Stochastic Geometric Analysis of Cellular Networks Enhanced With D2D and M2M Communication

UNIVERSITY OF LEEDS

Asma Afzal

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

School of Electronic and Electrical Engineering

University of Leeds

September 2017

Declaration

The candidate confirms that the work submitted is his/her own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that the work submitted is her own and that appropriate credit has been given within the thesis where reference has been made to the work of others.
The work in Chapter 2 has appeared in the following publication.

(a) Asma Afzal, Afef Feki, Merouane Debbah, Syed Ali Raza Zaidi, Mounir Ghogho and Desmond C. McLernon, "Leveraging D2D to Maximize the Spectral Efficiency of Massive MIMO Systems", published in Workshop on Spatial Stochastic Models for Wireless Networks (SpaSWiN), 2017.

The work in chapter 3 has appreared in the following publication.

(b) Asma Afzal, Syed Ali Raza Zaidi, Desmond C. McLernon and Mounir Ghogho "On the Analysis of Device-to-Device Overlaid Cellular Networks in the Uplink under 3GPP Propagation Model", published in IEEE Wireless Communications and Networking Conference (WCNC), 2016.

The work in chapter 4 is based on the following publications.

(c) Asma Afzal, Syed Ali Raza Zaidi, Desmond C. McLernon and Mounir Ghogho, "On the Analysis of Cellular Networks with Caching and Coordinated Device-to-Device Communication", published in IEEE International Conference on Communications (ICC) 2016.

(d) Asma Afzal, Syed Ali Raza Zaidi, Desmond C. McLernon and Mounir Ghogho, "Information-centric offloading in cellular networks with coordinated Device-to-Device Communication", submitted to IEEE Transactions on Wireless Communications.

The work in chapter 5 is based on the following publications.

(e) Asma Afzal, Syed Ali Raza Zaidi, Desmond C. McLernon, Mounir Ghogho and Afef Feki, "M2M meets D2D: Harnessing D2D Interfaces for the Aggregation of M2M Data", published in IEEE International Conference on Communications (ICC) 2017.

(f) Asma Afzal, Syed Ali Raza Zaidi, Afef Feki, Desmond C. McLernon and Mounir Ghogho, "A Poisson Hard Sphere Model for the Analysis of M2M Data Aggregation via D2D Links", submitted to IEEE Transactions on Communications.

The work in chapter 6 has appeared in the following publications.

(g) Asma Afzal, Syed Ali Raza Zaidi, Muhammad Zeeshan Shakir, Muhammad Ali Imran, Mounir Ghogho, Athanasios V. Vasilakos, Desmond C. McLernon, and Khalid Qaraqe, "The cognitive Internet of Things: a unified perspective", published in ACM Springer Mobile Networks and Applications (MONET) journal, 2015: pg. 72-85.

(h) Syed Ali Raza Zaidi, Asma Afzal, Maryam Hafeez, Mounir Ghogho, Desmond C McLernon, Ananthram Swami, "Solar energy empowered 5G cognitive metro-cellular networks", published in IEEE Communications Magazine, July 2015.

The work in all the publications listed above was carried out with the advice and assistance of Dr Syed Ali Zaidi, Dr Des McLernon, and Professor Mounir Ghogho. Dr Afef Feki and Dr Merouane Debbah mentored the candidate during her research visit to France and provided their useful feedback on the results. Dr Muhammad Zeeshan Shakir, Dr Muhammad Ali Imran, Dr Athanasios V. Vasilakos and Dr Khalid Qaraqe helped crystallize the idea for the joint publication (g). This is to assert that the candidate has contributed solely to the technical part and writing of the papers (a)-(g) under the guidance of her co-authors. The candidate contributed to the results and simulations for (h) and wrote some sections along with Dr Syed Ali Zaidi and Dr Maryam Hafeez.

To my parents, Sheikh Muhammad Afzal & Ulfat Parveen

Acknowledgments

First and foremost I would like to thank Allah Almighty whose benevolence and endless mercy made all my achievements possible.

I would like to acknowledge the technical guidance from Dr Des McLernon, Dr Syed Ali Zaidi and Professor Mounir Ghogho throughout my study. I am particularly grateful to Dr Ali for his expert advice and feedback in the discipline of stochastic geometry, which was the major underpinning theoretical component of this thesis. I would also like to thank Dr Des for his kind and caring attitude. I could reach out to him with the pettiest of issues and he would lend his ear and give his helpful advice.

I acknowledge the generous financial assistance of the Commonwealth Scholarship Commission in the UK without which, none of this work would have been possible.

I would like to extend my gratitude to Professor Mérouane Debbah and Dr Afef Feki for our fruitful discussions during my research visit to CentraleSupélec in France. Dr. Afef's friendly demeanor, openness and positive attitude made my stay in France enjoyable and a memorable one. I am very thankful for her eagerness to help me in every way.

I am grateful to my former and present colleagues including Sana Hamid, Dr Samaneh Igder, Dr Raul Aquino, Dr Edmond Nurellari, Dr Daniel Licea, Yen Nguyen, Ali Hayajneh and Tuan Tang for a friendly research environment and enlightening discussions on many different topics.

My sincere appreciation goes to my family and my in-laws, whose encouragement and support has always been a strong motivation for me. I am sure my late mother would be very proud of me, smiling in the heavens. Lastly, but most importantly, words fall short to show how grateful I am to my husband Zubair Lutfullah Kakakhel. His love, constant reassurance and belief in me kept me sane and helped me persevere. I am truly grateful for his companionship.

Abstract

Cellular networks have to undergo a complete transformation to meet the formidable capacity demands from the ever-increasing number of smart devices. A close look at the requested traffic profile of the devices reveals a deeper underlying challenge for 5G wireless networks. This is because, there is no "one-size fits all" solution and different categories of devices have contrasting requirements. For example, smart phones expect anytime, anywhere connectivity and because of the data hungry applications, they also expect a certain quality of experience. On the other hand, the massive amount of other smart things deployed in the near future will primarily require to stay connected with only a small payload to transmit/receive. Thus, the network operators have an arduous task ahead to design ultra-flexible networks that can easily accommodate a large number of devices with unique requirements and specifications.

Device-to-device (D2D) communication has been recently proposed as a promising solution to enhance the capacity of cellular networks by enabling direct communication between user equipments (UEs) located in close proximity without the intervention of the base station (BS). In this thesis, we borrow tools from stochastic geometry to analyze the gains in throughput achieved by offloading UEs to communicate via D2D. We study distance-based and and content-based mode selection strategies. In the distance based mode selection, it is assumed that the D2D pair already exists and D2D mode is selected if the distance separation between them is below a certain threshold. On the other hand, in content based mode selection, the D2D pairs are created subject to requested content availability.

We also study how the concept of D2D communication could be extended to establish D2D connections between UEs and the machine-type-devices (MTDs) located in their close proximity to aggregate M2M data. This could potentially reduce the burden of massive access of MTDs on the BS. We employ a novel Poisson hard sphere model for the association between UEs and the MTDs. We quantify the number of MTDs from which a UE can successfully aggregate data

without compromising its quality of service (QoS) requirements.
Finally, another contribution of this dissertation is comprehensive statistical modeling of the coverage of M2M networks operating in the same spectrum as the cellular networks. The MTDs employ cognition to satisfy the strict QoS constraints of the primary cellular networks. We consider that the MTDs are energy limited as they harvest energy from the sun. The transmit power of the MTDs is therefore shown to play a critical role to efficiently utilize the harvested energy and also maximize the spectrum access opportunities.

Contents

List of Acronyms and Symbols

Acronyms

3GPP	3rd generation partnership project
5G	fifth generation
AWGN	Additive white Gaussian noise
BS	Base station
CR	Cognitive radio
D2D	Device-to-Device
DL	Downlink
EPC	Evolved packet core
FS	Free space
HPPP	Homogeneous Poisson point process
i.i.d	independent identically distributed
ICN	Information-centric networking
IoT	Internet of Things
LoS	Line of sight
M2M	Machine-to-machine
MIMO	Multiple input multiple output
NLoS	Non line of sight
PHS	Poisson hard sphere

QoS	Quality of Service
RAT	Radio access technology
SE	Spectral Efficiency
SIC	Successive interference cancellation
SINR	Signal to interference and noise ratio
UE	User equipment
UL	Uplink

Symbols

$\mathbb{P}[A]$	Probability of event A
α	Path loss exponent*
ϵ	Fraction of UEs operating in D2D mode
η	Ratio of densities*
λ	Density of point process*
$\mathbb{E}_X[.]$	Expectation with respect to random variable X
$\mathcal{L}_I(.)$	Laplace transform of aggregate interference
$\mathcal{P}$	Overall success probability*
$\mathcal{R}$	Probability of rate coverage*
$\mathcal{S}$	Probability of SINR coverage*
$\mathcal{T}$	Average throughput*
Φ	Point process
σ^2	Noise power*
τ	Fraction of time reserved for Aggregating M2M data*
θ	SINR threshold*

c	Requested content index
E	Irradiance*
h	Channel power*
L	Size of library
M	Number of antennas at the base station
N	Number of user equipments
P	Transmit power*
r	Desired link distance separation*
W	Bandwidth reserved for communication*

*These symbols appear with superscripts and subscripts in the thesis for further classification.

List of Figures

1 Introduction

In this chapter, the background, motivation and the organization of the research work carried out is presented. More specifically, we identify the challenges involved in integrating D2D and M2M communications with the legacy cellular networks and comprehensively review the existing literature to address these challenges. Finally, we highlight the contributions of this thesis.

1.1. Motivation

The past decade has witnessed a tremendous growth in data transfers over the wireless medium. This growth is a result of proliferation in ubiquitous devices such as smart phones, tablets and other hand-held devices. Social networking, e-gaming, navigation, multimedia content downloads and sharing including ultra high-definition/4K videos, and other real-time data hungry applications are now a part of our lives and have become a necessity rather than a luxury. To add to this, the current vision of urban development is to empower everyday objects with internet connectivity creating what we call the internet of things (IoTs), leading to *smart cities*. The connectivity of these objects is pivotal as simple object level interactions can collectively result in powerful and intelligent decision making. As a consequence, the internet service providers face an onerous task of not only ensuring anytime/anywhere connectivity, but also maintaining users' satisfaction and quality of experience (QoE).

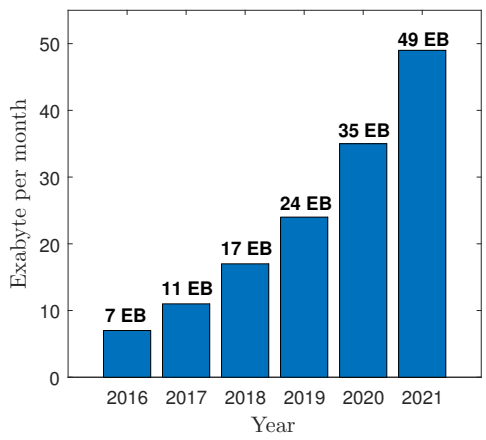
In the context of cellular networks, there has been a continuous evolution in cellular capabilities over the years. The cellular technology was initially designed to support voice calls only and the first digital cellular technology in 1993, commonly known as 2G, offered extremely low data rates of only a few kbps with a severely limited range. The growing trends towards video calling and web browsing soon resulted

in a major paradigm shift and cellular networks continued to push the barriers with technologies like code division multiple access (CDMA), orthogonal frequency division multiple access (OFDMA), multi-user multiple-input-multiple-output (MIMO), etc., supporting higher data rates and simultaneous user connections. The present day cellular technology of 4G, LTE-Advanced, supports data rates of up to 100Mbps with a range of over tens of miles.

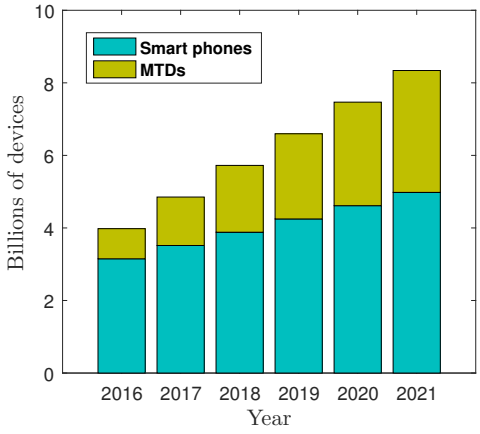
This technological advancement in cellular networks may at first seem impressive, but it falls short of meeting the explosive data demand and the number of simultaneously active devices. According to a recent mobile data forecast, the global mobile data traffic grew 18 fold over the past 5 years and will continue to grow in the future [1] and as aptly stated by Cisco, "*with no signs of slowing, data just keeps growing*" [2]. In 2016 alone, the global mobile data traffic grew 63% with a record total of 7.2 exabytes per month, where an exabyte is equivalent to a billion gigabytes. The annual data traffic is predicted to hit a staggering 49 exabytes at the end of 2021 as shown in Figure 1.1a. Smart phones are a major contributor to this growth, making up nearly 86% of this traffic by 2021. However, it may come as a surprise that the number of smart phones will remain fairly stagnant with only roughly 5% increase over the span of five years, while the number of machine-type-devices (MTDs) will steadily increase with the expected compound annual growth (CAGR) of 34% by the end of 2021 as displayed in Figure 1.1b.

