k_r,n_r}$, where

$$\mathrm{Var}[I^c_{k_r,n_r}] \approx I^v_{k_r,n_r} = \sum_{l' \in \mathcal{L}_{active} \backslash l_r} P_{k_{l'},n_r} \cdot {d_{k_r,k_{l'}}}^{-\alpha}, \tag{3.34}$$

and $P_{k_{l'},n_r}$ is the transmit power of $D2D_{k_{l'}}$ on subcarrier n_r, $d_{k_r,k_{l'}}$ is the distance between $D2D_{k_{l'}}$ to $D2D_{k_r}$, which is determined by the locations of both $D2D_{k_{l'}}$ and $D2D_{k_r}$.

Transmit power on subcarrier n_r, P_{k_l,n_r} Similar to the single sub-cell structure scenario, P_{k_l,n_r} is approximated by the average power of $D2D_{k_l}$ over allocated subcarriers, as

$$P_{k_l,n_r} \approx \frac{P_{max}}{N_{k_l}}, \tag{3.35}$$

where N_{k_l} is the number of allocated subcarriers to $D2D_{k_l}$, and determined by the single sub-cell resource allocation scheme.

As shown in Section 3.2.2, SAR scheme achieves near optimal ergodic sum capacity in single sub-cell structure, hence it is considered to be adopted by each active sub-cell in L sub-cell structure, and the pdf of $N_{k_{l'}}$ is given by (3.22).[2] Note that N_{k_l} is related to the number of D2D pairs K_l within the active sub-cell l and the pmf of $K_{l'}$ can be further derived from (3.27) and

[2]For the purpose of illustration, SAR scheme is assumed in this section. However, the analysis can be applied to any sub-cell resource allocation schemes in the same manner.

(3.30), as

$$\mathbb{P}(K_l = k|\text{active}) = \frac{\left(\frac{\bar{K}}{L}\right)^k \cdot e^{-\frac{\bar{K}}{L}}}{k!} \cdot \frac{1}{1 - e^{-\frac{\bar{K}}{L}}}. \tag{3.36}$$

Pdf of $d_{k_r,k_{l'}}$ Since the locations of the interference D2D pairs are random, it is assumed that θ_{k_r} interference D2D pairs of $D2D_{k_r}$ are uniformly distributed over the cell except l_r, and Appendix C shows that conditioned on the location of $D2D_{k_r}$, $A_{k_r}(\gamma_{k_r}, \phi_{k_r})$, the cumulative distribution function (cdf) $F_{d_{k_r,k_{l'}}|\gamma_{k_r}}(x|y)$ is given by

$$F_{d_{k_r,k_{l'}}|\gamma_{k_r}}(x|y) = \begin{cases} \frac{x^2 \arccos\frac{R_l^2 - x^2 - y^2}{2xy} - R_l^2 \arccos\frac{R_l^2 + y^2 - x^2}{2R_l y} + 2M(R_l)}{\pi(R^2 - R_l^2)}, \\ \qquad\qquad R_l - y < x \le R_l + y; \\ \frac{x^2 - R_l^2}{R^2 - R_l^2}, \qquad R_l + y < x \le R - y; \\ \frac{x^2 \arccos\frac{x^2 + y^2 - R^2}{2xy} + R^2 \arccos\frac{R^2 + y^2 - x^2}{2Ry} - 2M(R) - \pi R_l^2}{\pi(R^2 - R_l^2)}, \\ \qquad\qquad R - y < x \le R + y; \end{cases} \tag{3.37}$$

where

$$M(r) = \sqrt{\frac{r+x+y}{2}\left(\frac{r+x+y}{2} - r\right)\left(\frac{r+x+y}{2} - x\right)\left(\frac{r+x+y}{2} - y\right)}, \tag{3.38}$$

and γ_{k_r} is the distance between $D2D_{k_r}$ and the center of the cell with the pdf

$$f_{\gamma_{k_r}}(x) = \frac{2x}{R_l^2}, \qquad 0 \le x \le R_l. \tag{3.39}$$

By substituting (3.34), (3.35) into (3.33), the average ergodic sum capacity of the reference sub-cell can be expressed as

$$\begin{aligned}\bar{C}_r = \sum_{k=1}^{\infty} k \cdot \mathbb{P}(K_r = k|\text{active}) \int_0^N n \cdot f_{N_{k_r}}(n) \int_0^{\infty} f_{|h_{k_r,n}|^2}(h) \int_0^{R_l} f_{\gamma_{k_r}}(d_r) \\ \cdot \Bigg[\sum_{l' \in \mathcal{L}_{active} \backslash l_r} \sum_{k'=1}^{\infty} \mathbb{P}(K_{l'} = k'|\text{active}) \int_{2(L_{active}-1)} \cdots \int C_s \prod_{l' \in \mathcal{L}_{active} \backslash l_r} f_{N_{k_{l'}}}(n_{l'})\, dn_{l'} \\ \cdot f_{d_{k_r,k_{l'}}|\gamma_{k_r}}(d_{l'}|d_r) \mathrm{d}d_{l'} \Bigg]\, dd_r\, dh\, dn,\end{aligned} \tag{3.40}$$

where

$$C_s = B \cdot \log_2 \left(1 + \frac{P_{max} \cdot h}{n \left(\sum_{l' \in \mathcal{L}_{active} \backslash l_r} \frac{P_{max}}{n_{l'}} \cdot {d_{l'}}^{-\alpha} \right) + \sigma^2} \right). \tag{3.41}$$

It is seen in (3.40) that multi-fold integral operation is required to obtain the numerical results of (3.32), as the cumulative property of inter D2D interference leads to $3 \cdot (L_{active} - 1)$ random variables.

Note that the transmit power $P_{k_{l'},n_r}$ is i.i.d. in (3.34), the cumulative inter D2D interference for $D2D_{k_r}$ can be greatly simplified as

$$I^v_{k_r,n_r} \approxeq \tilde{I}^v_{k_r,n_r} = \frac{P_{max}}{\mathbb{E}\left(N_{k_{l'}}\right)} \cdot \sum_{l' \in \mathcal{L}_{active} \backslash l_r} {d_{k_r,k_{l'}}}^{-\alpha}, \tag{3.42}$$

where the equality holds when $\frac{L}{K} \to \infty$, i.e., each D2D pair reuses N subcarriers universally.

As (3.37) is a piecewise function with complicated sub-domain functions, (3.42) is difficult to tackle. Fortunately, it can be easily observed that when $D2D_{k_r}$ located at the central of the l_r, i.e. $\gamma_{k_r} = 0$, the sub-domain of first and third sub-functions of $F_{d_{k_r,k_{l'}}}$ are empty. Thus, (3.37) can be simplified as

$$F_{d_{k_r,k_{l'}}|\gamma_{k_r}=0}(x) = \frac{x^2 - R_l^2}{R^2 - R_l^2}, \qquad R_l < x < R \tag{3.43}$$

According to (3.4) and (3.43), the pdf of large scale fading gain, ${d_{k_r,k_{l'}}}^{-\alpha}$, can be obtained as

$$f_{{d_{k_r,k_{l'}}}^{-\alpha}|\gamma_{k_r}=0}(x) = \frac{2 \cdot x^{-\frac{2}{\alpha}-1}}{\alpha \cdot (R^2 - R_l^2)}, \qquad R^{-\alpha} < x < R_l^{-\alpha} \tag{3.44}$$

which indicates that ${d_{k_r,k_{l'}}}^{-\alpha}$ is a bounded Pareto-distributed random variable when $\gamma_{k_r} = 0$.

Theorem 2. *Let* $Y = \sum_{n=1}^{N} X_n$, *and* $r = \frac{Y}{\max\{X_1,\cdots,X_N\}}$. *If* $X_1, \cdots, X_n$ *follow i.i.d. bounded Pareto distribution,*

$$f(x) = \frac{\beta \cdot x^{-1-\beta}}{1 - c^{-\beta}}, \quad 1 < x < c, \quad \beta > 0 \tag{3.45}$$

then

$$\mathbb{E}(r) = \frac{1}{1-\beta} + N\left[1 - \frac{1}{1-c^{-\beta}} + \frac{1}{1-\beta}\left(\frac{c^{-\beta} - \beta c^{-1}}{1-c^{-\beta}} - \frac{B_{inc}(1-c^{-\beta}, N, \frac{1}{\beta})}{(1-c^{-\beta})^N}\right)\right] \tag{3.46}$$

where $B_{inc}(\cdot,\cdot,\cdot)$ *is the incomplete beta-function.*

Proof. See Appendix D. □

By exploring the heavy-tailed property of bounded Pareto distribution, the summation of $(L_{active}-1)$ i.i.d. Pareto random variables can be further simplified, as shown in Theorem 2. Therefore, the conditional cumulative interference received at $D2D_{k_r}$, when $\gamma_{k_r}=0$, can be approximated as

$$\tilde{I}^{v}_{k_r,n_r|\{\gamma_{k_r}=0\}} = \frac{P_{max}}{\mathbb{E}\left(N_{k_{l'}}\right)} \cdot \Psi(\alpha, u, L_{active}) \cdot {d^*_{k_r}}^{-\alpha} \tag{3.47}$$

where $u = \left(\frac{R_l}{R}\right)^{-\alpha} = \left(\frac{4L-1}{3}\right)^{-\frac{\alpha}{2}}$, $d^*_{k_r}$ denotes the smallest distance among the distances of $(L_{active}-1)$ inter D2D interference, and $\Psi(\alpha, u, L_{active})$ can be seen as the offset parameter conditional on $d^*_{k_r}$, which is

$$\begin{aligned}\Psi(\alpha, u, L_{active}) =& \mathbb{E}\left(\frac{\sum_{l'\in\mathcal{L}_{active}\setminus l_r} {d_{k_r,k_{l'}}}^{-\alpha}}{{d^*_{k_r}}^{-\alpha}}\right)\\ =& \frac{2}{2-\alpha} + (L_{active}-1)\left[1 - \frac{1}{1-u^{-\frac{\alpha}{2}}}\right.\\ &\left. + \frac{\alpha}{2}\left(\frac{u^{-\frac{\alpha}{2}} - \frac{2}{2-\alpha}u^{-1}}{1-u^{-\frac{\alpha}{2}}} - \frac{B_{inc}(1-u^{-\frac{\alpha}{2}}, (L_{active}-1), \frac{2}{\alpha})}{(1-u^{-\frac{\alpha}{2}})^{L_{active}-1}}\right)\right]\end{aligned} \tag{3.48}$$

Finally, the pdf of $d^*_{k_r}$ is obtained according to the order statistics theory [81], as

$$f_{d^*_{k_r}|L_{active}}(x|l) = (l-1)\cdot\left(1-F_{d_{k_r,k_{l'}}}(x)\right)^{l-2}\cdot f_{d_{k_r,k_{l'}}}(x), \qquad R_l \le x \le R \tag{3.49}$$

For $0 < \gamma_{k_r} \le R_l$, due to the growing dominant effect of $d^*_{k_r}$ to the cumulative interference received at $D2D_{k_r}$, $I^v_{k_r,n_r}$, the average cumulative interference $\mathbb{E}\{I^v_{k_r,n_r}\}$ is a monotonic increasing function from $\mathbb{E}\left[I^v_{k_r,n_r|\{\gamma_{k_r}=0\}}\right]$ to $\mathbb{E}\left[I^v_{k_r|\{\gamma_{k_r}=R_l\}}\right]$ with γ_{k_r} increases. Moreover, it can be easily observed that $\mathbb{E}\left[I^v_{k_r|\{\gamma_{k_r}=R_l\}}\right] = \infty$, since $d_{k_r,k_{l'}} \to 0$ when $\gamma_{k_r} = R_l$. Therefore, an upper bound of the $\bar{C}_{TOT}$ can be derived by using $\tilde{I}^v_{k_r|\{\gamma_{k_r}=0\}}$ in (3.47) as a lower bound of the cumulative interference.