MTDs are a key enabler of the Internet of Things (IoTs) as machine-to-machine (M2M) communication is finding its use in a wide variety of applications including health care, vehicle to everything (V2X) communication, industrial automation, banking, smart grid utilities, etc. [3]. The proliferation in MTDs has been fueled by the miniaturization of electronics with high computational capabilities. In particular, the industry of smart wearable devices is gaining traction and contributing significantly to the rise of MTDs. Wearable devices are equipped with the capability to connect and communicate to the network either directly through embedded cellular connectivity or through another device (typically a smartphone) using Wi-Fi, Bluetooth, or another technology. These devices come in various shapes and forms, ranging from smart watches, fitness trackers, health monitors, smart glasses, heads-up displays (HUDs), wearable navigation devices and scanners, smart clothing, etc.

Because of the staggering statistics of data traffic and connected devices, the wireless community is already working towards the fifth generation (5G) cellular



(a) Projected global mobile data traffic.



(b) Projected number of MTDs and smart phones.

Figure 1.1.: Global mobile data trends and forecast [1].

system to support the requirements of the projected challenging use cases. 5G networks are expected to provide 1000× capacity and support 100× simultaneous connections and data rates with extremely low latency [4]. A straightforward way to improve network capacity is to increase the available spectrum, but strict regulations for spectrum licensing and excessive costs make it prohibitive. Many new enhancements to the current LTE infrastructure have been proposed by the 3rd generation partnership project (3GPP) to materialize the concept of 5G; each with some benefits and pitfalls. One such proposed technique is the proximity service (ProSe), which was first introduced in Release 12 of the 3GPP specifications [5]. Commonly referred to as direct device-to-device (D2D) communication, it is a technology that allows LTE devices to communicate directly without the intervention of the base station (BS). Since traditional communication between UEs takes place in two hops, i.e. to and from the BS, the initial intent of ProSe was to enable autonomous communication between devices in case of network failure. However, the potential of D2D communication soon became evident and its applications extended to social networking, pushing targeted advertisements based on user's location and interests and cooperative communication, to name but a few.

1.2. D2D Communication

What makes D2D communication so attractive for 5G networks is the innate fact that users in close proximity share common interests and are more inclined to request a certain content at the same time [6]. The internet infrastructure is continuously evolving from the traditional host-centric networking architecture to a more versatile, content-aware information-centric networking (ICN) architecture. With the help of ICN, intelligent ways of data dissemination and storage can be developed [7]. Creating caching zones by storing popular content at the user equipments (UEs) will allow neighboring UEs to exchange content directly and offload the burden of multiple duplicate requests at the core network. The short range direct communication naturally improves coverage, which in turn improves the data rate experienced by the users and also reduces latency and required transmit power. D2D communication can also help extend coverage of the cell or improve the coverage of the cell edge users.

In many aspects, the concept of D2D communication is quite similar to communication in M2M and ad hoc networks with a few fundamental differences. In the case of M2M communication, there is generally a missing human element and the transmissions are assumed to take place between machines [8]. M2M communication is technology-independent and application oriented and there is no explicit range restriction for communicating devices, while D2D relies on connectivity based on proximity between devices. The presence of a centralized control offered by the cellular network is what sets a D2D network apart from the ad hoc networks such as mobile ad hoc networks (MANET) or cognitive radio networks (CRNs), which have been thoroughly explored in the past [9–11]. The lack of coordination between the infrastructure and CRNs causes a lot of energy and time to be wasted on spectrum sensing and interference management becomes critical. Similarly in MANETs, collision avoidance, synchronization and multi-hop routing are major issues. D2D communication assumes localized single-hop transmissions, thus relieving the need for complex multi-hop routing protocols. The transparency of a D2D network also enables a cellular network to authorize, authenticate and synchronize a D2D connection. The network can also optimize resources, guarantee a certain quality-of-service (QoS) and even arrange billing for provisioning D2D services.

1.2.1. D2D Design Conundrums

Currently, there is yet to be a unanimous agreement on how D2D communication is going to be integrated with existing cellular networks [8]. The challenges faced on several fronts are explained as follows.

- *Device Discovery*: The two underlying components of D2D specified by [5] are device discovery and D2D communication. An accurate device discovery procedure is imperative for discovering nearby devices and establishing D2D links. The device discovery mechanisms can be further broken down into direct discovery and evolved packet core (EPC)-level discovery. Direct discovery is based on UE expressions; the UEs may either make announcements to notify other UEs of their identity and availability or listen for such announcements for suitable D2D transmitters. LTE Direct is one such proposed D2D schemes using application layer direct device discovery [12]. The EPC-level discovery, on the other hand, is based on proximity calculation based on the location information present in the mobility management entity (MME) in the core network. Direct discovery incurs the overhead of repeated signaling and battery drainage in the UEs, but can be more accurate due to periodic updates by the UEs [13].

- *Mode Selection*: Traditional cellular communication between UE-A and UE-B takes place when UE-A transmits to the BS in the cellular uplink (UL) slot and the BS forwards this message to UE-B in the downlink (DL). When UE-A and UE-B are physically close together, they can switch from cellular mode to D2D mode. Mode switching not only depends on the proximity between the UEs, but also on the policies set by the network operator. As a general rule, the UEs operating in D2D mode should be able to peacefully coexist within the cellular network. UEs can be allowed to switch to D2D mode to improve individual device utilities such as improving link spectral efficiency or minimizing energy consumption [14], but it is more important from the network's perspective that the interference generated from D2D links is kept at bay and the overall network level utilities are optimized. For example, in [15] D2D links are activated only when the constraint of ensuring a certain QoS for cellular UEs is satisfied [15].

- *Spectrum Access*: The classification of spectrum access techniques for D2D

communication proposed in literature is illustrated in Figure 1.2 and explained as follows.

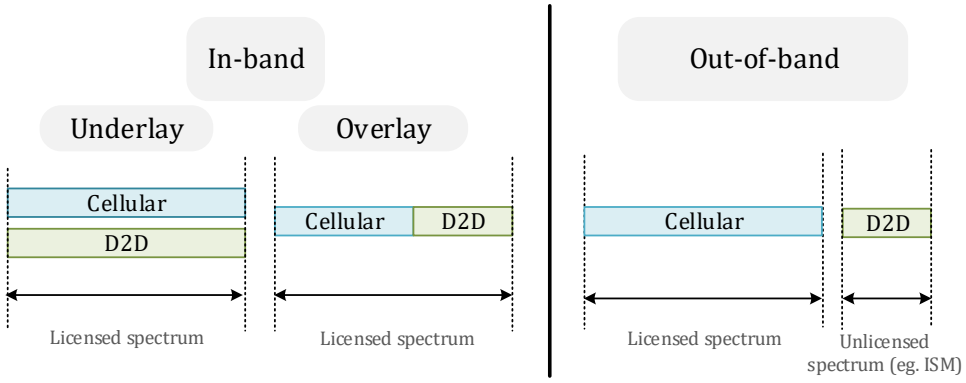


Figure 1.2.: Spectrum access for D2D communication.

- In-band D2D: It is when the D2D UEs access the licensed spectrum for communication. It is further subdivided into underlay and overlay. Underlay D2D is when both the cellular and D2D UEs share the spectrum resources, while overlay D2D is when a portion of the spectrum resources is reserved for D2D communication only. A large volume of research focuses on underlay D2D as it helps efficiently utilize the spectrum and enhances the overall spectral efficiency [14, 16–24]. However, a major concern is mitigating interference generated from D2D UEs [8, 20, 25, 26]. The problem of interference is eliminated in the case of overlay D2D, but dynamic resource management is required to efficiently utilize the available spectrum [27]. In this thesis, we mainly consider in-band overlay D2D communication except for one chapter, where we consider underlay communication.
- Out-of-band D2D: It is when the D2D UEs operate in the unlicensed spectrum such as 2.4 GHz ISM band. Examples include Zigbee, Bluetooth and WiFi Direct. A clear advantage of this approach is that it makes D2D and cellular communications impossible to interfere with each other [28]. However, all the services on unlicensed bands are best effort services and do not guarantee a QoS [8]. It is also extremely challenging for the network to coordinate out-of-band access and manage mode switching.

- *Infrastructure control*: It is necessary to ascertain what level of control the network should provide for successful coexistence of D2D communication. The availability of multiple radio access technologies (RATs) at the UEs further makes coordinating access a challenging task. We briefly discuss the possible configurations in which D2D can be implemented.
  - Cellular controlled D2D, also commonly known as network-assisted D2D, is when the cellular network is fully responsible for establishing, managing and arbitrating a D2D connection. D2D communication takes place in the licensed spectrum, which ensures security and reliability of a D2D link and improves the overall capacity of the network [19].
  - Cellular aware D2D, is when the cellular network is aware of D2D but not fully responsible to maintain a successful D2D connection. Cellular aware D2D takes place on a different RAT, for example, a D2D pair can be offloaded to operate on WiFi, but the cellular network may not be able to provide any QoS or security guarantees.
  - Autonomous D2D, is when the cellular network is completely oblivious to the D2D communication. Cellular unaware D2D finds its use in social networking, peer-to-peer communication and advertising. On a large scale, autonomous D2D can behave exactly like ad hoc networks in emergency situations, where the infrastructure is compromised. Although autonomous D2D communication may usually take place in the unlicensed band, D2D pairs may also access the licensed spectrum following the principles of CRNs.

- *Cellular link*: Another challenge particularly relevant to network-assisted in-band underlay D2D is whether D2D communication should take place in the UL or DL slot. In case of UL, the victim of the interference generated by the D2D pairs are the BSs, while in DL it is the cellular UEs. Conversely, the D2D receivers experience interference caused by cellular UEs in the UL and the BSs in DL. There is no clear suitable choice between UL or DL for D2D, but it is reported by [16] that the interference generated by the UEs in the UL is much lower than that from the BSs in DL and it better to take advantage of D2D in the UL. A recent study in [26] however, points out that the interference power generated by the BSs in DL may be powerful, but we can exploit the fact that the interference in temporally correlated due to static BSs. This gives us

better performance of D2D in DL.

- *Architecture enhancements*: To facilitate seamless integration of D2D communication with LTE networks, it is necessary to modify the EPC network by introducing new functional entities and develop standardized protocols. The inclusion of a D2D server is proposed in [29], where the D2D server is responsible for storing and updating identifiers of the D2D enabled UEs attached to the network, authorization of D2D UEs, policy management and D2D link establishment. The description of interfaces to interconnect the new entities is given in [30] and the references therein. The architectural changes required in the core network are outside the scope of this thesis and will not be discussed further.

1.3. Thesis Outline and Contributions

In this thesis, we focus on modeling cellular networks enhanced with network-assisted D2D and/or M2M communications. Since the future deployment of BSs and devices will be fairly dense and random, point-to-point characterization of a cellular or D2D link is insufficient and a more holistic view of the network is needed. In particular, the interference generated from the other nodes in the network is significant and cannot be ignored. To add to the complexity, this interference is also dependent on the physical location of the nodes and their distance to the receiver of interest. As pointed out by [31], the assumption of a deterministic grid deployment of BSs results in an upper bound on SINR coverage, but the analysis of grid topology involves long and cumbersome calculations resulting in intractable expressions offering very few insights. A recently emerged tool called stochastic geometry was introduced to circumvent this issue, by considering that the location of all nodes are independent of each other. Although, this is not the case in practice, the performance models derived using stochastic geometry are shown to provide lower bounds on the performance computed using actual deployment scenarios [32]. The difference between independent and grid deployment essentially boils down to the simplicity and increased tractability of the derived expressions using independent deployment. The independence of node locations can allow for substantial tools to be brought forward from stochastic geometry such as palm probability or generating functional, which are among others, described in detail in the next chapter.