The upper bound of the $\bar{C}_{TOT}$ can thus be obtained by substituting (3.47), (3.33) into (3.32), as

$$\begin{aligned}\tilde{C}^u_{TOT} =& B\cdot\sum_{l=1}^{L} l\cdot\mathbb{P}(L_{active}=l)\cdot\sum_{k_r=1}^{\infty} k_r\cdot\mathbb{P}(K_r=k_r|\text{active})\cdot\int_0^N n\cdot f_{N_{k_r}}(n)\\ &\cdot\int_0^{\infty} f_{|h_{k_r,n}|^2}(h)\int_{R_l}^{R} f_{d^*_{k_r}|(\gamma_{k_r}=0)}(x)\\ &\cdot\log_2\left(1+\frac{P_{max}\cdot h}{n_1\cdot\left(\frac{P_{max}}{\mathbb{E}\left(N_{k_{l'}}\right)}\cdot\Psi(\alpha,u,l)\cdot x^{-\alpha}+\sigma^2\right)}\right)dx\,dh\,dn\end{aligned} \tag{3.50}$$

Since the sub-cell resource allocation is independent with each other and the fading channels among D2D pairs are i.i.d., the integrations and summations in (3.50) can be interchanged, and $\tilde{C}^u_{TOT}$ can be further simplified

according to Theorem 1, as,

$$\bar{C}^{u}_{TOT} = \sum_{l=1}^{L} l \cdot \mathbb{P}(L_{active} = l) \sum_{k_r=1}^{\infty} \mathbb{P}(K_r = k_r | \text{active}) \int_{R_l}^{R} f_{d^*_{k_r}|(\gamma_{k_r}=0)}(x) \Omega(k_r, N, \xi_I(x, \mathbb{E}\left(N_{k_{l'}}\right))) \, dx \quad (3.51)$$

where $\xi_I(x, \mathbb{E}\left(N_{k_{l'}}\right)) = \frac{P_{max}}{\Psi(\alpha,u,l) \cdot P_{max} \cdot x^{-\alpha} / \mathbb{E}\left(N_{k_{l'}}\right) + \sigma^2}$.

3.3.2 Simulation Results and Discussions

In this sub-section, simulation results will be presented to demonstrate the performance of the proposed resource allocation algorithm with sub-cell division strategy by using Monte Carlo method with 10^4 sampling. As described in Section 3.1, we assume that K users are distributed in a single circular according to PPP, where each user is paired with a receiver located on its' surrounding circle with unit radius. The cell radius is then considered to be ten unit, according to the definition of proximity service in 3GPP [11,21]. It is assumed that the bandwidth of each subcarrier $B = 15$ KHz and the transmit power to noise ratio of each D2D pair, $\frac{P_{max}}{\sigma^2}$, is 30dB.

Fig. 3.6 shows the simulation results of average ergodic sum capacity of the cell under L sub-cell structure vary with the average number of D2D pairs $\bar{K}$. Also, the numerical results of $\bar{C}^{u}_{TOT}$ derived in (3.51) are shown in Fig. 3.6 as the upper bound of the average ergodic sum capacity.

Note that when $\bar{K}$ increases, the total transmit power within the cell increases linearly, as each D2D pair have individual transmit power P_{max}. Therefore, due to the combined effect of the multi-user diversity and the total transmit power, the average ergodic sum capacity of the cell under fixed L

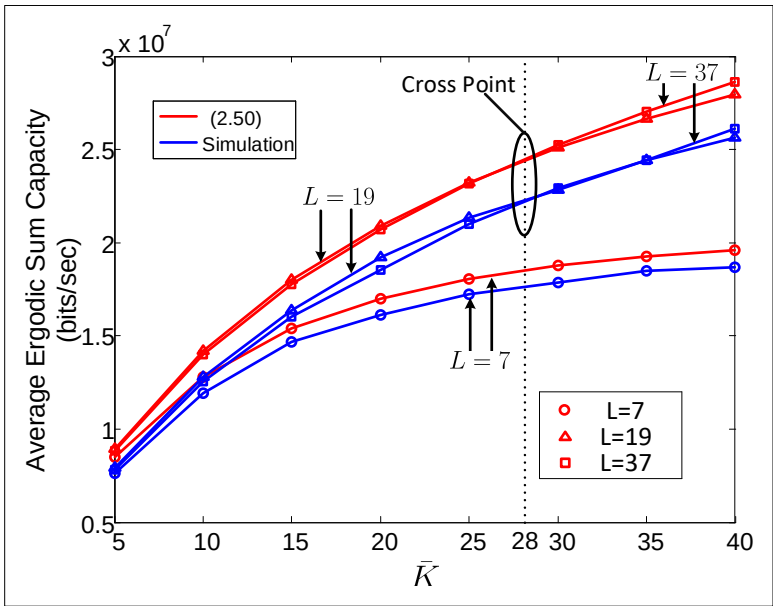


Figure 3.6: Ergodic sum capacities versus $\bar{K}$ with SAR-based sub-cell resource allocation scheme, where $\alpha = 3$ and $N = 32$.

sub-cell structure increases as $\bar{K}$ increases. Moreover, it can be clearly seen in Fig. 3.6 that the increasing slope of average ergodic sum capacity for different L structures are not the same. For instance, the gap between $L = 7$ and $L = 19$ increases when $\bar{K}$ increases. A cross point is shown between $L = 19$ and $L = 37$ at $\bar{K} = 28$, which indicates that when $\bar{K} < 28$, the average ergodic sum capacity of the system under $L = 19$ structure is larger than the one under $L = 37$. This phenomenon is caused by the effects of subcarriers being spatially reused in each active sub-cell and the accompanied cumulative inter D2D interference. Specifically, when L increases, the number of active sub-cells, L_{active} increases according to (3.31). As N subcarriers are reused in each active sub-cells, the frequency reuse factor increases as L_{active} increases, whereas in the mean time, average ergodic sum capacity is strongly limited by the cumulative inter D2D interference, which is a function of L_{active}, as shown in (3.47). Therefore, it concludes that the optimal number of sub-cells $\hat{L}_{opt}$

exists, which can maximize the average ergodic sum capacity of our proposed resource allocation scheme. As the cross point of $\bar{C}^u_{TOT}$ matches the simulation result in Fig. 3.6, (3.51) can be used to characterise the average system sum capacity under different L structure.

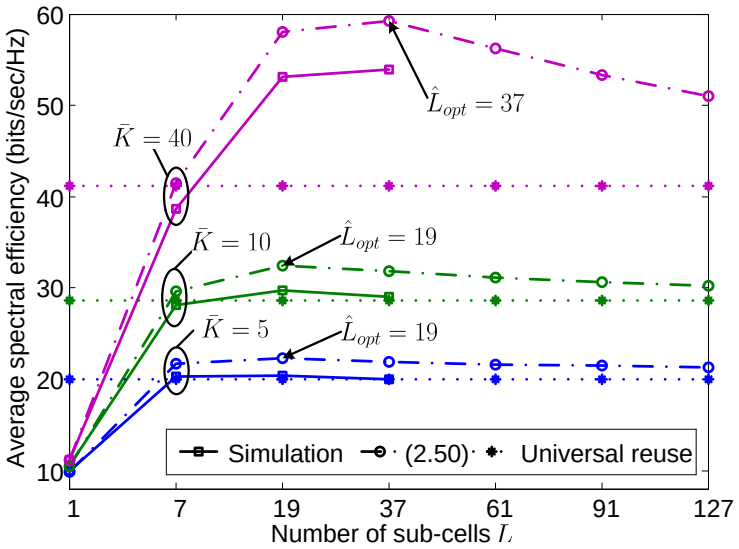


Figure 3.7: Average spectrum efficiencies versus number of subcells with SAR-based sub-cell resource allocation scheme when $\alpha = 3$ and $\frac{N}{K} = 1$.

For a better view of the optimal number of sub-cells, $\hat{L}_{opt}$, Figs. 3.7 and 3.8 present the simulation results of the average spectral efficiency as a function of number of sub-cells with various path loss exponents. Please note that the simulation results for the system with number of subc-cells larger than 37 is not achievable due to the extremely high complexity. The numerical result of $\frac{\bar{C}^u_{TOT}}{N \cdot B}$ is also shown as an upper bound of the average spectral efficiency, where $\bar{C}^u_{TOT}$ is given in (3.51). To eliminate the effect of total transmit power to the average spectral efficiency, the number of subcarriers N is set to be the same value as $\bar{K}$.

Since the CSI of interference channels between D2D pairs is not available at

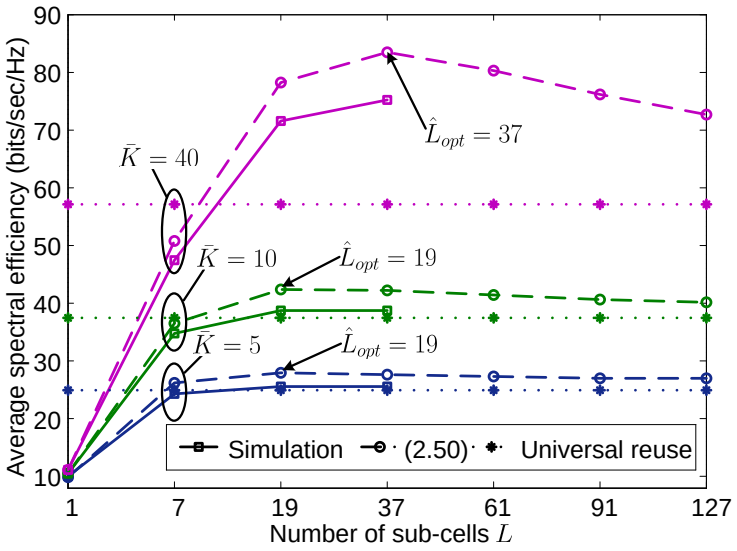


Figure 3.8: Spectral efficiency versus number of subcells with SAR-based sub-cell resource allocation scheme when $\alpha = 4$ and $\frac{N}{K} = 1$.

BS, without our proposed D2D sub-cell division strategy, the subcarriers are either orthogonally assigned to all D2D pairs (refer to as orthogonal scheme) or fully reused by all the D2D pairs (refer to as universal reuse scheme) in the cell. The average spectral efficiency of both schemes are shown in Figs. 3.7 and 3.8 as the benchmarks to our proposed sub-cell division based resource allocation schemes, respectively. Please note that the orthogonal resource allocation is equivalent to the single sub-cell scenario in our proposed scheme, and the numerical results of average spectral efficiency can be obtained via (3.25). It can be seen from Figs. 3.7 and 3.8 that the average spectral efficiency improves dramatically when L increases from one and concave down after reach a certain value. This is consistent with Fig. 3.6 and the analysis that there exists $\hat{L}_{opt}$ to maximize the average spectral efficiency. Further, the average spectral efficiency converge to the universal reuse results as L tent to be infinity. The reason is that as L increases, L_{active} will finally converge

to the number of D2D pairs when L tent to be infinity. Furthermore, it is noticeable that the gap between the optimal points to the universal reuse results increases significantly as $\bar{K}$ increases, which reveals that the average spectral efficiency can be greatly improved with $\hat{L}_{opt}$ obtained via (3.51) under proposed resource allocation scheme in dense D2D scenarios (e.g. $\bar{K} \geq 10$). For sparse D2D scenarios (e.g. $\bar{K} < 10$), $\hat{L}_{opt}$ can be approximated to 7 as the gap between optimal value to the value of $L = 7$ is insignificant.

3.4 Summary

In this chapter, the resource allocation for cellular D2D underlaid OFDMA system has been studied to maximize the average spectral efficiency. By taking the advantage of the spatial dispersive nature of D2D pairs, a simple, yet effective resource allocation scheme was proposed with significantly reduced CSI feedback compared to existing schemes. The theoretical analysis of system sum capacity/spectral efficiency was then performed through in-depth study of cumulative inter D2D interference. A tractable analytic expression of the average ergodic sum capacity was derived. The simulations results corroborate the accuracy of the obtained optimal number of sub-cells, which increases with number of D2D pairs. It also demonstrated that with small CSI feedback needed, substantial capacity/spectral efficiency gains can be achieved by cellular D2D communications with proposed resource allocation scheme. Finally, it was suggested that seven clusters is enough to obtain near optimal spectral efficiency for the practical system, where the number of D2D pairs in a single cell is small (e.g. less than 10).