The main objectives of this thesis can be broken down into three main categories

Chap.	Technology		Resource		Spectrum access		Research Objective	Salient feature(s)
	D2D	M2M	UL	DL	Underlay	Overlay		
3	x			x	x		Capacity enhancement using D2D	Massive MIMO
4	x		x			x		3GPP path loss models
5	x			x		x		Content caching
6	x	x	x			x	Exploiting D2D interfaces to aggregate M2M data	Data aggregation
7		x		x	x		Cognitive M2M network	Energy harvesting at MTDs

Table 1.1.: Schematic view of the chapter-wise outline of this thesis.

as shown in table 1.1. The first three chapters focus on the capacity enhancement of cellular networks with the help of D2D communication. In chapters 3 and 4, we consider that the D2D pairs have been created previously, while in chapter 5, we study how D2D pairs will be established. Since we consider network assisted D2D, the network can control mode switching of UEs to maximize the average throughput. Chapters 3 takes a proactive approach towards mode selection, where UEs are offloaded to D2D mode primarily to maximize the overall cell capacity. Chapters 4 proposes a simple distance based mode selection, where the UEs are offloaded to D2D mode if the distance separation between the pair exceeds a certain threshold. This threshold can then be tuned to adjust the average rate experienced by an arbitrary UE. Chapter 5 provides an enhanced model for the selection of the pairing device/helper for a UE to establish a D2D connection based on the content availability in the UE's neighboring devices. If the content is available, then the UE must be served in D2D mode by the selected helper. The content popularity, and other physical parameters are shown to affect the average rate experienced by a UE. The number of candidate helpers that can possibly serve the UE is a critical controlling parameter and may be tuned to obtain the desired performance. In chapter 6, we present a completely different use case of D2D communication as we study how UEs can establish D2D links with MTDs in their close proximity to

relay their data to the BS. In chapter 7, we investigate the coexistence of M2M networks with cellular networks. The cellular network is assumed to be unaware of the underlying M2M network, which follows the concept of CRNs guaranteeing a certain pre-defined QoS for the cellular UEs.

A detailed overview of the chapters is given as follows.

Chapter 2

This chapter reviews the fundamentals of modeling the signal-to-interference-and-noise ratio experienced by a receiver in a communication system. We describe the point process theory and its key role in developing tractable expressions to assess the performance of large scale networks. We specifically highlight the channel models, performance metrics and the modeling assumptions used throughout the thesis.

Chapter 3

In this chapter, we study the effectiveness of introducing D2D communication in a massive MIMO system, which is also envisaged to be an essential component of 5G networks. We consider an isolated cell, where the BS is equipped with massive MIMO capabilities, i.e. the number of antennas at the BS is much larger compared to the number of UEs. The extra degrees of freedom in a massive MIMO help achieve near-optimal performance with simple linear processing techniques. In the cellular downlink, we assume that the UEs can be offloaded by the BS to operate in D2D mode and share the cellular resources in the underlay. We answer the following research question: *given a certain number of UEs uniformly distributed inside a cell, what is the optimal offload fraction which maximizes the sum capacity in a massive MIMO system?* We consider a simple scenario where each UE has a dedicated D2D transmitter at a certain distance separation to serve the UE when it is offloaded to D2D mode. We develop tractable bounds for the overall spectral efficiency of the system and show that offloading a fraction of UEs to D2D mode does in fact improve the overall capacity. This offload fraction is strongly coupled with the number of antennas at the BS, the transmit powers of BS and UEs and also the D2D link distance.

Chapter 4

Simplistic assumptions are often used in the literature for channel modeling to obtain tractable expressions at the cost of accuracy. One such commonly used assumption is of a power-law path loss model for modeling large scale path loss according to which, the signal power attenuates by $r^{-\alpha}$, where r is the distance separation and α is the path loss exponent. This chapter considers the 3GPP recommended path loss model for the analysis of a cellular network overlaid with D2D communication. The practical path loss model is a multi-slope path loss model taking into account the presence of free space (FS), line-of-sight (LoS) and non-line-of-sight (NLoS) links. We use stochastic geometry to characterize the average throughput of the network in cellular uplink and compare our analysis with the reference scenerio, where a simple power-law path loss model is used. For D2D links, we assume that each UE has a dedicated D2D receiver at a random distance separation and if this separation is below a certain threshold, then D2D mode is selected. We show that in contrast to the previous findings, where the throughput saturates with the intensity of BSs, the average network throughput does not saturate, but there exists a unique BS intensity which maximizes the throughput.

Chapter 5

The previous chapters assume the presence of a dedicated D2D transmitter/receiver associated with each UE and the primary factor controlling mode selection is the distance separation between the pair. A fundamental assumption made in these chapters and the vast majority of existing literature is that the D2D transmitter is in possession of the content desired by its respective receiver and there is no explicit mention of how D2D pairs are created in the first place. In this chapter, we study how D2D pairs are created based on not only proximity but also content availability. We consider a large-scale cellular network model, where the BSs, requesting UEs and potential D2D helpers are distributed according to independent homogeneous Poisson point processes (HPPPs). We consider that content requests in UEs follow the Zipf distribution of content popularity and each D2D helpers has pre-cached some of the contents. A requesting UE is offloaded to D2D mode in downlink in the overlay if one of its k nearest D2D helpers has the requested content. We consider two helper selection schemes, namely i) nearest selection (NS), where the closest helper possessing the requested content is selected, and a baseline ii) uniform selec-

tion (US), where one of the k nearest helpers is uniformly selected without checking for content availability. Our main contribution is to characterize the distribution of distance between the UE and its ith nearest helper within the macrocell, which to the best of our knowledge, does not yet exist in the literature. This distribution effectively captures the fact that when the BS is aware of the devices in its coverage region only, the D2D links must also lie within the coverage region. We further evaluate the probabilities of a generic requesting UE to be served in D2D mode under both the NS and US schemes and derive the key performance metrics including the coverage and average rate experienced by the UE. The following conclusions can be drawn from this work:

- The probability of a UE to be served in D2D mode increases with the popularity of the requested content, implying that it is easier to find a D2D helper possessing the requested content.
- Always associating to the nearest helper ($k = 1$) in D2D is sub-optimal and even the D2D helpers located farther away can provide high-capacity links compared to the cellular link.
- Enhancing D2D opportunities does not always result in performance gains; as the number k of candidate D2D helpers increases, the UE is pushed to connect with helpers located farther resulting in performance degradation.

Chapter 6

Cellular networks were primarily designed to cater for human-to-human (H2H) communication, which is based on fewer, longer interactive sessions such as voice and video calls. The emphasis of cellular networks has always been to increase network capacity, throughput and the QoE of UEs. With the penetration of a bulk of connected MTDs, the cellular networks have to adapt and scale to support massive access. The characteristics of M2M traffic are drastically different as millions of MTDs transmit only a small amount of data sporadically. One way to address the issue of massive access of MTDs is to aggregate the generated data. This chapter explores how we can exploit the D2D connectivity between the cellular UEs and nearby MTDs for the aggregation of M2M data. UEs are ideal candidates for data aggregation as they are abundantly present everywhere. We develop a comprehensive analytical framework borrowing tools from stochastic geometry to analyze the

aggregation and trunking of M2M data using D2D interfaces in the cellular uplink. The uplink slot is divided such that a portion is used by the UEs to aggregate M2M data and the remaining time is used to transmit this data along with the UE's own data. We consider two different transmission strategies of the MTDs in the aggregation phase namely, i) Multiplexed, where the transmissions of the MTDs associated to a UE are muliplexed in frequency or time, and ii) Combined, where the MTDs associated to each UE transmit to it simultaneously. The UEs employ successive interference cancellation (SIC) to decode the signals of interest. We study the trade offs in adopting either the multiplexed or combined scheme. While the multiplexed scheme is more robust, it requires extensive signaling and control from the UE. The combined transmission, on the other hand, results in a lot of interference and also requires sophisticated decoding techniques.

The contributions of this work are highlighted as follows.

- The commonly used Voronoi tessellation used extensively to model BS coverage in stochastic geometry works is inappropriate for UE coverage, because of the low power of UEs. We therefore consider a Poisson hard sphere (PHS) model for UE coverage, where the UE coverage region (referred to as S-cell) is a disk centered at the UE with a random radius depending on its distance to the neighboring UEs.

- We derive the distributions of distance between the UE and the MTDs within the S-cell and characterize the probability that the rate requirements of MTDs and UEs are satisfied in aggregation and trunking phases respectively under both the multiplexed and combined scheme.

- We explore the tradeoff between a UE's performance in the uplink and the number of MTDs per UE. Our results reveal that a UE can successfully aggregate data from a few MTDs without compromising its QoS.

Chapter 7

In this chapter, we study the coexistence of a cellular network underlaid with an M2M network in a cognitive manner, where the MTDs reuse the same resources as the cellular UEs in the uplink. The MTDs in a cognitive M2M network (or CM2M) are assumed to harvest solar energy and those with sufficient harvested energy may transmit to a generic IoT controller (could be a smart phone, access point, etc.).

The transmissions of MTDs are also subject to the QoS constraint of the cellular UEs in terms of maximum outage probability threshold. Since the MTDs have to control the interference at the BS in the uplink, there can either be a few active transmitting MTDs at a high power or a large number of MTDs transmitting at a low power. Transmitting at high power results in better coverage and lower bit error rate, but it also means that the energy required to transmit will be high. Since energy is also a limited resource, there exists an inherent trade off which maximizes the probability of successful transmission of an MTD. We use stochastic geometry to model both the cellular network and underlying M2M network. To accurately analyze the energy outage, we include the effect of cloud cover on the harvested energy. The clearness index, which is the ratio between the measured irradiance and the clear sky irradiance, is a random variable depending on the intensity of clouds. We use the empirical distributions of clearness index for the city of Leeds, United Kindom to characterize the energy outage for various cloud cover intensities. Using the energy outage probability and the cellular QoS constraint, we derive the maximum permissible activation probability and the probability of successful transmission for the MTDs. We also demonstrate that the transmit power plays a key role in maximizing the success probability.

Chapter 8

This chapter concludes the thesis and discusses possible future directions for research.

1.4. List of Publications

- Asma Afzal, Syed Ali Raza Zaidi, Muhammad Zeeshan Shakir, Muhammad Ali Imran, Mounir Ghogho, Athanasios V. Vasilakos, Desmond C. McLernon, and Khalid Qaraqe. "The cognitive Internet of Things: a unified perspective. "published in ACM Springer Mobile Networks and Applications (MONET) journal, 2015: pg. 72-85.

- Syed Ali Raza Zaidi, Asma Afzal, Maryam Hafeez, Mounir Ghogho, Desmond C McLernon, Ananthram Swami, "Solar energy empowered 5G cognitive metro-cellular networks", published in IEEE Communications Magazine, July 2015.

- Asma Afzal, Syed Ali Raza Zaidi, Desmond C. McLernon and Mounir Ghogho "On the Analysis of DevicetoDevice Overlaid Cellular Networks in the Uplink under 3GPP Propagation Model" published in IEEE Wireless Communications and Networking Conference (WCNC), 2016.

- Asma Afzal, Syed Ali Raza Zaidi, Desmond C. McLernon and Mounir Ghogho, "On the Analysis of Cellular Networks with Caching and Coordinated Device-to-Device Communication" published in IEEE International Conference on Communications (ICC) 2016.

- Asma Afzal, Syed Ali Raza Zaidi, Desmond C. McLernon, Mounir Ghogho and Afef Feki, "M2M meets D2D: Harnessing D2D Interfaces for the Aggregation of M2M Data" published in IEEE International Conference on Communications (ICC) 2017.

- Asma Afzal, Afef Feki, Merouane Debbah, Mounir Ghogho and Desmond C. McLernon, "Leveraging D2D to Maximize the Spectral Efficiency of Massive MIMO Systems" published in Workshop on Spatial Stochastic Models for Wireless Networks (SpaSWiN), 2017.

- Asma Afzal, Syed Ali Raza Zaidi, Afef Feki, Desmond C. McLernon and Mounir Ghogho, "A Poisson Hard Sphere Model for the Analysis of M2M Data Aggregation via D2D Links" submitted to IEEE Transactions on Communications.

- Asma Afzal, Syed Ali Raza Zaidi, Desmond C. McLernon and Mounir Ghogho, "Information-centric offloading in cellular networks with coordinated Device-

to-Device Communication" submitted to IEEE Transactions on Wireless Communications.

2 Theory Preamble

In this chapter, we briefly describe the modeling techniques and methodology used for this thesis. We introduce the tool of stochastic geometry and present the mathematical preliminaries to create the foundations for easy understanding of this thesis.

2.1. Notations

Throughout this thesis, we use the following mathematical notations. The probability density of a random variable X is represented by $f_X(x)$ and the corresponding cumulative distribution function is denoted by $F_X(x)$, where the lowercase letter x is a particular realization of the random variable X in upper case. The expectation of a function with respect to X is represented by $\mathbb{E}_{\mathbb{X}}[.]$. A bold face lower-case letter, for e.g. $\mathbf{y}$, is used to denote a vector on $\mathbb{R}^2$ and its Euclidean norm is represented by $||\mathbf{y}||$. Symbol $\setminus$ refers to the exclusion of elements from a set, for instance, $[1,2,3]\setminus[1] = [2,3]$. The expression $b(o,r)$ represents a ball of radius r centered at the origin.

2.2. Channel Model

The wireless channel suffers from impairments including attenuation due to path loss, scattering and multi-path fading, where multiple copies of the same signal arrive at the receiver resulting in a constructive superposition at some points and destructive superposition at the others. Furthermore, the imperfection in hardware and electronics also introduces thermal noise, commonly known as the additive white Gaussian noise (AWGN). It is important to take into account all these impairments

to effectively analyze the characteristics of the received signal. The commonly used power law path loss model is defined as

$$PL = 10\log_{10}(A) + 10\alpha\log_{10}(r)\ \text{dB}, \tag{2.1}$$

where A is the carrier-frequency dependent path loss co-efficient, α is the path loss exponent and r is the distance separation. In chapter 4, we consider a more comprehensive multi-slope path model recommended by the 3GPP for the analysis of cellular and D2D links. The multi-slope path loss model accounts for the presence of line-of-sight (LoS) and the non-line-of-sight (NLoS) links. The probability of having a NLoS link increases with the distance separation r. The LoS probability expressions and 3GPP recommended values of A and α are all discussed in detail in chapter 4.

We consider Rayleigh fading to model multi-path propagation. It is a fitting model for rich-scattering urban environments, where the channel gain $g \sim \mathcal{C}(0,1)$ is a zero mean circularly symmetric Gaussian random variable with unit variance. This implies that the power gain of the faded signal $h =\mid g \mid^2 \sim \exp(1)$ will follow a unit mean exponential distribution, where $\mid g \mid$ is a Rayleigh distributed random variable.

2.3. Stochastic Geometry to Model Cellular Networks

A number of tools to model cellular networks such as game theory or graph theory are based on the instantaneous approach, i.e. they consider a snapshot of the network, where the devices have been deployed and the channel gains and interference powers already computed [33,34]. This information is then used to develop ways to optimize resources, maximize capacity and rate experienced by a given set of users at a given instant. Stochastic geometry modeling on the other hand makes use of the statistics of the network and takes into account the distribution of the location of devices, the distances between them and the distribution of the channel gains. These statistics are then used to characterize network performance in the average sense. Stochastic geometry helps us answer questions like, what is the average ergodic capacity experienced by an arbitrary user in the network and what is the outage

probability of an arbitrary user?

Stochastic geometry may only have been recently recognized as a useful tool for analyzing wireless networks, but it has been used extensively in the past to model biological and physical phenomena [35,36]. It builds on the theory of point processes, where the point process is a random collection of points in an d-dimensional space. For instance, the positions of plants and trees in a forest, imperfections in a material at a microscopic level, spread of an epidemic, etc., can all be represented as a collection of points. In the context of wireless networks, the BSs, UEs and other devices can be represented as points distributed randomly in space. Formally, a point process (PP) is denoted as $\Phi = \{x_i \in \mathbb{R}^d, i = 1, 2, 3, ...\}$ and the counting measure on Φ, can be written as $\Phi(\mathcal{A}) = n$, where n is the random number of points in the Borel set $\mathcal{A} \in \mathbb{R}^d$. The intensity measure of points in $\mathcal{A}$ can be written as

$$\Lambda(\mathcal{A}) = \mathbb{E}\left[\Phi(\mathcal{A})\right] = \int_{\mathbb{R}^d} \lambda(d\mathbf{x}), \tag{2.2}$$

where $\lambda(\mathbf{x})$ is the location dependent intensity function of a point process. For the case of homogeneous point processes (HPP), $\lambda(\mathbf{x}) = \lambda$ is a constant implying that the density of points is independent of the location and they are homogeneously distributed in $\mathbb{R}^d$. The intensity measure then reduces to

$$\Lambda(\mathcal{A}) = \lambda\, l_d(\mathcal{A}), \tag{2.3}$$

where $l_d(\mathcal{A})$ is a d-dimensional volume measure, also known as the Lebesgue measure on $\mathbb{R}^d$. When $d = 2$ and $\mathcal{A}$ is a disk of radius r, then the intensity measure of a homogeneous point process simplifies to $\Lambda(\mathcal{A}) = \lambda\pi r^2$.

Thinning

This is a process which removes the points from a process. This process is dictated by a thinning function $p(\mathbf{x}) \in [0, 1]$, which is the location-dependent probability. The intensity $\Lambda_t(\mathcal{A})$ of the resulting process after thinning can be written as

$$\Lambda_t(\mathcal{A}) = \int_{\mathcal{A}} p(\mathbf{x})\Lambda(d\mathbf{x}). \tag{2.4}$$

When the thinning probability $p(\mathbf{x}) = p$ is constant, the point process is homogeneously thinned in the entire space and the resulting intensity of the thinned process

can be simplified as $\Lambda_t(\mathcal{A}) = p\Lambda(\mathcal{A})$.

2.3.1. Poisson Point Process

Throughout this thesis, we will make use of the Poisson point process (PPP) due to its many convenient properties. The PPP of intensity measure Λ is defined by the means of its d-dimensional distributions as

$$\mathbb{P}\left\{\Phi(\mathcal{A}_1) = n_1, \Phi(\mathcal{A}_2) = n_2, .., \Phi(\mathcal{A}_k) = n_k\right\} = \prod_{i=1}^{k} \left(\frac{\Lambda(\mathcal{A}_i)^{n_i}}{n_i!} \exp\left(-\Lambda(\mathcal{A}_i)\right) \right) \quad (2.5)$$

for every bounded, mutually disjoint set $\mathcal{A}_k$, $k = 1, 2, ...$ In other words, the number of points in the Borel set $\mathcal{A}_i$ is Poisson distributed with mean $\Lambda(\mathcal{A}_i)$, i.e. $\Phi(\mathcal{A}_i) \sim \text{Poisson}\left(\Lambda(\mathcal{A}_i)\right)$.

Void probability

It is the probability that the point process does not have a point in a given bounded set. It can be written as $v_{\mathcal{A}} = \mathbb{P}\left[\Phi(\mathcal{A}) = 0\right]$. For a PPP, the void probability reduces to

$$v_{\mathcal{A}} = \exp\left(-\Lambda(\mathcal{A})\right) \quad (2.6)$$

Slivnyak's Theorem

In the study of point processes, it is quite common to analyze the statistics at a given point of the process. Particularly in the case of wireless networks, we are interested in quantifying coverage and rate at the typical user, i.e. the user located at the origin. Palm theory formalizes the concept of a general point process to have a fixed point at a certain location. For the point process with a property or an event E, we can write the palm distribution as

$$\begin{aligned} \mathbb{P}\left\{\Phi \text{ has a property } E \parallel o\right\} &= \mathbb{P}\left\{\Phi \text{ has a property } E \mid o \in \Phi\right\} \\ &= \mathbb{P}^o[E], \end{aligned} \quad (2.7)$$

where $o = \{0, 0\}$ is the origin and $\mathbb{P}^o[E]$ is a shorthand representation of the palm distribution of the point process Φ with a fixed point at o. In some applications however, it is important to exclude the point under consideration from the analysis.

For instance in measuring the interference in a cellular downlink, all the active BSs are generating interference apart from the one the user is associated with. The reduced palm distribution is the probability that the point on which we condition on is not included in the distribution. It can be represented as

$$\begin{aligned}\mathbb{P}\left\{\Phi\backslash\{x\} \in E \mid x \in \Phi\right\} &= \mathbb{P}\left\{\Phi\backslash\{x\} \in E \parallel x\right\} \\ &= \mathbb{P}^{!x}[E], \end{aligned} \tag{2.8}$$

where $\mathbb{P}^{!x}[E]$ implies the reduced palm distribution where the point located at x is removed from Φ. According to Slivnyak's theorem, the reduced palm distribution of a PPP is equivalent to the original palm distribution $\mathbb{P}^{!x}[E] = \mathbb{P}^{x}[E]$, implying that the effect of removing or adding a point does not affect the properties of the point process. This is a very powerful theorem enabling us to analyze the performance at the typical user without any loss of generality.

Probability Generating Functional

The probability generating functional of a point process with respect to a non-negative bounded function f on $\mathbb{R}^d$ such that $0 \leq f(x) \leq 1$ is analogous to the probability generating function of a non-negative integer-valued random variable. It is defined as

$$G(f) = \mathbb{E}\left[\prod_{x\in\Phi} f\left(x\right)\right], \tag{2.9}$$

where the product is performed for all points in the point process Φ. For a PPP, the generating functional can be written as

$$G(f) = \exp\left(-\int_{\mathbb{R}^d} \left(1 - f\left(x\right)\right) \Lambda(d\mathbf{x})\right). \tag{2.10}$$

2.4. Modeling Cellular Networks: An Example

We now present a simple example of the characterization of the distribution of the average received power at a typical UE using stochastic geometry.

Let us consider that the base stations are distributed in a 2D network according to a HPPP Φ with intensity λ. We assume nearest association, i.e. an arbitrary

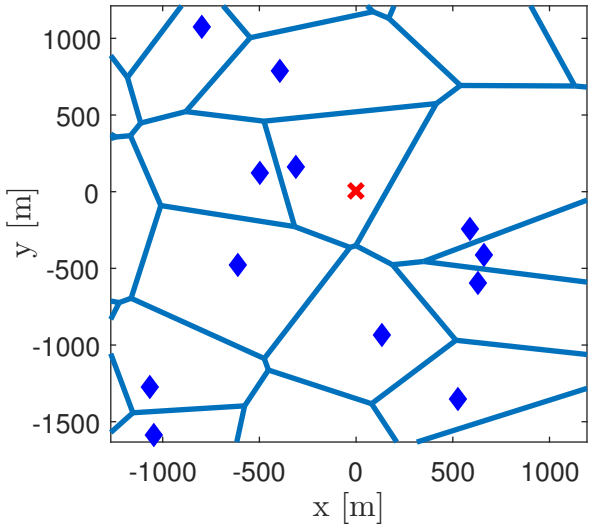


Figure 2.1.: A realization of a cellular network: BSs are represented by blue diamonds and the typical UE by a red cross.

UE will be served by the closest BS. The boundaries of coverage regions, or the macrocells of the BSs, correspond to a Voronoi tessellation as displayed in Figure 2.1.

The power received at the typical UE from the BS at a distance $R = r$ can be written as

$$P_r = P_t h r^{-\alpha} \tag{2.11}$$

where P_t is the transmit power, $h \sim \exp(1)$ is the channel power and α is the path loss exponent. Notice that we assume that A takes the value 1 in (2.1) without loss of generality. The distribution of R is obtained using the concept of void probability of point processes as follows.

$$\begin{aligned}\mathbb{P}[R > r] &= \mathbb{P}[\text{No BS inside the disk of radius } r] \\ &= \exp\left(-\lambda\pi r^2\right).\end{aligned}$$

Differentiating $1 - \mathbb{P}[R > r]$, we get the PDF as follows.

$$f_R(r) = 2\pi\lambda r \exp\left(-\lambda\pi r^2\right). \tag{2.12}$$

The CDF of the received power P_r can then be written as

$$\begin{aligned} F_{P_r}(p) &= \mathbb{P}\left[P_t h r^{-\alpha} < p | R\right] \\ &= \mathbb{E}_R\left[\mathbb{P}\left[h < \frac{p}{P_t} r^{\alpha} | R\right]\right] \\ &= 1 - \int_0^{\infty} \exp\left(-\frac{p}{P_t} r^{\alpha}\right) f_R(r)\, dr. \end{aligned} \tag{2.13}$$

When $\alpha = 4$, (2.13) simplifies to

$$F_{P_r}(p) = 1 - \kappa\sqrt{\frac{\pi}{p}} \exp\left(-\frac{\kappa^2}{p}\right) \operatorname{erfc}\left(\frac{\kappa}{\sqrt{p}}\right), \tag{2.14}$$

where $\kappa = \frac{1}{2}\lambda\pi\sqrt{P_t}$ and $\operatorname{erfc}(z) = \frac{2}{\pi}\int_z^{\infty} \exp(-t^2)\, dt$ is the complementary error function. We compare our analysis in (2.14) with Monte Carlo simulations. For the simulations, we consider a large circular area of radius 1 km and conduct 10^4 iterations. In each iteration, the number of BSs is drawn from a Poisson distribution with intensity λ and the BSs are distributed uniformly in the region. The power received at the origin from the nearest BS is then calculated for each iteration. We see from Figure 2.2 that the expression in (2.14) accurately captures the distribution of power received at the typical UE.