Chapter 4

D2D Assisted Mobile Edge Computing Networks

In this chapter, a novel use case for cellular assisted D2D communication, D2D assisted MEC networks will be investigated. From the disciplines of both mobile computing and wireless communications, an admission access control scheme of D2D communication will be presented with the objective of minimizing the overall energy consumption of computation-intensive and latency-critical tasks at mobile devices.

This chapter is organized as follows. Section 4.1 describes the system model of D2D assisted MEC network. Section 4.2 presents the infeasibility problem and proposes a low complexity task allocation algorithm. The energy minimization problem is studied with a supplementary algorithm developed in section 4.3. Simulation results and discussions are given in Section 4.4, followed by the conclusions in Section 4.5.

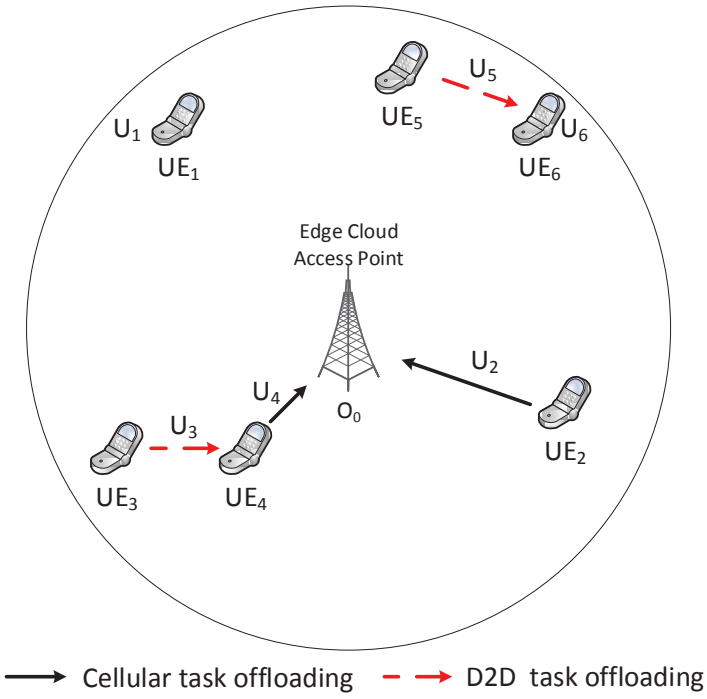


Figure 4.1: System model of D2D assisted MEC network.

4.1 System Model and Preliminary Analysis

As shown in Fig. 4.1, we consider that there are N active UEs, denoted by $\mathcal{N} = \{1, 2, ..., N\}$, each of which has one computation-intensive task to be executed. Let $\mathcal{U} = \{U_1, \cdots, U_i, \cdots U_N\}$ denotes the set of tasks to be executed at one time, where task U_i is assumed to be generated at UE i. For any task $U_i \in \mathcal{U}$, let F_i, D_i, $T_{i,max}$ denote the total number of the CPU cycles to be computed, the data size of offloading package and the corresponding time constraint of task U_i, respectively. [1]

It is assumed that D2D communication is available between any two UEs in the cell. As a result, each task can be offloaded either to the edge access point

[1]Note that F_i and D_i are system parameters defined by the specific cloud computing systems [83]. Without loss of generality, it is assumed that D_i is a fixed value and F_i is a uniformly distributed random variable in this chapter.

(AP) located at centre of the cell or to any other UE via D2D communication. The set of possible computing entities, including the edge AP and all the UEs that each task can be allocated to, is defined as $\mathcal{M} = \{0, \mathcal{N}\}$, with $|\mathcal{M}| = M$, where 0 is the index of the edge AP. $a_{i,j}$, $i \in \mathcal{N}$, $j \in \mathcal{M}$ is defined as the task indication parameter to show that whether the task from the UE i is allocated to the computing entities j. Thus, one has

$$a_{i,j} \in \{0, 1\}, \forall i \in \mathcal{N}, \forall j \in \mathcal{M}, \tag{4.1}$$

where $a_{i,j} = 1$ denotes that the UE i decides to allocate its task U_i to the computing entity j, otherwise task U_i is not allocated to the computing entity j. As each task can only be executed in one computing entity or may not be able to be completed anywhere due to lack of communication or computation resources, it can be easily obtained that

$$\sum_{j \in \mathcal{M}} a_{i,j} \leq 1, \forall i \in \mathcal{N}. \tag{4.2}$$

4.1.1 Energy Consumption for Task Execution

In this chapter, it is considered that the computing resource of computing entity j is denoted as f_j. Multiple tasks can be executed at one computing entity and its computing resource is shared by the tasks allocated to it. That is,

$$\sum_{i \in \mathcal{N}} a_{i,j} \cdot f_{i,j} \leq f_j, \forall j \in \mathcal{M}, \tag{4.3}$$

where $f_{i,j}$ denotes the allocated computing resource for task U_i at computing entity j.

The execution time of task U_i at computing entity j, $T^C_{i,j}$, is then obtained as

$$T^C_{i,j} = \frac{F_i}{f_{i,j}}, \quad \forall i \in \mathcal{N}, \forall j \in \mathcal{M}, \tag{4.4}$$

and the corresponding computation energy consumption, $E^C_{i,j}$, is give by [84]

$$E^C_{i,j} = T^C_{i,j} \cdot p^C_{i,j} = \kappa_j (f_{i,j})^{\nu_j - 1} F_i, \quad \forall i \in \mathcal{N}, \forall j \in \mathcal{M}, \tag{4.5}$$

where $p^C_{i,j}$ is the execution power of task U_i with allocated computing resource $f_{i,j}$. $\kappa_i > 0$ is the effective switched capacitance and $\nu_j \geq 2$ is the positive constant.

4.1.2 Energy Consumption for Task Offloading

Note that task U_i is not offloaded if $a_{i,i} = 1, \forall i \in \mathcal{N}$, since UE i executes its own task. If UE i decides to offload its task U_i to computing entity j where $j \in \mathcal{M} \setminus i$, the task data D_i has to be transmitted to the destination computing entity j. By assuming that the transmit power of each device is fixed to p^{Tr}, the data rate, $r_{i,j}$, can be obtained as

$$r_{i,j} = B \log_2 \left(1 + \frac{p^{Tr} \cdot |g_{i,j}|^2}{\sigma^2} \right), \quad \forall i \in \mathcal{N}, \forall j \in \mathcal{M} \setminus i, \tag{4.6}$$

where $g_{i,j}$ is the channel coefficient from UE i to computing entity j, and $|g_{i,j}|^2 = {d_{i,j}}^{-\alpha} \cdot |h_{i,j}|^2$. Here, $h_{i,j}$ is small-scale fading with $h_{i,j} \sim \mathcal{CN}(0,1)$, $d_{i,j}$ is the distance between UE i and computing entity j and α is the path-loss

exponent. σ^2 describes the additive white Gaussian noise which is assumed to be distributed as $\mathcal{CN}(0, \sigma^2)$. Note that dedicated spectrum with fixed bandwidth, B, is assumed to be allocated to each UE for task offloading, and the consumed transmission time of task U_i to computing entity j is given by

$$T_{i,j}^{Tr} = \begin{cases} \frac{D_i}{r_{i,j}} & \forall i \in \mathcal{N}, j \in \mathcal{M} \setminus i; \\ 0 & \forall i \in \mathcal{N}, i = j. \end{cases} \tag{4.7}$$

The transmission energy consumption of offloading task U_i can be then obtained as

$$E_{i,j}^{Tr} = p^{Tr} \cdot T_{i,j}^{Tr}, \ \ \forall i \in \mathcal{N}, j \in \mathcal{M}. \tag{4.8}$$

The energy consumption of each executed task consists of computation energy consumption and offloading energy consumption. Thus, the total energy consumption can be expressed as

$$E_{tot} = \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{M}} a_{i,j} \left(E_{i,j}^{Tr} + E_{i,j}^{C} \right) \tag{4.9}$$

It is obvious that if all the tasks in $\mathcal{U}$ are dropped, i.e. $\sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{M}} a_{i,j} = 0$, the minimum energy consumption can be achieved, with $E_{tot} = 0$, and the problem becomes trivial. Moreover, due to the limited computing resource of each entity, one latency-critical task may not be successfully completed under D2D assisted edge cloud offloading network. Therefore, in this chapter, we focus on the energy minimization problem while achieving the maximum number of feasible task under D2D assisted edge cloud offloading networks. For the sake of clarity, a two-stage optimization problem is formulated, which

contains the initial feasibility problem and the energy minimization problem.

4.2 Initial Feasibility Problem

4.2.1 Problem Formulation and Analysis

At the first stage, we aim to maximize the number of tasks which can be successfully completed, either through offloading to other UEs, the cloud or executing themselves. The initial feasibility optimization problem can be formulated as

$$\mathcal{P}1: \max_{\{a_{i,j}\},\{f_{i,j}\}} \sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{M}} a_{i,j}$$

$$\text{Subject to}: (4.1), (4.2), (4.3);$$

$$\sum_{j\in\mathcal{M}} a_{i,j} \left(T_{i,j}^{C}(f_{i,j}) + T_{i,j}^{Tr}\right) \leq T_{i,max}, \forall i \in \mathcal{N}, \tag{4.10}$$

where (4.10) follows the constraints that each task U_i is delay-sensitive and needs to be finished within $T_{i,max}$.

Since (4.2) indicates that each task can be assigned to one computing entity at most, the time constraint of each task can be replaced by the computing resource requirement of each task at each computing entity, $f_{i,j}$, by substituting (4.4) and (4.7) into (4.10), which is

$$f_{i,j} \geq f_{i,j}^{min} = \frac{F_i}{T_i - D_i/r_{i,j}}, \tag{4.11}$$

where $f_{i,j}^{min}$ denotes the minimum computing resource requirement of task U_i at computing entity j.

It can be easily seen that the equality of (4.11) holds for the optimal solution of $\mathcal{P}1$. Therefore, the optimal solution of $\mathcal{P}1$ can be shown as

$$f_{i,j}^{*} = a_{i,j}^{*} \cdot f_{i,j}^{min}. \tag{4.12}$$

Since it is not possible to allocate negative computing resource, $f_{i,j}^{*}$ needs to be positive value. That is, $a_{i,j}^{*} = 0$ when $f_{i,j}^{min} \leq 0$. Therefore, by combining (4.3) and (4.12), $\mathcal{P}1$ can be reformulated as

$$\mathcal{P}2: \quad \min_{\{a_{i,j}\}} \quad \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{M}} a_{i,j}$$

$$\text{Subject to}: (4.1), (4.2);$$

$$\sum_{i \in \mathcal{N}} f_{i,j}^{min} \cdot a_{i,j} \leq f_j, \forall j \in \mathcal{M}; \tag{4.13}$$

$$a_{i,j} = 0, \text{if } f_{i,j}^{min} \leq 0, \forall i \in \mathcal{N}, j \in \mathcal{M}. \tag{4.14}$$

The optimization problem $\mathcal{P}2$ reduces to a multiple knapsack problem, when $f_{i,j}^{min}$ is the same for task U_i at all the computing entities. Since the multiple knapsack problem is NP-hard and $\mathcal{P}2$ is at least as hard as the multiple knapsack problem, there is no computational efficient approach to solve $\mathcal{P}2$ optimally. Therefore, a low-complexity task selection algorithm is developed in next subsection.

4.2.2 Algorithm Design

As discussed in the previous subsection, the offloading cost for task U_i to computing entity j would lead to different minimum computing resource requirement, $f_{i,j}^{min}$, and the offloading cost for task U_i to computing entity j is only related to the offloading channel between computing entity i and j, $g_{i,j}$. Since UEs are assumed to be independently and identically distributed (i.i.d.) within the cell and dedicated spectrum is allocated for each possible offloading transmission, the offloading costs are independent with each other. That is, $f_{i,j}^{min}$ will not be affected by the task allocation result. However, each computing entity has limited amount of avaliable computing resource, f_j, as shown in (4.13). To illustrate the effect of (4.13) and (4.14), we therefore define the auxiliary weight variable of task U_i at computing entity j as $\Delta f_{i,j}$, shown as

$$\Delta f_{i,j} = \begin{cases} \frac{f_{i,j}^{min}}{f_j'} & \text{if } f_{i,j}^{min} > 0; \\ \infty & \text{otherwise,} \end{cases} \tag{4.15}$$

where f_j' denotes the current avaliable computing resource for entity j.

Based on the defined auxiliary variable $\Delta f_{i,j}$, an iterative task allocation (ITA) algorithm can be provided to solve $\mathcal{P}2$. The main idea is to allocate N tasks through N iterations, as each task can only be allocated to one computing entity. During iteration k, task U_{i^*} is allocated to computing entity j^* and $a_{i^*,j^*} = 1$, where $(i^*, j^*) = \arg\min[\Delta f_{i,j}]_{(N-k+1)\times M}$ and $[\Delta f_{i,j}]_{(N-k+1)\times M}$ denotes the auxiliary weight variable matrix for the remaining $N-k+1$ tasks at M computing entities. After iteration k, $\{\Delta f_{i,j^*} : \forall i \in \mathcal{N} \setminus i^*\}$ needs to be

updated according to (4.15), where

$$f'_{j^*} = f_{j^*} - \sum_{i \in \mathcal{N}} f^{min}_{i,j^*} \cdot a_{i,j^*}. \tag{4.16}$$

Algorithm 2 Iterative Task Allocation (ITA) algorithm

Initialization: Iterative indicator k=1, the maximum number of iterations as N. Set the task allocation indicator matrix $[a_{i,j}]_{N\times M} = \mathbf{0}$, calculate $[f^{min}_{i,j}]_{N\times M}$ according to (4.11), $[\Delta f_{i,j}]^{(0)}_{N\times M}$ according to (4.15) .
if $k < N$ and $\min[\Delta f_{i,j}]^{(k)}_{(N-k)\times M} < 1$ **then**
  Based on $[\Delta f_{i,j}]^{(k-1)}_{(N-k+1)\times M}$, calculate $\mathcal{U}_q^{(k-1)}$.
  if $\mathcal{U}_q^{(k-1)} \neq \emptyset$ **then**
    update $[\Delta f_{q,j}]^{(k)}_{|\mathcal{U}_q|\times M}$,then search for $(q^*, j^*) = \arg\min[\Delta f_{q,j}]^{(k)}_{|\mathcal{U}_q|\times M}$ and set $a_{q^*,j^*} \leftarrow 1$
  else
    according to $[\Delta f_{i,j}]^{(k-1)}_{(N-k+1)\times M}$, search for $(i^*, j^*) = \arg\min[\Delta f_{i,j}]^{(k-1)}_{(N-k+1)\times M}$ and update $[a^*_{i,j}]_{N\times M}$.
  end if
  Update $[\Delta f_{i,j}]^{(k)}_{(N-k)\times M}$ according to (4.15) and (4.16).
else
  Output $[a^{(p1)}_{i,j}]_{N\times M} = [a^*_{i,j}]_{N\times M}$.
end if

Note that during the iterative process, there might exist some special tasks, which have only one feasible entity, so called *exclusive* tasks. Specifically, since the mobile UEs are randomly distributed in the system, the number of available computing entity for each task is different due to its' location and the individual time constraint (4.10). The *exclusive* task can only be process at one idle entity, which fulfils its' time constraint. Therefore, to maximize the number of feasible tasks, we propose to give priority to the *exclusive* task set, denoted by $\mathcal{U}_q \subseteq \mathcal{U}$. That is, if $\mathcal{U}_q \neq \emptyset$ at the beginning of iteration k, then

one of the *exclusive* task U_{q^*} needs to be allocated to entity j^* at iteration k and $a_{q^*,j^*} = 1$, where $(q^*, j^*) = \arg\min[\Delta f_{q,j}]_{|\mathcal{U}_q| \times M}$.

It can be easily seen that after N iterations at most, the iterative task allocation (ITA) algorithm converges, and the task allocation matrix $[a_{i,j}^{(p1)}]_{N \times M}$ can be obtained and the ITA algorithm is summarized in Algorithm 2.

4.3 Energy Minimization Problem

In this section, we focus on investigating the energy minimization problem while maintaining the maximum number of feasible tasks achieved by solving the initial feasibility problem.

4.3.1 Problem Formulation

As discussed in section III, minimum computing resource requirement of task U_i at computing entity j, $f_{i,j}^{min}$, can be obtained from its individual time constraint. Note that the offloading energy consumption, $E_{i,j}^{Tr}$, is determined by the task allocation, and the computation energy consumption $E_{i,j}^{C}$ is proportional to the allocated computing resource $f_{i,j}$. Thus, by substituting (4.11) into (4.9), a lower bound of the total energy consumption can be expressed as

$$E_{tot} \geq \tilde{E}_{tot} = \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{M}} a_{i,j} \left(\frac{p^{Tr} \cdot D_i}{r_{i,j}} + \kappa_j F_i (f_{i,j}^{min})^{\nu_j - 1} \right), \tag{4.17}$$

and the energy minimization problem can be therefore shown as

$$\mathcal{P}3: \min_{\{a_{i,j}\}} \tilde{E}_{tot}$$

$$\text{subject to}: (4.1), (4.2), (4.13), (4.14);$$

$$\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} a_{i,j} = |\mathcal{U}_F|, \tag{4.18}$$

where $\mathcal{U}_F$ denotes the largest feasible task set and $|\cdot|$ denotes the cardinality of the set.

In general, with any given feasible task set $\mathcal{U}_F$, (4.2) and (4.18) can be integrated into one constraint, shown as

$$\sum_{j\in\mathcal{M}} a_{i,j} = 1, \ \forall i \in \mathcal{U}_F, \tag{4.19}$$

and $\mathcal{P}3$ with given feasible task set $\mathcal{U}_F$, can be seen as a typical general assignment problem, which is NP-hard. Therefore, a low complexity energy minimization algorithm is developed in the following subsection.

4.3.2 Algorithm Design

Based on the feasible task allocation matrix, $[a_{i,j}^{(p1)}]_{N\times M}$, obtained via ITA algorithm, the largest feasible task set is defined as $\mathcal{U}_F^{(p1)} = \{U_f : \sum_{j\in\mathcal{M}} a_{f,j}^{(p1)} = 1, \forall U_f \in \mathcal{U}\}$. For each task U_f, it can be seen from (4.11) that with known data rate between entity f and entity j, $r_{f,j}$, $f_{f,j}^{min}$ is a deterministic parameter, which monotonically decreases as $r_{f,j}$ increases.

Moreover, (4.5) shows that $E_{i,j}^{C}$ is a monotonically increasing function of

$f_{i,j}^{min}$, as $(\nu_j - 1)$ is always larger than 1. Therefore, it can be easily observed that for one feasible task U_f, the energy consumption at the feasible entity j, $E_f = E_{f,j}^{Tr} + E_{f,j}^{C}$, increases with $f_{f,j}^{min}$. That is, according to (4.17), $\tilde{E}_{tot}$ becomes larger when the entity j with a larger $f_{f,j}^{min}$ is chosen. Note that E_f is feasible only when $f_{i,j}^{min} > 0$, as shown in (4.14).

Based on the above analysis, an iterative *switching* task (IST) algorithm is then proposed to minimize the energy consumption of the largest feasible task set $\mathcal{U}_F^{(p1)}$.

Specifically, the *switching* operation is the switching of each task U_i from entity j, with $a_{i,j}^{(p1)} = 1$, to entity j'. There are two necessary conditions need to be satisfied to form a possible *switching* pair (j, j') for task U_i: (1) $f_{i,j'}^{min} < f_{i,j}^{min}$; and (2) entity j' must have enough computing resource left to fulfil $f_{i,j'}^{min}$. That is,

$$f_{j'}^{Re} = f_{j'} - \sum_{U_l \in \mathcal{U}_F^{(p1)}} a_{l,j'} \cdot f_{l,j'}^{min} \geq f_{i,j'}^{min}, \tag{4.20}$$

where $f_{j'}^{Re}$ denotes the residual computing resource at the computing entity $j', \forall j' \in \mathcal{M}$.

For each task $U_f \in \mathcal{U}_F^{(p1)}$, if there exists multiple entities to pair with entity j, denoted by $\{j'_o : o = 1, \cdots, O\}$, that satisfy the condition of the *switching* operation, task U_f is assigned to entity j^* with the minimum f_{f,j'_o}^{min}, i.e., $j^* = \arg\min f_{f,j'_o}^{min}$.

After $\left|\mathcal{U}_F^{(p1)}\right|$ iterations, the task allocation matrix $[a_{i,j}]_{N \times M}$ can be obtained. However, since only one task is checked for the *switching* operation at one iteration, if *switching* operation for task U_k is done at iteration k, the previous $k-1$ tasks might also have available switching pairs since the resid-

ual computing resource for the two switched entities changed. Therefore, an outer iteration is needed if *switching* operation is performed during the $\left|\mathcal{U}_F^{(p1)}\right|$ iterations. The IST algorithm is summarized in Algorithm 3.

Algorithm 3 Iterative Switching Task (IST) algorithm

1: **Initialization:** Outer iteration number $ot = 1$. Set the task allocation indicator matrix $[a_{i,j}^{(1)}]_{N\times M} = [a_{i,j}^{(p1)}]_{N\times M}$ and calculate $[f_{i,j}^{min}]_{N\times M}$ according to (4.11).
2: **repeat**
3: Set task iterative number $k = 1$, number of switching operations $N_s = 0$, and the maximum number of iterations as $\left|\mathcal{U}_F^{(p1)}\right|$.
4: **repeat**
5: Calculate the residual computing resource for each entity $j \in \mathcal{M}$, $\{f_j^{Re}\}^{(k-1)}$. For $U_k \in \mathcal{U}_F^{(p1)}$, obtain j_k, where $a_{k,j_k}^{(ot)} = 1$.
6: **while** $f_{k,j_k}^{min} \neq \min f_{k,j}^{min}$ **do**
7: do *switching* operation, $N_s \leftarrow N_s + 1$ and update $[a_{i,j}^{(ot)}]_{N\times M}$.
8: **end while**
9: $k \leftarrow k + 1$;
10: **until** $k = \left|\mathcal{U}_F^{(p1)}\right|$
11: $ot \leftarrow ot + 1$;
12: **until** $N_s \neq 0$
13: Output $[a_{i,j}^{(p3)}]_{N\times M} = [a_{i,j}^{(ot)}]_{N\times M}$

4.4 Simulation Results

In this section, simulation results are presented to show the benefits of D2D assisted MEC networks and evaluate the performance of the proposed algorithms. It is assumed that N UEs are uniformly i.i.d. in the cell with radius 100 meters and the edge AP is located at the centre of the cell. The edge cloud AP is assumed to have $f_0 = 20$GHz computing capability. Each user has random computing capability f_j varying from 0GHz to

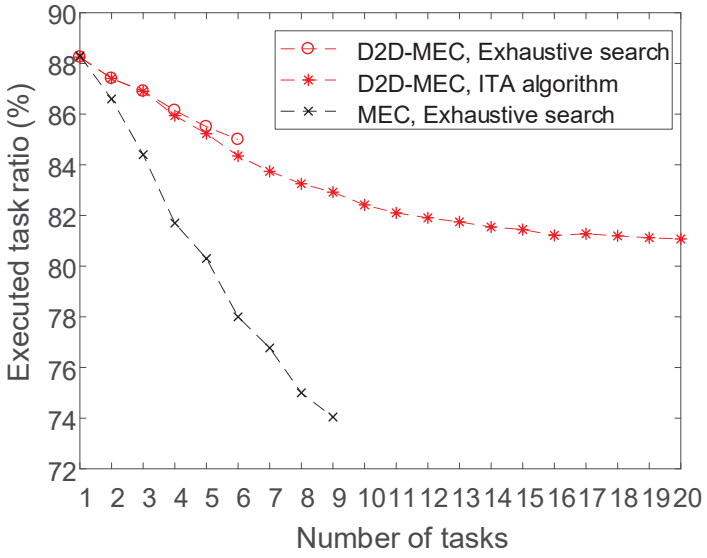


Figure 4.2: Executed task ratio versus the number of tasks in traditional MEC networks and D2D assisted MEC networks with various admission algorithms.