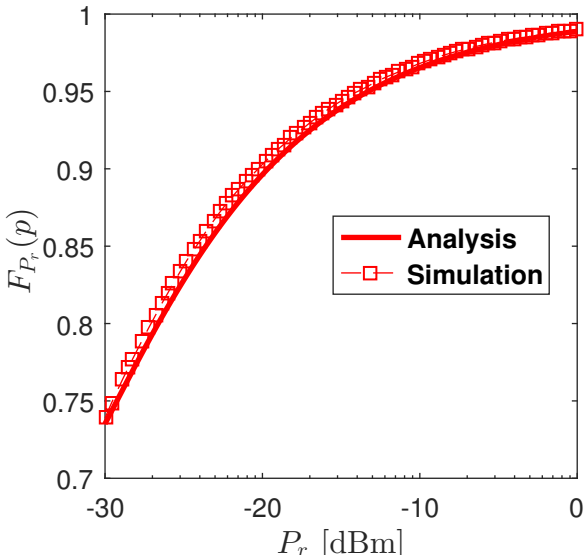


Figure 2.2.: CDF of desired power at the typical UE: $\lambda = 1.27 \times 10^{-4}$, $P_t = 1$W.

It is evident from the simple example above that stochastic geometry analysis can

help develop tractable expressions for the network performance, thereby alleviating the need to perform cumbersome and time consuming simulations. In the upcoming technical chapters, we study more complicated scenarios where the cellular networks are enhanced with D2D or M2M communications. We then analyze the statistics of the generated interference to characterize the key performance metrics.

3 Offloading to D2D in a Single-Cell Scenario

In this chapter, we consider a single-cell scenario where the BS is equipped with large, but finite number of antennas and the total number of UEs is kept fixed. By offloading a fraction of UEs to communicate via short-range direct D2D communication, the overall cell capacity can be improved. The key design question is that what fraction of UEs should be offloaded to D2D mode in order to maximize the aggregate cell level throughput. We demonstrate that there exists an optimal user offload fraction, which is strongly coupled with the network parameters such as the number of antennas at the BS, D2D link distance and the transmit SNR of both the UE and the BS. By carefully tuning the offload fraction, capacity gains of up to 5× can be achieved.

3.1. Introduction

Massive MIMO is an attractive emerging technology as it enables simultaneous communication between the BS and multiple UEs at the same time/frequency resource [37]. The distinct feature of massive MIMO is that the number of antennas deployed at the BS is much larger than the number of UEs to be served. This allows for significant improvements in link reliability and data rates due to increased spatial directivity. The additional degrees of freedom alleviate the need for sophisticated signal processing techniques and simple linear processing achieves near-optimal performance [38]. In this chapter we will study massive MIMO in conjunction with D2D communication. With the help of D2D, UEs in close proximity can communicate directly without the intervention of the BS [16]. The short range of D2D

communication improves coverage and it also reduces the burden of access on the BS and the core network.

Even though D2D been studied extensively in the context of cellular networks with BSs equipped with a single antenna, the analysis of D2D with massive MIMO is still in its infancy. In [39] and [21], the authors analyze an isolated cell with a single cellular UE and D2D pair and investigate how the excess antennas at BS can eliminate the interference at the D2D receiver. The sum capacity of an isolated cell with a fixed number of cellular UEs and a random number of D2D pairs has been studied in [40] for the case of cellular UL. Expressions for signal-to-interference-and-noise ratio (SINR) are derived for both the cellular and D2D cases for fixed spatial locations of UEs and the randomness is accounted for in simulations. The corresponding DL analysis is conducted in [41] and the density of D2D pairs maximizing the sum capacity is explored.

The research on massive MIMO with D2D thus far does not consider dynamic mode selection for the UEs. It is only in [42] that the authors consider mode switching for a UE (between cellular and D2D) in cellular UL for a simple network setting with a single D2D pair. The optimality region for D2D mode satisfying the link spectral efficiency (SE) requirements is defined around the D2D transmitter. The obtained results, however, cannot be directly translated to DL and scaled for multiple D2D pairs case as the location of interfering UEs is assumed to be fixed. The interference from the active D2D pairs is highly dependent on their distance from the UE under consideration and will significantly impact the findings. Also, the link SE metric does not capture the overall cell capacity.

Motivated by this, we study the offloading problem for a single cell scenario in DL, where a fixed number N of UEs is distributed uniformly around the BS. We focus on D2D in DL time slot as it is more suited for massive MIMO scenario. This is because the BS can make use of the excess degrees of freedom to mitigate interference at the D2D receivers, whereas it is not possible in the UL with single antenna UEs [21, 39, 43, 44]. While D2D communication between UEs in close proximity can provide high data rates, the transmit power of a BS is much higher than a UE and it is not clear under what circumstances offloading is a better choice. There exists an inherent tradeoff as offloading UEs to operate in D2D mode will improve the SE, but at the same time, a large number of D2D UEs will generate higher interference causing the SE to drop. The incentive of this work is to answer the following question: *Given a certain number of UEs inside a cell, what is the optimal*

offload fraction which maximizes the sum capacity in a massive MIMO system and how do the system parameters affect this fraction?. Our main contribution is to explore this trade off and derive closed-form expressions for the approximation of the unconditional overall capacity.

3.2. System Model

We consider a time division duplex (TDD) DL transmission scenario where the BS is equipped with M antennas and $N < M$ single antenna UEs distributed uniformly in an annular region of inner radius R_{min} and outer radius R_{max} centered at the BS as shown in Figure 3.1. K out of N UEs are served directly by the BS, while

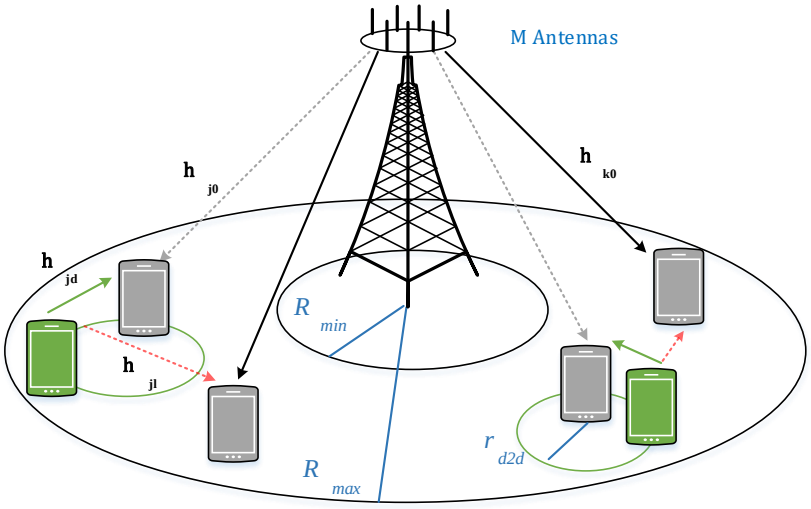


Figure 3.1.: System Model.

the remaining $(N - K)$ UEs are offloaded to D2D mode. Each of the $(N - K)$ D2D receiving UEs is associated to a unique D2D transmitter UE located randomly at the perimeter of a disk of radius r_{d2d} centered at the UEs. These transmitters can be thought of as UEs who are willing to establish D2D connections with their neighboring UEs to share files of common interest [45].

Without any loss of generality, the set of all N UE locations can be written as $\mathcal{U} = \{\mathbf{x}_1, .., \mathbf{x}_K, \mathbf{x}_{K+1}, .., \mathbf{x}_N\}$. Assuming that the BS is located at the origin, the

distance between the k^{th} UE and the BS $r_{k0} = \|\mathbf{x}_k\|$ is distributed as

$$f_{r_{k0}}(x) = \frac{2x}{R_{max}^2 - R_{min}^2}, R_{min} \leq x \leq R_{max}. \tag{3.1}$$

We adopt a simple power-law path loss model where the signal power attenuates according to $r^{-\alpha_m}, m = \{b, d\}$, where r is the distance separation and α_m denotes the path loss exponent in mode m where $m = b$ stands for cellular mode when the UE is served by the BS and $m = d$ stands for D2D mode. The BS-UE and UE-UE links suffer from small scale Rayleigh fading. This implies that the channel gain is independent and identically distributed (i.i.d) complex Gaussian variable with zero mean and unit variance. We further assume that the D2D pairs share the same resources as the cellular UEs and hence, both the BS-UE and UE-UE links interfere with each other. The BS is considered to have full channel state information (CSI) of the UEs and it employs zero-forcing beamforming (ZFBF) precoding. As a result, there is no signal leakage within the cellular UEs. The BS transmits a total power P_b, which is equally distributed for cellular UEs and the D2D UEs transmit a fixed power P_d, where $P_d < P_b$. The preliminary analysis for the SINR at the UEs in cellular and D2D modes is presented as follows.

3.2.1. Cellular Mode

The signal received at the k^{th} cellular UE under ZFBF can be written as

$$\begin{aligned} y_k &= \sqrt{\frac{P_b r_{k0}^{-\alpha_b}}{K}} \left(\mathbf{h}_{k0}^{BS-UE}\right)^H \mathbf{w}_{k0}^{BS} s_{k0}^{BS-UE} \\ &\quad + \sqrt{P_d} \sum_{l=1}^{N-K} \sqrt{r_{kl}^{-\alpha_d}} h_{kl}^{UE-UE} s_l^{UE} + v_k^{BS}, \end{aligned} \tag{3.2}$$

where, $\mathbf{h}_{k0}^{BS-UE} \in \mathbb{C}^{M\times 1}$ is a vector of M Rayleigh fading channel gains, v_k^{BS} is the zero mean additive white Gaussian noise (AWGN) with variance σ_{BS}^2, s_{k0}^{BS-UE} is the complex scalar signal, $\mathbf{w}_{k0} \in \mathbb{C}^{M\times 1}$ is the precoding vector and $(.)^H$ denotes the Hermitian transpose operator. To satisfy the maximum BS power constraint, s_{k0}^{BS-UE} and $\mathbf{w}_{k0}^{BS} = \mathbf{g}_{k0}^{BS}/\|\mathbf{g}_{k0}^{BS}\|^2$. The precoding vector and the complex signal is normalized such that $\mathbb{E}\left[\left|s_{k0}^{BS-UE}\right|^2\right] = \|\mathbf{w}_{k0}\|^2 = 1$. The un-normalized precoding

matrix $\mathbf{G}^{BS-UE} = \left[\mathbf{g}_{10}^{BS-UE}, .., \mathbf{g}_{K0}^{BS-UE}\right]$ under ZF precoding can be written as

$$\mathbf{G}^{BS-UE} = \mathbf{H}^{BS-UE}\left(\left(\mathbf{H}^{BS-UE}\right)^{H}\mathbf{H}^{BS-UE}\right)^{-1},$$

where $\mathbf{H}^{BS-UE} = \left[\mathbf{h}_{10}^{BS-UE}, .., \mathbf{h}_{K0}^{BS-UE}\right]$. The second term in (3.2) denotes the interference signal from all the $(N-K)$ active D2D transmitters to the k^{th} cellular UE, where $r_{kl} = \|\mathbf{x}_k - \mathbf{x}_l\|$ is the distance between the k^{th} UE and the l^{th} D2D transmitter and s_l^{UE} is the information symbol transmitted by the l^{th} D2D transmitter. The average SE for the k^{th} UE in cellular mode can be written as

$$SE_k^{BS-UE} = \mathbb{E}\left[\log_2\left(1 + SINR_k^{BS-UE}\right)\right], \tag{3.3}$$

where

$$SINR_k^{BS-UE} = \frac{\gamma_b r_{k0}^{-\alpha_b}\left|\left(\mathbf{h}_{k0}^{BS-UE}\right)^{H}\mathbf{w}_{k0}^{BS}\right|^2}{\gamma_d\sum_{l=1}^{N-K} r_{kl}^{-\alpha_d}\left|h_{kl}^{UE-UE}\right|^2 + 1} \tag{3.4}$$

and $\gamma_b = {}^{P_b}/_{\sigma_{BS}^2}$ and $\gamma_d = {}^{P_d}/_{\sigma_{UE}^2}$ are the cellular and D2D transmit signal-to-noise ratios (SNRs).