10GHz, for all $j \in \{1, \cdots, N\}$ and the transmit power is fixed at 10^{-3} watt. Each user randomly generates one required task with $D_i = 1, \forall i \in \mathcal{N}$, and $T_{i,max} = 1, \forall i \in \mathcal{N}$. [2] The required CPU cycles F_i is randomly distributed between 0GHz to 10GHz. Rayleigh small scale fading is assumed with zero mean and unit variance.

4.4.1 Maximum Executed Task Ratio δ

Fig. 4.2 presents the executed task ratio, defined as $\delta = \frac{|\mathcal{U}_F|}{N}$, in MEC networks and D2D assisted MEC networks. The maximum executed task ratio of traditional MEC networks and D2D assisted MEC (D2D-MEC) networks are obtained by using the exhaustive search method. Specifically, each task can

[2]The data size of the offloading package D_i depends on the specific system framework in mobile cloud computing networks [83]. For the sake of simplicity, D_i is fixed to one in the simulation. However, D_i can be set to any non-specific value.

be executed by the original UE or the edge AP for traditional MEC networks, otherwise the task cannot be accomplished. By exhaustive search through δ of all the possible 3^N combinations for N tasks, the maximum executed task ratio of traditional MEC networks is obtained. Similarly, the maximum task executed ratio of the D2D assisted MEC networks is obtained by exhaustive search through δ of all the possible N^{N+2} combinations as each task have $(N+2)$ process options. As the exhaustive search method has computation complexity of $O(N^{N+2})$ for D2D assisted MEC networks and $O(3^N)$ for MEC networks, the simulation results are obtained up to $N = 9$ for MEC networks and up to $N = 6$ for D2D assisted MEC networks. It is shown in Fig. 4.2 that ITA algorithm can achieve nearly the same executed task ratio as the optimal exhaustive search. It is clearly shown that with the assistance of D2D communication, the executed ratio converges gently to 81% when $N = 20$. By contrast, the executed ratio rapidly drops to 74% in traditional MEC network when $N = 9$.

It can be foreseen that in dense user scenario, although the computing resource of AP is much more than that of an individual UE, it cannot fulfil the large quantity of offloading requests. The executed task ratio will converge to 50% in traditional MEC network once the AP is fully occupied. This is because the computing capability of any UE j, f_j, and computing requirement of each task F_i are uniformly distributed with the same mean and variance. By contrast, in D2D assisted MEC network, the executed task ratio is much higher than that in the traditional MEC network. In fact, it is almost converges to more than 80% in dense user scenario, which implying that the user experience is improved significantly.

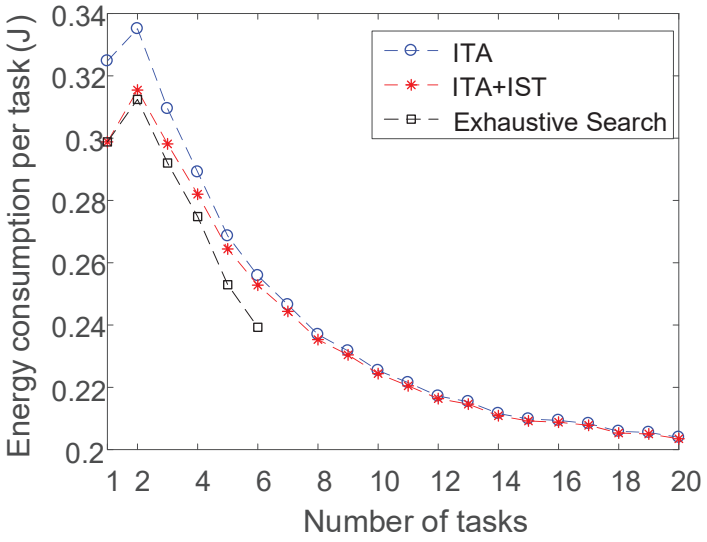
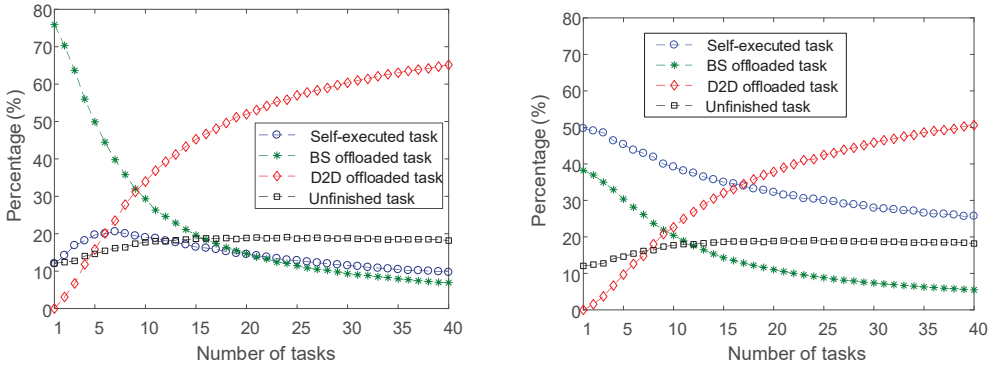


Figure 4.3: Energy consumption per executed task versus the number of tasks in D2D assisted MEC networks.

4.4.2 Energy Consumption per Executed Task

Fig. 4.3 presents the energy consumption per executed task as a function of the number of tasks in D2D assisted MEC networks. In general, it can be seen that the simulation results of ITA+IST decreases dramatically from 0.318 J at $N = 2$ to 0.203 J at $N = 20$. It indicates that the energy consumption per executed task can be reduced significantly in D2D assisted MEC networks, especially in dense user scenario. The increment of energy consumption from $N = 1$ to $N = 2$ is due to the fact that more power is desired in D2D assisted MEC as more task can be executed. However, this effect will be mitigated as the number of users increases. The optimal energy consumption results are obtained via exhaustive search from $N = 1$ to $N = 6$. It is shown that with the supplementary IST algorithm adopted after ITA algorithm, the energy consumption is greatly reduced and close to the optimal results when the



(a) ITA algorithm (b) ITA+IST algorithm

Figure 4.4: Task offloading percentage of different computing entities versus the number of tasks.

number of tasks is small. It is also shown in Fig. 4.3 that the performance of ITA algorithm converges to ITA+IST algorithms in dense user scenario. This is because when the number of tasks increases, almost full system computing resource is utilized, leading to little margin for further energy optimization.

Fig. 4(a-b) show the offloading percentage among different computing entities for both algorithms. It can be clearly seen that most tasks are offloaded to AP when the number of tasks is small in ITA algorithm. This is due to the fact that AP has much more computing resource than individual UEs. However, the IST algorithm revises the energy-inefficient allocations by letting more UEs compute their own generated tasks. Moreover, in dense user scenario, less than 10% of total tasks are offloaded to AP, implying that only some extremely computation-intensive tasks which cannot be handled by D2D offloading are offloaded to AP.

4.5 Summary

In this chapter, the task infeasibility problem and the energy minimization problem in D2D assisted MEC network are investigated. Specifically, a low-complexity ITA algorithm is provided for maximizing the number of executed tasks with the latency constraint of each individual task and the computing resource constraint of each computing entity. A supplementary IST algorithm is then developed for global energy minimization. Simulation results show that combining ITA and IST algorithms can achieve near-optimal performance by comparing with the optimal exhaustive search method. Moreover, it is further shown that the D2D assisted MEC network achieves higher executed task ratio than the traditional MEC network, and the energy consumption for executing each task is significantly reduced, especially in dense user scenarios.

Chapter 5

Conclusion and Future Work

5.1 Conclusion

The future mobile communication systems are expected to support various kinds of services for massive instant user devices. With centralized resource management of cellular networks, D2D communication can effectively utilize the ever increasing user density to achieve significant performance gains and is one of the key technologies for user-centric communications. However, the heterogeneity and complexity of the resource management brings severe challenges to both the theoretical evaluation and practical realization of D2D underlaid cellular network.

In this thesis, we demonstrated that the cellular assisted D2D communication can significantly improve the system performance with low complexity and practical resource management, and it has substantial value to the user centric networks, such as MEC networks, as the basic underlying component. Specifically, the resource management of cellular assisted D2D communication has been investigated in two aspects: 1) the resource allocation and performance analysis of cellular assisted D2D communication, and 2) the admission

access control in D2D assisted MEC networks.

In Chapter 3, a simple resource allocation scheme for cellular assisted D2D communication was proposed with significantly reduced CSI feedback compare to existing schemes. Specifically, the resource allocation was formed by the geography based sub-cell division strategy and sub-cell resource allocation. By in-depth study of inter D2D interference, a tractable analytical expression of average sum capacity for cellular D2D communication was firstly derived to show that D2D communication can greatly improve the spectral efficiency, especially in dense D2D user environments.

With the coordination of cellular network, D2D communication can be utilized to improve the performance of MEC network, which is a typical user centric network. In conventional MEC networks, a large potion of computation-intensive and latency-critical tasks cannot be supported, which affects the QoE greatly, especially for cell-edge users. To tackle this issue, Chapter 4 suggested that some tasks can be offloaded to the nearby devices via cellular assisted D2D communication. The admission access control of D2D communication should thus follow the disciplines from both mobile computing and wireless communication. It was shown that with optimized admission access control, both the number of infeasible tasks and the energy consumption per task can be reduced significantly.

5.2 Future Work

The work in this thesis only provides a starting point. As resource allocation of cellular assisted D2D communication was well studied in Chapter 3, it can be extended to the D2D communication underlaid cellular networks in

the future. That is, the resource allocation for both D2D pairs and cellular users can be jointly considered with the admission access control of D2D communications. In Section 3.1, the number of D2D pairs and the dedicated licensed spectrum allocated for D2D communications were assumed to be fixed system parameters and the cellular users were not considered. However, due to the heterogeneity of D2D communication, the resource allocation between cellular users and D2D pairs can be jointly optimized in D2D communication underlaid cellular networks with fixed total number of users within the cell. Moreover, as resource allocation of D2D communication for a single cell is investigated, how to expand the proposed resource allocation scheme to the multiple cell scenario is another interesting topic.

In the second part of this thesis, the admission access control of D2D assisted MEC networks was investigated by assuming that at most one dedicated subcarrier is allocated to one user for the task offloading. How to further optimize the resource allocation and admission access control design of D2D assisted MEC networks with spectral reuse to improve system performance is an interesting topic that deserves much attention in future study.

Finally, with the interference analysis in Section 3.3, the data rate and the access distance of D2D pairs can be characterized in the future studies to define the proximity requirements of cellular assisted D2D communications. The proximity requirements may serve as a key parameter for designing distributed D2D resource allocation schemes, and the D2D discovery and establishment protocols.