3.2.2. D2D Mode

The signal received at the j^{th} UE x_j in D2D mode from its corresponding dth D2D transmitter can be written as

$$y_j = \sqrt{P_d r_{d2d}^{-\alpha_d}}\, h_{jd}^{UE-UE} s_d^{UE} + I_j^{UE-UE} + I_j^{BS-UE} + v_d^{UE},$$

where v_d^{UE} is the zero mean additive white Gaussian noise (AWGN) with variance σ_{UE}^2, the interference signal received by the j^{th} UE in D2D mode from other active D2D transmitters can be written as

$$I_j^{UE-UE} = \sqrt{P_d}\sum_{l\neq d}^{N-K}\sqrt{r_{jl}^{-\alpha_d}}\, h_{jl}^{UE-UE} s_l^{UE},$$

and the interference from the BS is

$$I_j^{BS-UE} = \sqrt{\frac{P_b r_{j0}^{-\alpha_b}}{K}} \sum_{k=1}^{K} \left(\mathbf{h}_{j0}^{BS-UE}\right)^H \mathbf{w}_{k0}^{BS} s_{k0}^{BS-UE}.$$

The average SE for the j^{th} UE in D2D mode can then be written as $SE_j^{UE-UE} = \mathbb{E}\left[\log_2\left(1 + SINR_j^{UE-UE}\right)\right]$, where

$$SINR_j^{UE-UE} = \frac{\gamma_d r_{d2d}^{-\alpha_d} \left|h_{jd}^{UE-UE}\right|^2}{\left|I_j^{BS-UE}\right|^2/\sigma_{UE}^2 + \left|I_j^{UE-UE}\right|^2/\sigma_{UE}^2 + 1}. \tag{3.5}$$

3.3. Spectral Efficiency Analysis

The goal of this work is to evaluate the optimal fraction of UEs to be offloaded in D2D mode. We define our performance metric as follows.

Definition 3.1. Given a fixed number of UEs N, the maximum attainable overall capacity is given by the following optimization problem

$$C_{tot} = \max_{K} \left(\sum_{k=1}^{K} SE_k^{BS-UE} + \sum_{j=K+1}^{N} SE_j^{UE-UE}\right)$$

where K is the number of UEs in cellular mode. Our aim is to evaluate $\mu^* = {}^{(N-K^*)}/_{N}$, which is the optimal offload fraction, where K^* is the optimal number of remaining cellular UEs.

In the following subsections, we present our analysis pertaining to the cellular and D2D SEs.

3.3.1. Cellular Mode

To characterize the average SE of a generic UE operating in cellular mode in (3.3), we need to average the effect of randomness due to the UE locations and the channel gains. The following Lemma provides the SE of a UE in cellular mode conditioned on UE locations.

Lemma 3.1. *Conditioned on the location of the UEs, the average SE of a UE in*

cellular mode can be approximated as

$$SE_k^{BS-UE}|r_{k0}, r_{kl} \approx \log_2\left(1 + \frac{\gamma_b\,(M-K)\,r_{k0}^{-\alpha_b}}{K\left(1+\gamma_d\sum_{l=1}^{N-K} r_{kl}^{-\alpha_d}\right)}\right). \tag{3.6}$$

Proof. Since $\log\left(1+x^{-1}\right)$ is convex in x, we employ Jensen's inequality to obtain

$$SE_k^{BS-UE}|r_{k0}, r_{kl} \geq \log_2\left(1 + \mathbb{E}\left[\frac{1}{SINR_k^{BS-UE}}\right]^{-1}\right). \tag{3.7}$$

The inverse of the desired power in (3.4) is an inverse chi-squared random variable such that $\frac{1}{2\left|\left(\mathbf{h}_{k0}^{BS-UE}\right)^H \mathbf{w}_{k0}^{BS}\right|^2} \sim$ Inv-$\chi^2_{2(M-K+1)}$. This is because the isotropic M dimensional vector is projected onto $M-K+1$ dimensional beamforming space [43]. As a consequence, we have $\mathbb{E}\left[\frac{1}{\left|\left(\mathbf{h}_{k0}^{BS-UE}\right)^H \mathbf{w}_{k0}^{BS}\right|^2}\right] = (M-K)^{-1}$. The interference power from each D2D UE is a unit mean exponential random variable $\left|h_{kl}^{UE-UE}\right|^2 \sim \exp(1)$. Exploiting the independence of these random variables and substituting the expected power values in (3.7), we obtain (3.6). The approximation sign is used as the bound is very tight for a wide range of network parameters. □

We de-condition $SE_k^{BS-UE}|r_{j0}, r_{jl}$ in (3.6) with respect to distances in the following Lemma 3.2 and Proposition 3.1.

Lemma 3.2. *The average SE of an arbitrary UE in cellular mode conditioned on the location of interfering D2D UEs can be approximated in closed-form for $\alpha_b = \alpha_d = 4$ as*

$$SE_k^{BS-UE}|r_{kl}\left(\beta_{kl}\right) \approx \frac{\log_2\left(1+\beta_{kl}\right)}{R_{max}^2} + \frac{2\sqrt{\beta_{kl}}}{\ln\left(2\right)}\tan^{-1}\left(\sqrt{\frac{1}{\beta_{kl}}}\right), \tag{3.8}$$

where $\beta_{kl}\left(\psi\right) = \frac{\gamma_b(M-K)}{K(1+\gamma_d\psi)R_{max}^4}$ and $\psi = \sum_{l=1}^{N-K} r_{kl}^{-4}$.

Proof. The proof follows by averaging (3.6) over r_{k0} which is distributed according to (3.1). □

Proposition 3.1. *The bounds on the unconditional average SE of a UE in cellular mode $SE_{k,LB}^{BS-UE} \leq SE_k^{BS-UE} \leq SE_{k,UB}^{BS-UE}$ can be written in closed-form as*

$$SE_{k,UB}^{BS-UE} = SE_k^{BS-UE}|r_{kl}\left(\beta_c^{UB}\right), \tag{3.9}$$

where $\beta_c^{UB} = \beta_{kl}\left(\psi_c^{UB}\right)$ *with* $\psi_c^{UB} = (N-K)\,\mathbb{E}\left[r_{kl}\right]^{-4}$ *and* $\mathbb{E}\left[r_{kl}\right] = \frac{128}{4\pi} R_{max}$ *for the upper bound and*

$$SE_{k,LB}^{BS-UE} = SE_k^{BS-UE}|r_{kl}\left(\beta_c^{LB}\right), \tag{3.10}$$

where $\beta_c^{LB} = \beta_{kl}\left(\psi_c^{LB}\right)$, *with* $\psi_c^{LB} = (N-K)\,\mathbb{E}\left[r_{kl}^{-4}\right]$ *and*

$$\mathbb{E}\left[r_{kl}^{-4}\right] \approx \rho_g^{-1}\left[-\frac{3\sqrt{(4R_{max}^2-1)}}{4R_{max}^2} + \left(1+\frac{1}{2R_{max}^2}\right)\cos^{-1}\left(\frac{1}{2R_{max}}\right)\right]$$

$\rho_g = \sqrt{(4R_{max}^2-1)}\left(\frac{2R_{max}^2+1}{8R_{max}^2}\right)$ *for the lower bound.*

Proof. The first term in (3.8) is of the form $\log\left(1+\left(A+B\,r_{kl}^{-4}\right)^{-1}\right)$, which is concave in r_{kl} and convex in r_{kl}^{-4} for $A, B > 0$. The second term is of the form $g\left(r_{kl}, A, B\right) = \left(A+B\,r_{kl}^{-4}\right)^{-1/2}\tan^{-1}\left(\left(A+B\,r_{kl}^{-4}\right)^{1/2}\right)$, which can be easily shown to be a quasi-concave function. The function $g(r_{kl}, A, B)$ transitions from convex to concave and contains a point of inflection at $r_{kl} > 0$. Simulations reveal that this inflection point takes a very small value of $r_{kl} < 10$m for various values of A and B as seen in Figure 3.2 . Since $0 < r_{kl} < 2R_{max}$, we treat $g(r, A, B)$ as concave

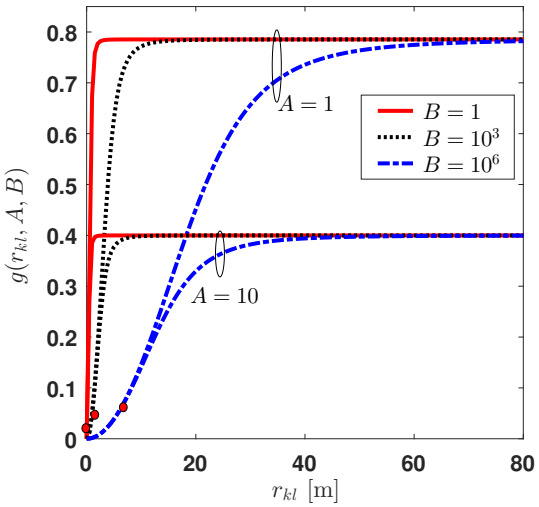


Figure 3.2.: Effect of A and B on $g(r_{kl}, A, B)$. Point of inflection is shown by red filled circles.

and make use of Jensen's inequality to shift the expectation operator inside these

functions to obtain the bounds[1]. The D2D UEs are i.i.d distributed and their respective transmitters are uniformly located at a fixed distance r_{d2d}. For tractability, we assume that $R_{min} = 0$. This does not impact the result as $R_{max} \gg R_{min}$. The effective distance between the k^{th} UE and l^{th} transmitting D2D UE is then distributed according to [46]

$$f_{r_{kl}}(x) = \frac{2x}{\pi R_{max}^2}\left(2\cos^{-1}\left(\frac{x}{2R_{max}}\right) - \frac{x}{R_{max}}\sqrt{1-\left(\frac{x}{2R_{max}}\right)^2}\right),\ 0 \le x \le 2R_{max},$$

where $\mathbb{E}[r_{kl}] = \frac{128}{4\pi}R_{max}$. It is tricky to obtain $\mathbb{E}\left[r_{kl}^{-4}\right]$ as the integral $\int_0^\infty x^{-4} f_{r_{kl}}(x)\,dx$ diverges on the lower bound. To tackle this issue and avoid singularity, we introduce a minimum separation distance of 1m. Therefore, we have $\mathbb{E}\left[r_{kl}^{-4}|r_{kl} \ge 1\right] = \int_{x=1}^{2R_{max}} x f_{r_{kl}}(x|r_{kl} \ge 1)\,dx$, where $f_{r_{kl}}(x|r_{kl} \ge 1) = {}^{f_{r_{kl}}(x)}/_{\mu_g}, 1 \le x \le 2R_{max}$ and $\mu_g = \mathbb{P}[r_{kl} \ge 1]$. This completes the proof. □

3.3.2. D2D Mode

Similar to the methodology adopted in the previous section, we obtain the SE of a UE in D2D mode conditioned on UE locations in the following Lemma.

Lemma 3.3. *Conditioned on the location of the UEs, the average SE of a UE in D2D mode can be approximated as*

$$SE_j^{UE-UE}|r_{j0}, r_{jl} \approx \log_2\left(1 + \frac{\gamma_d r_{d2d}^{-\alpha_d}}{1 + \gamma_d \sum_{l\neq d}^{N-K} r_{jl}^{-\alpha_d} + \gamma_b r_{j0}^{-\alpha_c}}\right). \tag{3.11}$$

Proof. We follow a different approach compared to the proof of Lemma 3.1. Since the desired power is exponentially distributed $\left|h_{jd}^{UE-UE}\right|^2 \sim \exp(1)$, the expected value of its inverse does not exist. We therefore exploit the concavity of $\log(1+x)$ to obtain $\log_2\left(1 + \mathbb{E}\left[SINR_j^{UE-UE}\right]\right)$. We can write

$$\mathbb{E}\left[SINR_j^{UE-UE}\right] = \int_0^\infty \mathrm{e}^{-s_z}\mathbb{E}\left[\exp\left(-s_z {I_j}^{BS-UE}\right)\right]\mathbb{E}\left[\exp\left(-s_z {I_j}^{UE-UE}\right)\right]dz,$$

where $s_z = {}^{z\, r_{d2d}^{\alpha_d}}/_{\gamma_d}$. Since $\left(\mathbf{h}_{j0}^{BS-UE}\right)^H$ is independent of $\mathbf{w}_k^{BS}$, $\left|\left(\mathbf{h}_{j0}^{BS-UE}\right)^H \mathbf{w}_k^{BS}\right|^2 \sim \exp(1)$. $\left|{I_j}^{BS-UE}\right|^2$ is the superposition of K independent data streams, which im-

[1]The analysis in the subsequent sections also assumes $g(r, A, B)$ to be concave.

plies $\sum_{k=1}^{K} 2\left|\left(\mathbf{h}_{j0}^{BS-UE}\right)^{H}\mathbf{w}_{k}^{BS}\right|^{2} \sim \chi_{2K}^{2}$. For the D2D interference power $\left|I_{j}^{UE-UE}\right|^{2}$, we have $\left|h_{jl}^{UE-UE}\right|^{2} \sim \exp(1)$. To ensure tractability, we invoke Jensen's inequality once again to draw the expectation inside the exponential to obtain (3.11). □

Similar to the analysis for cellular mode, we derive the expressions for unconditional SE of a UE in D2D mode as follows.