Appendix A

Validation of (3.22)

Since the small-scale channel fading factors are i.i.d. for all K D2D pairs, the mean value of N_k is $\frac{N}{K}$. As N_k is a discrete value, it is difficult to obtain the explicit probability mass function (pmf). By relaxing N_k into a continuous variable $\tilde{N}_k$, the gamma distribution can be used to approximate the pdf of N_k, with the shape parameter approximated by N and the rate parameter approximated by K. The corresponding pdf of $\tilde{N}_k$ is then given by

$$f_{\tilde{N}_k}^{SAR}(x; N, \frac{1}{K}) = \frac{x^{N-1} \cdot e^{-K \cdot x} \cdot K^N}{\Gamma(N)}. \tag{A.1}$$

In order to show the accuracy of the approximation, the pmf of N_k under 10^5 subcarrier assignment results with SAR-based resource allocation scheme were shown in Fig. A.1, compared with the numerical result of (A.1). It can be seen that given the number of D2D pairs $K = 4$, when the total number of subcarriers is relatively large, e.g. larger than 40, the approximated Gamma distribution for $\tilde{N}_k$ is very close to the practical simulation results of N_k. This is due to the fact that gamma function is continuous while the N_k is a positive integer. Hence, it is clear that the approximation becomes more accurate as $\frac{N}{K}$ enlarges.

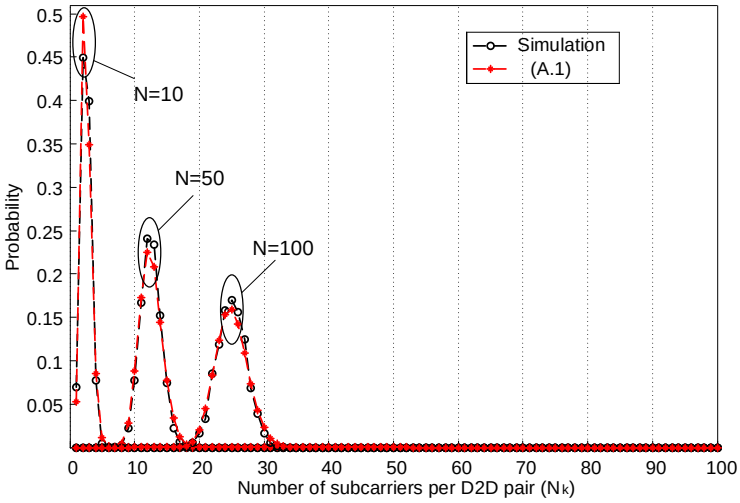


Figure A.1: The probability of number of allocated subcarriers per D2D pair when $\frac{P_{max}}{\sigma^2} = 20\text{dB}$, $K = 4$

Appendix B

Proof of Theorem 1

From the pdf of the best subchannel's channel fading factor $|h_{k,n_k}|^2$ in (3.19), the mean and variance of $|h_{k,n_k}|^2$ can be derived as

$$\mu_{h^*} = \int_0^\infty x \cdot K \cdot \left(1 - \exp^{-x}\right)^{K-1} \cdot \exp^{-x} \, dx = \mathbb{H}(K) \overset{K\to\infty}{\sim} \log(K), \quad \text{(B.1)}$$

and

$$\begin{aligned} \sigma_{h^*}^2 &= \int_0^\infty x^2 \cdot K \cdot \left(1 - \exp^{-x}\right)^{K-1} \cdot \exp^{-x} \, dx - \mathbb{H}(K)^2 \\ &= \frac{\pi^2}{6} - \psi^{(1)}(1+K) \overset{K\to\infty}{\sim} \frac{\pi^2}{6}, \end{aligned} \quad \text{(B.2)}$$

respectively, where $\mathbb{H}(K) = \sum_{k=1}^{K} \frac{1}{k}$ is the harmonic number with its asyptotic limit $\gamma + \log(K)$ and $\gamma \approx 0.5772156649$ is the Euler-Mascheroni constant. $\psi^{(1)}(z)$ is the polygamma function of order 1. According to Taylor's theory, the analytic function of (3.18) centred at $\bar{C}_{TOT}(\mu_{h^*})$ is shown as

$$\bar{C}_{TOT} = K \cdot B \cdot \mathbb{E}_{\mathcal{N}_k, \{h_{k,n_k}\}_{n_k \in \mathcal{N}_k}} \left\{ \sum_{n_k \in \mathcal{N}_k} \sum_{m=0}^{\infty} \frac{g^{(m)}(\mu_{h^*})}{m!} \cdot \left(|h_{k,n_k}|^2 - \mu_{h^*}\right)^m \right\}, \quad \text{(B.3)}$$

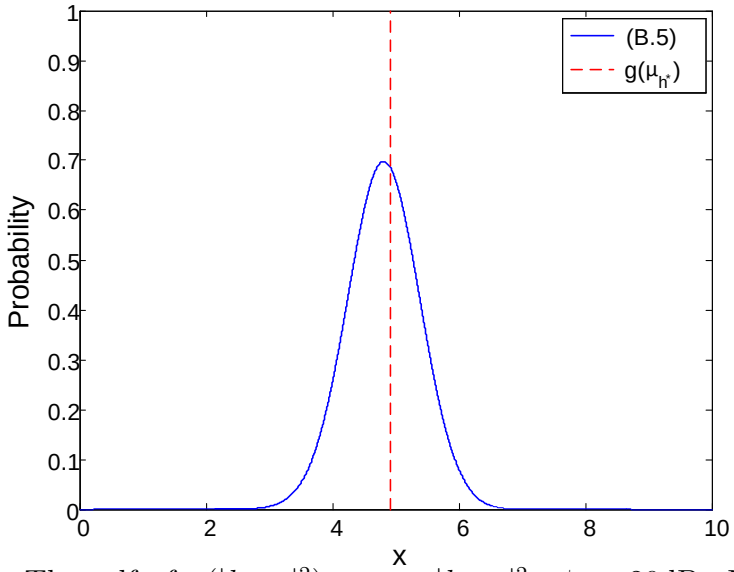


Figure B.1: The pdf of $g(|h_{k,n_k}|^2)$ versus $|h_{k,n_k}|^2$. $\xi = 20$dB, $N = 10$ and $K = 10$.

where $g^{(m)}(x)$ denotes the mth derivative of $g(x)$, and

$$g(x) = \log_2\left(1 + \frac{\xi \cdot x}{N_k}\right). \tag{B.4}$$

Based on the pdf of $|h_{k,n_k}|^2$ given by (20), the pdf of $g(|h_{k,n_k}|^2)$ can be expressed as

$$f_{g(|h_{k,n_k}|^2)}(x) = \frac{K \cdot \log 2 \cdot 2^y \cdot e^{\frac{N_k \cdot (1-2^y)}{\xi}} \cdot \left(1 - e^{\frac{N_k \cdot (1-2^y)}{\xi}}\right)^{K-1}}{\frac{\xi}{N_k}}. \tag{B.5}$$

From Fig. B.1, it can be seen that the pdf of $g(|h_{k,n_k}|^2)$ is approximately symmetric centred at $g(\mu_{h^*})$. Thus, (B.3) can be approximated by second-

order Taylor's polynomial, which is

$$\bar{C}_{TOT} \approx K \cdot B \cdot \mathbb{E}_{N_k} \left\{ N_k \cdot \left[\log_2 \left(1 + \frac{\xi \cdot \mu_{h^*}}{N_k} \right) - \left(\frac{\xi^2}{2 \log(2) \cdot {N_k}^2 \cdot \left(1 + \frac{\xi \cdot \mu_{h^*}}{N_k} \right)^2} \right) \cdot \sigma_{h^*}^2 \right] \right\}. \tag{B.6}$$

Similarly, as the pmf of N_k for BSCR-based sub-cell resource allocation schemes given in (3.20) is a symmetric function, $\bar{C}_{TOT}$ can be approximated by the second-order Taylor's polynomial centred at $\bar{C}_{TOT}(\mathbb{E}[N_k])$. The average ergodic capacity for BSCR-based sub-cell resource allocation scheme can be shown as

$$\bar{C}_{BSCR} \approx K \cdot B \cdot \left(q(\mu_{N_k}) + \frac{q^{(2)}(\mu_{N_k}) \cdot \sigma_{N_k}^2}{2} \right), \tag{B.7}$$

where $\mu_{N_k} = \frac{N}{K}, \sigma_{N_k}^2 = \frac{N(K-1)}{K^2}$ and

$$q(x) = x \left[\log_2 \left(1 + \frac{\xi \mu_{h^*}}{x} \right) - \left(\frac{\xi^2}{2 \cdot x^2 \left(1 + \frac{\xi \cdot \mu_{h^*}}{x} \right)^2 \cdot \log(2)} \right) \sigma_{h^*}^2 \right]. \tag{B.8}$$

After some manipulation, the closed form of $\bar{C}_{BSCR}$ can be expressed as

$$\bar{C}_{BSCR} = B \cdot N \cdot \log_2 \left(1 + \frac{\xi \cdot K \cdot \mathbb{H}(K)}{N} \right) - \frac{B \cdot \pi^2 \cdot \xi^2}{12 \cdot \log(2) \cdot (\xi \cdot K \mathbb{H}(K) + N)^2} - B \cdot (K-1) \cdot \underbrace{\frac{b^4 + 2N \cdot (1-c) \cdot b^3 + N^2 \cdot (1+c) \cdot b^2}{(b+N)^4}}_{\Delta}, \tag{B.9}$$

where $b = \xi \cdot K \cdot \mathbb{H}(K)$ and $c = \frac{\pi^2}{6\mathbb{H}(K)^2}$. Since $b \geq 1$ and it is monotonically increasing with K, we have $\Delta \approx 1$. Thus, the last term of (B.9) can be approximated to $B \cdot (K-1)$ and (3.24) can be obtained.

For SAR resource allocation scheme, since $\tilde{N}_k$ follows Gamma distribution, the reciprocal random variable $a = \frac{1}{\tilde{N}_k}$ follows inverse gamma distribution with $\mu_a = \frac{K}{N-1}$ and $\sigma_a^2 = \frac{K^2}{(N-1)^2 \cdot (N-2)}$. $\bar{C}_{SAR}^u$ can be then approximated by the second-order Taylor's polynomial centred at $\bar{C}_{SAR}^u(\mathbb{E}[a])$, given by

$$\bar{C}_{SAR}^u \approx K \cdot B \cdot \left(p(\mu_a) + \frac{p^{(2)}(\mu_a) \cdot \sigma_a^2}{2} \right), \tag{B.10}$$

where

$$p(x) = \frac{1}{x} \left[\log_2 \left(1 + \xi \cdot \mu_{h^*} \cdot x\right) - \left(\frac{\xi^2 \cdot x^2}{2 \left(1 + \xi \cdot \mu_{h^*} \cdot x\right)^2 \log(2)} \right) \sigma_{h^*}^2 \right]. \tag{B.11}$$

After some manipulation, the closed form of $\bar{C}_{SAR}^u$ can then be derived as

$$\begin{aligned} \bar{C}_{SAR}^u = \frac{B \cdot (N-1)^2}{N-2} \cdot \log_2 \left(1 + \frac{\xi \cdot K \cdot \mathbb{H}(K)}{N-1}\right) - \frac{B \cdot \pi^2 \cdot (N-1) \cdot \xi^2 \cdot K^2}{12 \cdot \log(2) \cdot (\xi \cdot K \cdot \mathbb{H}(K) + N - 1)^2} \\ - \frac{B \cdot (N-1)}{2 \log(2)(N-2)} \cdot \frac{(c+3)d^4 + 2(4-c)d^3 + 7d^2 + 2d}{(d+1)^4}, \end{aligned} \tag{B.12}$$

where $d = \frac{\xi \cdot K \cdot \mathbb{H}(K)}{N-1}$. For the sake of simplicity, the last term of (B.12) can be neglected since it is much smaller than the first and second terms. Thus, (3.25) is obtained.

Appendix C

Derivation of (3.37) and (3.38)

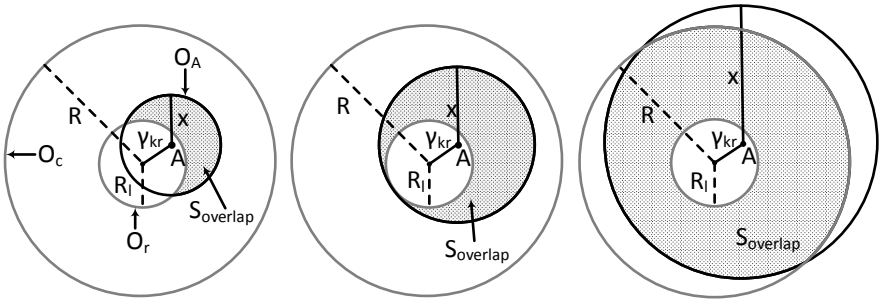
It is assumed that D2D pairs are distributed over a circular cell with radius R according to a homogeneous PPP Φ_A. Conditioned on the number of D2D pairs within each sub-cell, the D2D pairs are independently and uniformly distributed within each sub-cell. For $D2D_{k_r}$ within reference sub-cell l_r, the intra-cell interference is from the area between the boundary of central sub-cell circle O_r and the cell O_c. Therefore, the distance of interference link $d_{k_r,k_{l'}}$ has the following conditional cdf

$$F_{d_{k_r,k_{l'}}|\gamma_{k_r}}(x|y) = \frac{S_{overlap}}{\pi R^2 - \pi {R_l}^2}, \tag{C.1}$$

where $S_{overlap}$ is shown in Fig. C.1, which is the shaded intersection area of the circle O_A with center A, radius x and the annulus interference area.

As shown in the figure, $S_{overlap}$ is a discontinuous function regarding the location of $D2D_{k_r}$, $A_{k_r}(\gamma_{k_r}, \phi_{k_r})$. It can be easily observed that

$$S_{overlap} = \begin{cases} \pi x^2 - S_l & R_l - \gamma_{k_r} < x \leq R_l + \gamma_{k_r} \\ \pi x^2 - \pi R_l^2 & R_l + \gamma_{k_r} < x \leq R - \gamma_{k_r} \\ S_c - \pi R_l^2 & R - \gamma_{k_r} < x \leq R + \gamma_{k_r} \end{cases} \tag{C.2}$$



(a) $R_l-\gamma_{kr}< X \leq R_l+\gamma_{kr}$ (b) $R_l+\gamma_{kr}< X \leq R-\gamma_{kr}$ (c) $R-\gamma_{kr}< X \leq R+\gamma_{kr}$

Figure C.1: Graphic illustration of $S_{overlap}$.

where S_l and S_c are the intersection areas of circle O_A and circle O_r and the intersection area of circle O_A and circle O_c, respectively. According to [85], the intersection area of two circles, S_{inter} is given by

$$S_{inter}(r) = \pi x^2 \left(1 - \frac{1}{\pi} \arccos \frac{r^2 - x^2 - y^2}{2xy}\right) + r^2 \arccos \frac{r^2 + y^2 - x^2}{2ry}$$
$$-2\sqrt{\frac{r+x+y}{2}\left(\frac{r+x+y}{2} - r\right)\left(\frac{r+x+y}{2} - x\right)\left(\frac{r+x+y}{2} - y\right)}. \quad \text{(C.3)}$$

Therefore, we have $S_l = S_{inter}(R_l)$ and $S_c = S_{inter}(R)$. (3.37) can be then obtained by combining (C.1)-(C.3).

Appendix D

Proof of Theorem 2

As $X_1, \cdots, X_N$ are i.i.d., the expectation of r was derived in Appendix B of [86] as

$$\mathbb{E}[r] = N\left(1 - \int_1^c T(y) dF^{N-1}(y)\right), \tag{D.1}$$

where

$$T(y) = \frac{1}{y}\int_1^y \frac{1 - x^{-\beta}}{1 - c^{-\beta}}\, dx = \frac{1}{1 - c^{-\beta}} \cdot \left(1 - \frac{y^{-\beta} - \beta y^{-1}}{1 - \beta}\right). \tag{D.2}$$

By substituting (D.2), (3.45) into (D.1) and after some manipulation, we have

$$\begin{aligned}
\mathbb{E}[r] =& N\left[1 - \frac{1}{(1 - c^{-\beta})^N}\int_1^c \left(1 - \frac{y^{-\beta} - \beta y^{-1}}{1 - \beta}\right) d(1 - y^{-\beta})^{N-1}\right] \\
\overset{A=1-y^{-\beta}}{=}& N\left[1 - \frac{1}{((1 - c^{-\beta})^N)(1 - \beta)} \underbrace{\int_0^{1-c^{-\beta}} \left(A - \beta + \beta(1 - A)^{1/\beta}\right) dA^{N-1}}_{G(A)}\right],
\end{aligned} \tag{D.3}$$

where $G(A)$ can be further simplified by applying integration by parts, as

$$\begin{aligned} G(A) =&(1-c^{-\beta})^{N-1}\cdot(1-c^{-\beta}-\beta+\beta c^{-1}) - \int_0^{1-c^{-\beta}} A^{N-1}dA \\ &- \beta\int_0^{1-c^{-\beta}} A^{N-1}d(1-A)^{1/\beta} \\ =&(1-c^{-\beta})^{N-1}\cdot(1-c^{-\beta}-\beta+\beta c^{-1}) - \frac{(1-c^{-\beta})^N}{N} + B_{inc}\left(1-c^{-\beta}, N, \frac{1}{\beta}\right), \end{aligned} \tag{D.4}$$

where $B_{inc}(x,a,b) = \int_0^x t^{a-1}(1-t)^{b-1}dt$ is the incomplete Beta function. Finally, (3.46) can be obtained by combining (D.3) and (D.4).

Bibliography

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5g be?," *IEEE Journal on Selected Areas in Communications*, vol. 32, pp. 1065–1082, June 2014.

[2] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5g," *IEEE Communications Magazine*, vol. 52, pp. 74–80, February 2014.

[3] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5g cellular: It will work!," *IEEE Access*, vol. 1, pp. 335–349, 2013.

[4] R. Zhang, M. Wang, L. X. Cai, Z. Zheng, X. Shen, and L. L. Xie, "Lte-unlicensed: the future of spectrum aggregation for cellular networks," *IEEE Wireless Communications*, vol. 22, pp. 150–159, June 2015.

[5] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, Mar 1998.

[6] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive mimo for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, pp. 186–195, February 2014.

[7] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, "Femtocells: Past, present, and future," *IEEE Journal on Selected Areas in Communications*, vol. 30, pp. 497–508, April 2012.

[8] L. Dai, "An uplink capacity analysis of the distributed antenna system (das): From cellular das to das with virtual cells," *IEEE Transactions on Wireless Communications*, vol. 13, pp. 2717–2731, May 2014.

[9] K. Doppler, M. Rinne, C. Wijting, C. Ribeiro, and K. Hugl, "Device-to-Device communication as an underlay to lte-advanced networks," *Communications Magazine, IEEE*, vol. 47, pp. 42–49, Dec 2009.

[10] G. Fodor, E. Dahlman, G. Mildh, S. Parkvall, N. Reider, G. Miklos, and Z. Turanyi, "Design aspects of network assisted Device-to-Device communications," *IEEE Communications Magazine*, vol. 50, no. 3, pp. 170–177, 2012.

[11] X. Lin, J. G. Andrews, A. Ghosh, and R. Ratasuk, "An overview of 3gpp device-to-device proximity services," *IEEE Communications Magazine*, vol. 52, pp. 40–48, April 2014.

[12] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Communications Surveys Tutorials*, vol. 16, pp. 1801–1819, Fourthquarter 2014.

[13] E. Bastug, M. Bennis, and M. Debbah, "Living on the edge: The role of proactive caching in 5g wireless networks," *IEEE Communications Magazine*, vol. 52, pp. 82–89, Aug 2014.

[14] J. C. Haartsen, "The bluetooth radio system," *IEEE Personal Communications*, vol. 7, pp. 28–36, Feb 2000.

[15] J. S. Lee, Y. W. Su, and C. C. Shen, "A comparative study of wireless protocols: Bluetooth, uwb, zigbee, and wi-fi," in *IECON 2007 - 33rd*

Annual Conference of the IEEE Industrial Electronics Society, pp. 46–51, Nov 2007.

[16] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, pp. 388–404, Mar 2000.

[17] R. Wang, H. Yang, H. Wang, and D. Wu, "Social overlapping community-aware neighbor discovery for d2d communications," *IEEE Wireless Communications*, vol. 23, pp. 28–34, August 2016.

[18] K. W. Choi and Z. Han, "Device-to-device discovery for proximity-based service in lte-advanced system," *IEEE Journal on Selected Areas in Communications*, vol. 33, pp. 55–66, Jan 2015.

[19] X. Wu, S. Tavildar, S. Shakkottai, T. Richardson, J. Li, R. Laroia, and A. Jovicic, "Flashlinq: A synchronous distributed scheduler for peer-to-peer ad hoc networks," *IEEE/ACM Transactions on Networking*, vol. 21, pp. 1215–1228, Aug 2013.

[20] S. Balraj, "Lte direct overview," *Qualcomm Research, Tech*, 2013.

[21] Sophia-Antipolis, "Study on architecture enhancements to support proximity services (prose) (release 12)," *IEEE Communications Magazine*, vol. 3GPP TR 23.703 v. 1.0.0, 2013.

[22] B. Kaufman, J. Lilleberg, and B. Aazhang, "Spectrum sharing scheme between cellular users and ad-hoc device-to-device users," *IEEE Transactions on Wireless Communications*, vol. 12, pp. 1038–1049, March 2013.

[23] M. Usman, A. A. Gebremariam, U. Raza, and F. Granelli, "A software-defined device-to-device communication architecture for public safety applications in 5g networks," *IEEE Access*, vol. 3, pp. 1649–1654, 2015.

[24] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Stochastic geometry study on device-to-device communication as a disaster relief solution," *IEEE Transactions on Vehicular Technology*, vol. 65, pp. 3005–3017, May 2016.

[25] D. Xenakis, M. Kountouris, L. Merakos, N. Passas, and C. Verikoukis, "Performance analysis of network-assisted d2d discovery in random spatial networks," *IEEE Transactions on Wireless Communications*, vol. 15, pp. 5695–5707, Aug 2016.

[26] A. Prasad, A. Kunz, G. Velev, K. Samdanis, and J. Song, "Energy-efficient d2d discovery for proximity services in 3gpp lte-advanced networks: Prose discovery mechanisms," *IEEE Vehicular Technology Magazine*, vol. 9, pp. 40–50, Dec 2014.

[27] H. Wang and X. Chu, "Distance-constrained resource-sharing criteria for device-to-device communications underlaying cellular networks," *Electronics Letters*, vol. 48, pp. 528–530, April 2012.

[28] D. Feng, L. Lu, Y. Yuan-Wu, G. Y. Li, G. Feng, and S. Li, "Device-to-device communications underlaying cellular networks," *IEEE Transactions on Communications*, vol. 61, pp. 3541–3551, August 2013.

[29] C.-H. Yu, K. Doppler, C. Ribeiro, and O. Tirkkonen, "Resource sharing optimization for Device-to-Device communication underlaying cellular networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2752–2763, 2011.

[30] M. Lin, J. Ouyang, and W. P. Zhu, "Joint beamforming and power control for device-to-device communications underlaying cellular networks,"

IEEE Journal on Selected Areas in Communications, vol. 34, pp. 138–150, Jan 2016.

[31] G. A. Safdar, M. Ur-Rehman, M. Muhammad, M. A. Imran, and R. Tafazolli, "Interference mitigation in d2d communication underlaying lte-a network," *IEEE Access*, vol. 4, pp. 7967–7987, 2016.

[32] Y. Zhang, C. Zhao, L. Hong, J. Yu, and S. Li, "Interference alignment with random vector quantisation in device-to-device underlaying cellular networks," *IET Communications*, vol. 10, no. 10, pp. 1255–1262, 2016.

[33] Y. Ni, S. Jin, W. Xu, M. Matthaiou, S. Shao, and H. Zhu, "Beamforming and interference cancellation schemes for d2d communications," in *2015 IEEE International Conference on Communication Workshop (ICCW)*, pp. 608–613, June 2015.