Lemma 3.4. *The average SE of an arbitrary UE in D2D mode conditioned on the location of interfering D2D UEs can be approximated in closed-form for $\alpha_b = \alpha_d = 4$ as*

$$\begin{aligned} SE_{j}^{UE-UE}|r_{jl}\left(\beta_{jl1},\beta_{jl2}\right) \;\approx\; & \frac{1}{R_{max}^{2}}\log_{2}\left(1+\frac{\gamma_{d}r_{d2d}^{-4}}{\beta_{jl1}+\gamma_{b}/R_{max}^{4}}\right) \\ & +\frac{2}{\ln\left(2\right)}\left[\sqrt{\frac{\gamma_{b}/R_{max}^{4}}{\beta_{jl2}+\gamma_{d}r_{d2d}^{-4}}}\tan^{-1}\left(\sqrt{\frac{\beta_{jl2}+\gamma_{d}r_{d2d}^{-4}}{\gamma_{b}/R_{max}^{4}}}\right)\right. \\ & \left.-\sqrt{\frac{\gamma_{b}/R_{max}^{4}}{\beta_{jl2}}}\tan^{-1}\left(\sqrt{\frac{\beta_{jl2}}{\gamma_{b}/R_{max}^{4}}}\right)\right] \end{aligned} \tag{3.12}$$

where $\beta_{jl1}\left(\psi\right) = \beta_{jl2}\left(\psi\right) = 1+\gamma_{d}\psi$ and $\psi = \sum_{l\neq d}^{N-K} r_{jl}^{-4}$.

Proof. The proof follows by averaging (3.11) over r_{k0}. □

Proposition 3.2. *The bounds on the unconditional average SE of a UE in D2D mode $SE_{j,LB}^{UE-UE} \leq SE_{j}^{UE-UE} \leq SE_{UB}^{UE-UE}$ can be written in closed-form as*

$$SE_{j,UB}^{UE-UE} = SE_{j}^{UE-UE}|r_{jl}\left(\beta_{d}^{UB},\beta_{d}^{LB}\right), \tag{3.13}$$

where $\beta_{d}^{UB} = \beta_{jl1}\left(\psi_{d}^{UB}\right)$ with $\psi_{d}^{UB} = \left(N-K-1\right)\mathbb{E}\left[r_{kl}\right]^{-4}$ and $\beta_{d}^{LB} = \beta_{jl1}\left(\psi_{d}^{LB}\right)$ with $\psi_{d}^{LB} = \left(N-K-1\right)\mathbb{E}\left[r_{kl}^{-4}\right]$ for the upper bound and

$$SE_{j,LB}^{UE-UE} = SE_{j}^{UE-UE}|r_{kl}\left(\beta_{d}^{LB},\beta_{d}^{UB}\right), \tag{3.14}$$

for the lower bound.

Proof. The proof is similar to that of Prop. 3.1 with the exception that there is a negative sign inside the second term of (3.12). It can be easily shown that for $g\left(r_{jl},A,B\right)$, we have $g\left(r_{jl},A_{1},B\right) \leq g\left(r_{jl},A_{2},B\right)$ for $A_{1} \geq A_{2}$ or $A_{1}-A_{2} = \gamma_{d}r_{d2d}^{-4} > 0$. Therefore, if we re-write $SE_{j}^{UE-UE}|r_{jl} = T_{1}+T_{2}$, then T_{2} exhibits the opposite

behavior of T_1. It is concave in r_{jl}^{-4} and convex in r_{jl}. The coefficient $(N-K-1)$ in ψ_1 and ψ_2 denotes the number of interfering D2D pairs, excluding the one on which the performance is being measured. □

3.4. Results and Discussion

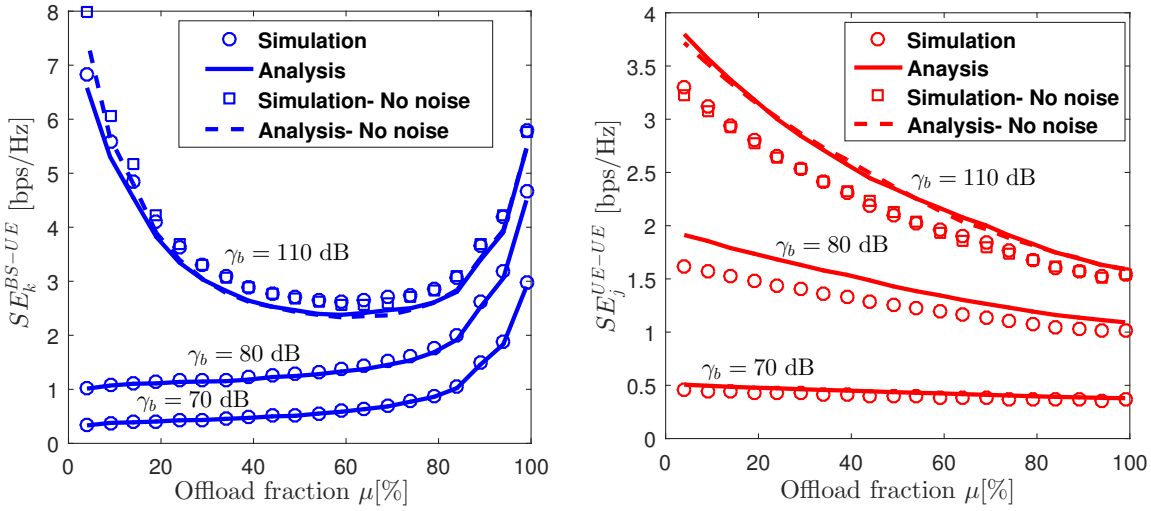
Parameter	Value
BS antennas M, Total users N	200, 100
BS coverage radius (Outer R_{max}, Inner R_{min})	200 m, 2m
D2D range r_{d2d}	12 m
Path loss exponents: α_c, α_d	4, 4
Ratio of cellular and D2D SNR $(\gamma_b/\gamma_d)_{\text{dB}}$	30 dB

Table 3.1.: List of simulation parameters.

In this section, we present the evaluation procedure adopted to assess the performance of the offloading mechanism. As a first step, we validate our analysis in (3.6) and (3.11) with the help of Monte Carlo simulations in Figures 3.3a and 3.3b respectively. The simulation parameters are listed in Table 3.1 unless stated otherwise. The simulations are repeated for 10^4 network realizations and for each realization, the SE is measured at an arbitrary UE operating in cellular or D2D mode. The SE obtained from (3.6) and (3.11) and the simulations is averaged over all the realizations and hence, the effect of link distances is also averaged out. We study the average SE per UE as a function of the offload fraction $\mu = (N-K)/N$. Clearly, the analysis for both SE_k^{BS-UE} and SE_j^{UE-UE} is in good agreement with the simulations for various transmit SNR[2] values γ_b and γ_d. In addition, we study the SE for the case when there is no noise, i.e. $v_k^{BS}, v_d^{UE} = 0$ or alternatively $\gamma_d, \gamma_b \to \infty$. In that case, the analysis in (3.6) and (3.11) reduces to

$$\lim_{\gamma_b,\gamma_d\to\infty} SE_k^{BS-UE} | r_{k0}, r_{kl} \approx \log_2\left(1 + \frac{\gamma_b/\gamma_d\,(M-K)\,r_{k0}^{-\alpha_c}}{K\sum_l^{N-K} r_{kl}^{-\alpha_d}}\right)$$

[2]The variation in transmit SNRs $\gamma_b = P_b/\sigma_{BS}^2$ and $\gamma_d = P_d/\sigma_{UE}^2$ is governed by several parameters including the BS and UE transmit powers P_b and P_d, the noise spectral density, etc. We will implicitly treat the effect of these parameters by varying γ_b and γ_d directly to assess the performance of our setup. To ensure a fair comparison, a fixed, positive ratio γ_b/γ_d is maintained.



(a) Cellular mode. **(b)** D2D mode.

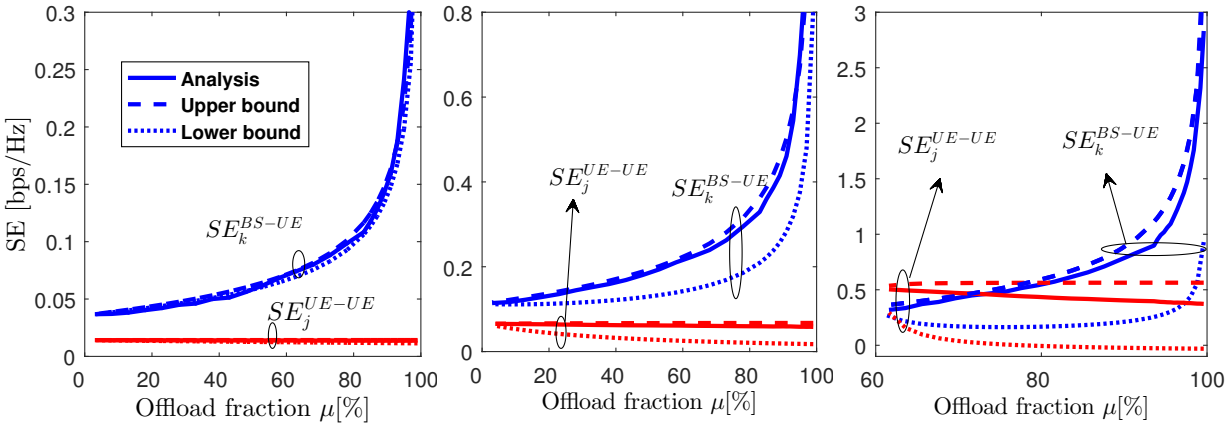
Figure 3.3.: Effect of the offload fraction on the SE of an arbitrary UE.

and

$$\lim_{\gamma_b,\gamma_d\to\infty} SE_j^{UE-UE}|r_{j0}, r_{jl} \approx \log_2\left(1 + \frac{r_{d2d}^{-\alpha_d}}{\sum_{l\neq d}^{N-K} r_{jl}^{-\alpha_d} + {\gamma_b}/{\gamma_d} r_{j0}^{-\alpha_c}}\right).$$

We observe that for low transmit SNR values γ_b and γ_d, SE_k^{BS-UE} increases monotonically, while there is a drop in SE_j^{UE-UE} with the increase in μ. We refer to this as the low-SNR (LS) regime. The rise in SE_k^{BS-UE} with the increase in μ is because as more UEs are offloaded to D2D mode, the number K of the cellular UEs inside the cell decreases and the power allocated to each cellular UE by the BS increases. The fall in SE_j^{UE-UE}, on the other hand, is due to the increasing interference from D2D UEs. This gap becomes more pronounced for higher values of γ_d. A different behavior is observed for SE_k^{BS-UE} in the high SNR (HS) regime; it closely resembles the case when $\gamma_d, \gamma_b \to \infty$, i.e. the system is interference-limited. The SE_k^{BS-UE} in HS regime is initially quite high when no UEs are offloaded. At offload percentage of 1%, the BS power is being distributed over $N-1$ UEs. Because of negligible D2D interference, a smaller allocated power is still sufficient to counter the $BS-UE$ link path loss in the HS scheme. As more UEs are offloaded, the allocated power for the cellular UE increases, but SE_k^{BS-UE} decreases steadily. This is because the increase in the BS power per UE is unable to cope with the increase in the D2D interference.