[34] S. L. Chiu, K. C. J. Lin, G. X. Lin, and H. Y. Wei, "Empowering device-to-device networks with cross-link interference management," *IEEE Transactions on Mobile Computing*, vol. 16, pp. 950–963, April 2017.

[35] Y. Cao, T. Jiang, and C. Wang, "Cooperative device-to-device communications in cellular networks," *IEEE Wireless Communications*, vol. 22, pp. 124–129, June 2015.

[36] Y. Wu, S. Wang, W. Liu, W. Guo, and X. Chu, "Iunius: A cross-layer peer-to-peer system with device-to-device communications," *IEEE Transactions on Wireless Communications*, vol. 15, pp. 7005–7017, Oct 2016.

[37] J. Zhao, Y. Liu, K. K. Chai, Y. Chen, and M. Elkashlan, "Joint subchannel and power allocation for noma enhanced d2d communications," *IEEE Transactions on Communications*, vol. 65, pp. 5081–5094, Nov 2017.

[38] B. Ozbek, M. Pischella, and D. L. Ruyet, "Dynamic shared spectrum allocation for underlaying device-to-device communications," *IEEE Wireless Communications*, vol. 24, pp. 88–93, October 2017.

[39] L. Liang, G. Y. Li, and W. Xu, "Resource allocation for d2d-enabled vehicular communications," *IEEE Transactions on Communications*, vol. 65, pp. 3186–3197, July 2017.

[40] Y. Qian, T. Zhang, and D. He, "Resource allocation for multichannel device-to-device communications underlaying qos-protected cellular networks," *IET Communications*, vol. 11, no. 4, pp. 558–565, 2017.

[41] M. Naslcheraghi, S. A. Ghorashi, and M. Shikh-Bahaei, "Full-duplex device-to-device collaboration for low-latency wireless video distribution," in *2017 24th International Conference on Telecommunications (ICT)*, pp. 1–5, May 2017.

[42] P. Wu, P. C. Cosman, and L. B. Milstein, "Resource allocation for multicarrier device-to-device video transmission: Symbol error rate analysis and algorithm design," *IEEE Transactions on Communications*, vol. 65, pp. 4446–4462, Oct 2017.

[43] X. Wang, M. Chen, T. Taleb, A. Ksentini, and V. C. M. Leung, "Cache in the air: exploiting content caching and delivery techniques for 5g systems," *IEEE Communications Magazine*, vol. 52, pp. 131–139, February 2014.

[44] M. Ji, G. Caire, and A. F. Molisch, "Wireless device-to-device caching networks: Basic principles and system performance," *IEEE Journal on Selected Areas in Communications*, vol. 34, pp. 176–189, Jan 2016.

[45] D. Malak, M. Shalash, and J. Andrews, "Spatially correlated content caching for device-to-device communications," *IEEE Transactions on Wireless Communications*, vol. PP, no. 99, pp. 1–1, 2017.

[46] D. Malak, M. Al-Shalash, and J. G. Andrews, "Optimizing content caching to maximize the density of successful receptions in device-to-device networking," *IEEE Transactions on Communications*, vol. 64, pp. 4365–4380, Oct 2016.

[47] M. Ji, G. Caire, and A. F. Molisch, "The throughput-outage tradeoff of wireless one-hop caching networks," *IEEE Transactions on Information Theory*, vol. 61, pp. 6833–6859, Dec 2015.

[48] L. Wang, H. Wu, Y. Ding, W. Chen, and H. V. Poor, "Hypergraph-based wireless distributed storage optimization for cellular d2d underlays," *IEEE Journal on Selected Areas in Communications*, vol. 34, pp. 2650–2666, Oct 2016.

[49] N. Golrezaei, P. Mansourifard, A. F. Molisch, and A. G. Dimakis, "Base-station assisted device-to-device communications for high-throughput wireless video networks," *IEEE Transactions on Wireless Communications*, vol. 13, pp. 3665–3676, July 2014.

[50] M. Naslcheraghi, S. A. Ghorashi, and M. Shikh-Bahaei, "Fd device-to-device communication for wireless video distribution," *IET Communications*, vol. 11, no. 7, pp. 1074–1081, 2017.

[51] M. Gregori, J. Gmez-Vilardeb, J. Matamoros, and D. Gndz, "Wireless content caching for small cell and d2d networks," *IEEE Journal on Selected Areas in Communications*, vol. 34, pp. 1222–1234, May 2016.

[52] N. Giatsoglou, K. Ntontin, E. Kartsakli, A. Antonopoulos, and C. Verikoukis, "D2d-aware device caching in mmwave-cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, pp. 2025–2037, Sept 2017.

[53] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution toward 5g multi-tier cellular wireless networks: An interference management perspective," *IEEE Wireless Communications*, vol. 21, pp. 118–127, June 2014.

[54] C. Y. Wong, R. Cheng, K. Lataief, and R. Murch, "Multiuser ofdm with adaptive subcarrier, bit, and power allocation," *Selected Areas in Communications, IEEE Journal on*, vol. 17, pp. 1747–1758, Oct 1999.

[55] H.-Y. Hsieh and R. Sivakumar, "On using peer-to-peer communication in cellular wireless data networks," *IEEE Transactions on Mobile Computing*, vol. 3, pp. 57–72, Jan 2004.

[56] H. Zhu and J. Wang, "Chunk-based resource allocation in OFDMA systems - part i: chunk allocation," *IEEE Transactions on Communications*, vol. 57, pp. 2734–2744, September 2009.

[57] H. Zhu and J. Wang, "Chunk-based resource allocation in OFDMA Systems-Part II: joint chunk, power and bit allocation," *IEEE Transactions on Communications*, vol. 60, pp. 499–509, February 2012.

[58] H. Song, J. Y. Ryu, W. Choi, and R. Schober, "Joint power and rate control for device-to-device communications in cellular systems," *IEEE Transactions on Wireless Communications*, vol. 14, pp. 5750–5762, Oct 2015.

[59] Z. Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, pp. 57–73, Feb 2008.

[60] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Transactions on Communications*, vol. 54, pp. 1310–1322, July 2006.

[61] R. Tang, J. Zhao, H. Qu, and Z. Zhang, "User-centric joint admission control and resource allocation for 5g d2d extreme mobile broadband: A sequential convex programming approach," *IEEE Communications Letters*, vol. 21, pp. 1641–1644, July 2017.

[62] H. Tang and Z. Ding, "Mixed mode transmission and resource allocation for d2d communication," *IEEE Transactions on Wireless Communications*, vol. 15, pp. 162–175, Jan 2016.

[63] R. Zhang, X. Cheng, L. Yang, and B. Jiao, "Interference graph-based resource allocation (ingra) for d2d communications underlaying cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 64, pp. 3844–3850, Aug 2015.

[64] H. Zhang, L. Song, and Z. Han, "Radio resource allocation for device-to-device underlay communication using hypergraph theory," *IEEE Transactions on Wireless Communications*, vol. 15, pp. 4852–4861, July 2016.

[65] Q. Ye, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "Distributed resource allocation in device-to-device enhanced cellular networks," *IEEE Transactions on Communications*, vol. 63, pp. 441–454, Feb 2015.

[66] R. Yin, C. Zhong, G. Yu, Z. Zhang, K. K. Wong, and X. Chen, "Joint spectrum and power allocation for d2d communications underlaying cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 65, pp. 2182–2195, April 2016.

[67] R. Yin, G. Yu, H. Zhang, Z. Zhang, and G. Y. Li, "Pricing-based interference coordination for d2d communications in cellular networks," *IEEE Transactions on Wireless Communications*, vol. 14, pp. 1519–1532, March 2015.

[68] D. Burghal and A. F. Molisch, "Efficient channel state information acquisition for device-to-device networks," *IEEE Transactions on Wireless Communications*, vol. 15, pp. 965–979, Feb 2016.

[69] N. Lee, X. Lin, J. G. Andrews, and R. W. Heath, "Power control for d2d underlaid cellular networks: Modeling, algorithms, and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 33, pp. 1–13, Jan 2015.

[70] H. Min, J. Lee, S. Park, and D. Hong, "Capacity Enhancement Using an Interference Limited Area for Device-to-Device Uplink Underlaying Cellular Networks," *IEEE Transactions on Wireless Communications*, vol. 10, pp. 3995–4000, December 2011.

[71] K. Kumar and Y.-H. Lu, "Cloud computing for mobile users: Can offloading computation save energy?," *Computer*, vol. 43, pp. 51–56, April 2010.

[72] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," *IEEE Communications Surveys Tutorials*, vol. PP, no. 99, pp. 1–1, 2017.

[73] W. Zhang, Y. Wen, K. Guan, D. Kilper, H. Luo, and D. Wu, "Energy-optimal mobile cloud computing under stochastic wireless channel," *IEEE Transactions on Wireless Communications*, vol. 12, pp. 4569–4581, September 2013.

[74] C. You, K. Huang, H. Chae, and B. H. Kim, "Energy-efficient resource allocation for mobile-edge computation offloading," *IEEE Transactions on Wireless Communications*, vol. 16, pp. 1397–1411, March 2017.

[75] X. Chen, L. Jiao, W. Li, and X. Fu, "Efficient multi-user computation offloading for mobile-edge cloud computing," *IEEE/ACM Transactions on Networking*, vol. 24, pp. 2795–2808, October 2016.

[76] L. Lei, Z. Zhong, C. Lin, and X. Shen, "Operator controlled Device-to-Device communications in LTE-advanced networks," *IEEE Wireless Communications*, vol. 19, no. 3, pp. 96–104, 2012.

[77] J. Wang and L. Milstein, "Cdma overlay situations for microcellular mobile communications," *Communications, IEEE Transactions on*, vol. 43, pp. 603–614, Feb 1995.

[78] J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE Journal on Selected Areas in Communications*, vol. 21, pp. 171–178, Feb 2003.

[79] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *Information Theory, IEEE Transactions on*, vol. 44, pp. 2619–2692, Oct 1998.

[80] W. Yu and J. M. Cioffi, "Constant-power waterfilling: performance bound and low-complexity implementation," *IEEE Transactions on Communications*, vol. 54, pp. 23–28, Jan 2006.

[81] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. Communications and signal processing, McGraw-Hill, 1991.

[82] G. George, R. K. Mungara, A. Lozano, and M. Haenggi, "Ergodic spectral efficiency in mimo cellular networks," *IEEE Transactions on Wireless Communications*, vol. 16, pp. 2835–2849, May 2017.

[83] L. Yang, J. Cao, S. Tang, T. Li, and A. Chan, "A framework for partitioning and execution of data stream applications in mobile cloud computing," in *2012 IEEE 5th International Conference on Cloud Computing (CLOUD)*, pp. 794–802, June 2012.

[84] K. Wang, K. Yang, and C. Magurawalage, "Joint energy minimization and resource allocation in C-RAN with mobile cloud," *IEEE Transactions on Cloud Computing*, vol. PP, no. 99, pp. 1–1, 2016.

[85] L. Dai, "A comparative study on uplink sum capacity with co-located and distributed antennas," *Selected Areas in Communications, IEEE Journal on*, vol. 29, pp. 1200–1213, June 2011.

[86] Y. Y. Zaliapin, I. V.and Kagan and F. P. Schoenberg, "Approximating the distribution of pareto sums," *pure and applied geophysics*, vol. 162, no. 6, pp. 1187–1228, 2005.

List of Publications

Journal Papers

1. Yuan Kai, Huiling Zhu and Jiangzhou Wang, "Resource Allocation and Performance Analysis of Cellular-assisted OFDMA Device-to-Device Communication," submitted to *IEEE Trans. Wireless Commun.*

Conference Papers

1. Yuan Kai and Huiling Zhu, "Resource allocation for multiple-pair D2D communications in cellular networks," in *Proc. IEEE ICC*, pp. 2955-2960, Jun. 2015.

2. Yuan Kai, Junyuan Wang and Jiangzhou Wang, "Energy Minimization for D2D assisted Mobile Edge Computing Networks," submitted to *Proc. IEEE ICC*, Jun. 2018.