After a certain fraction of UEs has been offloaded ($\mu \sim 50\%$), SE_k^{BS-UE} begins



(a) $\gamma_b = 50$dB **(b)** $\gamma_b = 60$dB **(c)** $\gamma_b = 70$dB

Figure 3.4.: Bounds on SE_k^{BS-UE} and SE_j^{UE-UE} from Prop. 3.1 and 3.2.

to rise again. This rise is now dominated by the increase in the allocated power per cellular UE. The value of SE_k^{BS-UE} at $\mu = 99\%$ is lower compared to that at $\mu = 1\%$ because of the adverse effects of the aggregate D2D interference power. In the rest of this chapter, we will focus on the LS regime as the HS regime is more suited for a multi-cell environment, where inter-cell interference also plays a critical role. An interesting observation from Figures 3.3a and 3.3b is that while SE_k^{BS-UE} monotonically increases in the LS regime and SE_j^{UE-UE} monotonically increases, there must exist an optimal offload fraction $\mu = \mu^*$ which maximizes C_{tot}.

After validation of our analysis, we study the accuracy of the bounds derived in Prop. 3.1 and 3.2. Figure 3.4 shows that the bounds closely match SE_k^{BS-UE} and SE_j^{UE-UE} from (3.6) and (3.11) respectively. The bounds are fairly tight especially for low values of γ_b and γ_d. For high values of γ_b and γ_d, the bounds on SE_j^{UE-UE} begin to deviate significantly while the bounds on SE_k^{BS-UE} still remain tight. The upper bound is tighter compared to the lower bound for both SE_j^{UE-UE} and SE_k^{BS-UE}. For the rest of the discussion, we use the upper bounds $SE_{j,UB}^{UE-UE}$ and $SE_{k,UB}^{BS-UE}$ to analyze the overall capacity C_{tot}.

We study the behavior of C_{tot} with respect to μ in Figures 3.5a-3.5c. We also explore the impact of key design parameters on the optimal offload fraction μ^* and the corresponding C_{tot}. These parameters include, i) the number of antennas M at the BS, ii) D2D link distance r_{d2d}, and iii) the transmit SNRs γ_b and γ_d. From (3.8), we see that the SE of cellular UE SE_k^{BS-UE} increases with the increase in M, while

the SE of D2D UE SE_j^{UE-UE} in (3.12) does not depend on M. As M increases, more and more UEs can be offlaoded to D2D mode as seen from Figure 3.5a. When M=N=100, it is better to offload 75% UEs in D2D mode while only 6% UEs should to be offloaded when M is increased to 400, while $N = 100$. Another important observation is that the selection of μ is crucial for smaller M. We can see that when $M = N = 100$, $C_{tot} = 2$ bps/Hz for $\mu = 3\%$, whereas $C_{tot} = 10$ bps/Hz for $\mu = 75\%$ giving 5 times better performance.

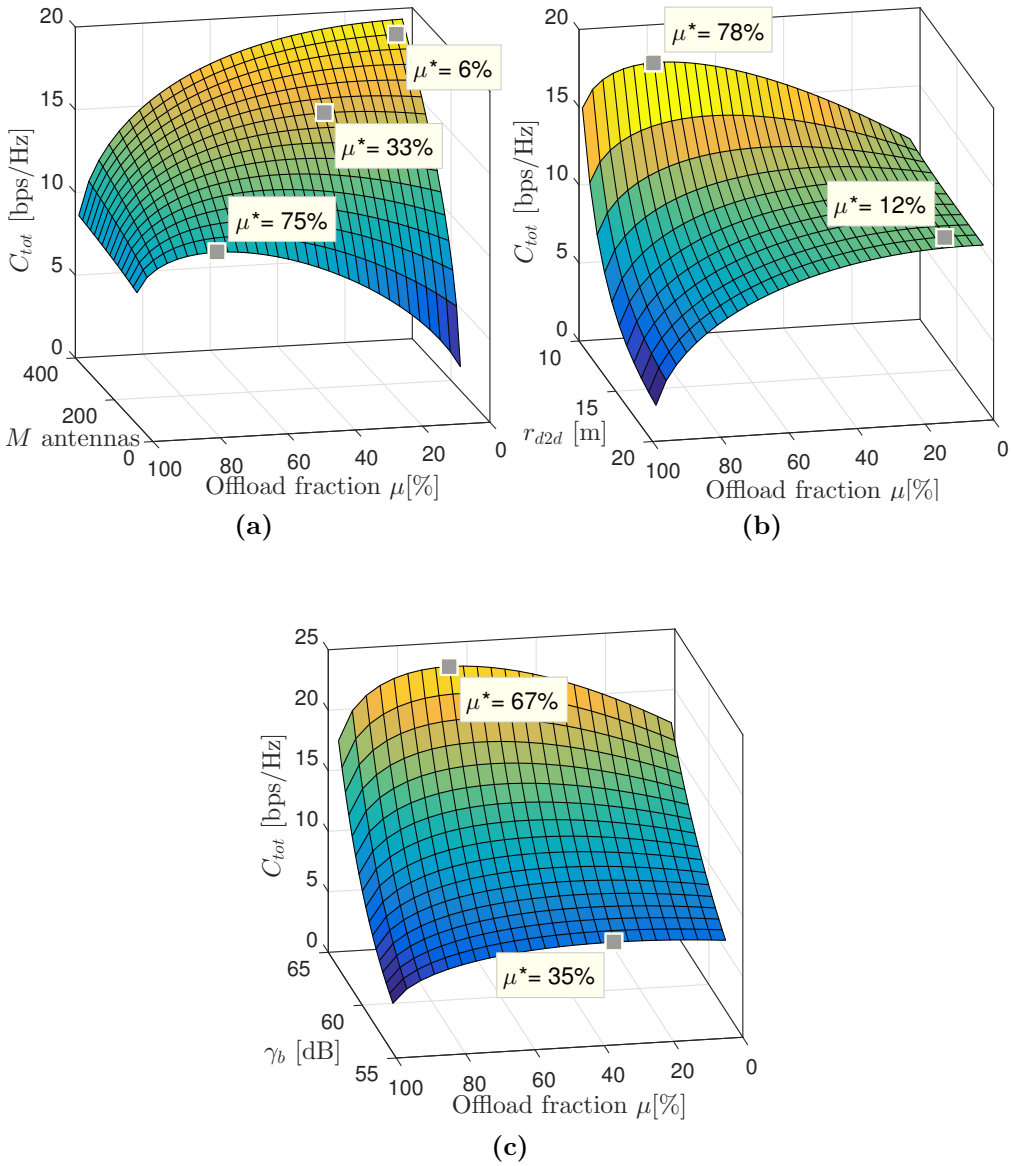


Figure 3.5.: Effect of the number of antennas M, D2D link distance r_{d2d} and γ_d on C_{tot} and $\mu^* = {}^{(N-K^*)}/_{N}$: $\gamma_b = 60$dB.

Figure 3.5b shows the effect of D2D link distance r_{d2d} on C_{tot} and μ^*. The increase in r_{d2d} aggravates D2D link path loss and degrades SE_j^{UE-UE}, while SE_k^{BS-UE} is independent of r_{d2d}. We see that a high overall capacity C_{tot} can be achieved with smaller values of r_{d2d} and it is better to offload UEs in D2D mode if their respective D2D transmitter is located close by. We further notice that even a slight increase of a few meters in r_{d2d} significantly reduces gains in C_{tot} from offloading, thereby causing μ^* to drop. As a consequence, the BS has to carefully evaluate the offloading strategy based on the D2D link distances before scheduling UEs for transmission.

We also study the effect of γ_b and γ_d in Figure 3.5c. We observe from (3.8) and (3.12) that as γ_b increases while γ_b/γ_d is fixed, both SE_k^{BS-UE} and SE_j^{UE-UE} increase causing C_{tot} to rise. The increase in SE_j^{UE-UE}, however, is more that the increase in SE_k^{BS-UE} as evident from Figures 3.3b and 3.4. This implies that with higher SNR, more UEs should be offloaded to D2D mode to maximize C_{tot}. We see that for a 10dB rise in γ_b and γ_d, up to 30% more UEs can be offloaded to maximize C_{tot}.

3.5. Conclusion

In this chapter we studied the performance gains achieved by employing network-assisted D2D communication in a massive MIMO system. We consider a finite number M of antennas at the BS, where M is larger than the number of UEs in the cell. We derived closed-form expressions for the spectral efficiency of an arbitrary UE in the cell in both D2D and cellular modes. Our results reveal that there exist a trade off between the number of UEs a BS can offload and the maximum achievable capacity of the cell. With careful selection of the offload fraction, the overall capacity of the cell can be improved up to a factor of five.

4 Offloading to D2D in a Multi-cell Scenario with Realistic Path loss Model

In this chapter, we employ the third generation partnership project (3GPP) recommended path loss models for the analysis of cellular networks overlaid with D2D communication and channel inversion power control in the uplink. The 3GPP model accounts for the presence of line-of-sight (LoS), non-line-of-sight (NLoS) and free space (FS) links. We develop tractable expressions for the coverage and rate in cellular and D2D modes using tools from stochastic geometry. Our theoretical work differs significantly from the traditional methodology using simple power law path loss models. We demonstrate that such classification of links significantly impacts the inference which can be derived from the analysis for the design of overlaid D2D networks. In particular, we show that, contrary to the previous findings, the area spectral efficiency (ASE) of the network does not saturate with the increase in the density of base stations (BS), but there exists an optimal mode selection threshold and BS density, which maximizes the ASE.

4.1. Introduction

Currently available literature on D2D communication in the uplink (UL) focuses on the analysis of spectrum sharing, interference mitigation, power control and mode selection techniques [8,22,27]. However, the analytical findings and even simulations in these works consider simplistic and idealized power law path loss model, where the signal power attenuates according to $r^{-\alpha}$, where r is the distance separation and α is the path loss exponent. In reality however, α depends on r and may increase

with r. Considering a single path loss exponent for all distance separation may greatly impact the computation of average received power and interference powers. Furthermore, the presence of line-of-sight (LoS) and non-line-of-sight (NLoS) links is also neglected in the simple path loss model. It is well established that transmissions from the user equipment (UE) face a lot of obstructions as the distance to the intended receiver gets large because of the low antenna heights of the UEs. This effect is worsened in urban environments where D2D communication is most applicable. Therefore, it is necessary for a path loss model to distinguish between cellular and D2D links.

Recent studies on the analysis of LoS and NLoS communication focus only on single tier downlink cellular networks [47–50]. In this chapter, we build upon the network model discussed in [27] for the overlaid D2D communication in the cellular UL with channel inversion power control by employing practical path loss models recommended by 3GPP for the transmissions from the UE to the BS in cellular mode [51] and the transmissions from the UE to UE in D2D mode [52].

4.1.1. Contributions

The contributions of this work are as follows:

- We borrow tools from stochastic geometry to fully characterize and obtain closed-form expressions for the average transmission power and coverage in cellular and D2D modes under the realistic 3GPP propagation model.

- We observe that for a given noise floor, the cellular coverage in the baseline model in [27] remains constant with the variation in BS density. But, our enhanced model based on the 3GPP standards indicates otherwise and shows that the cellular coverage in fact decreases with an increase in the BS density.

- The area spectral efficiency (ASE) of the network using the reference model saturates after a certain BS density threshold and increasing the BS density after that does not have any effect. On the contrary, our analysis with the 3GPP path loss model shows that there exists an optimal BS density, which maximizes the ASE of the network.

4.1.2. Organization

The remaining chapter is organized as follows. Section 4.2 outlines the hybrid network setting. Section 4.3 discusses the preliminary analysis, which includes the derivation of the expected power in cellular and D2D modes. Section 4.4 provides the main results of cellular and D2D coverage. Section 4.5 verifies the analysis of coverage with network simulations and discusses useful insights. Section 4.6 concludes the chapter.

4.2. System Model

We consider a UL scenario of a single tier cellular network overlaid with D2D communication and channel inversion power control. In this section, we briefly outline the important device, link and network level parameters, which dictate the network performance.

4.2.1. Spatial Model and Mode Selection

The BSs with intensity λ_b are placed inside a regular hexagonal grid. The transmitting UEs are distributed in space according to a homogeneous Poisson point process (HPPP) $\Phi_u \in \mathbb{R}^2$ with intensity λ_u. We assume that only a fraction ε of the UEs can participate in D2D communication. The intended receiver of each D2D enabled user is placed at distance L from the user, where L is a Rayleigh distributed RV with probability density function (PDF) $f_L(x) = 2\pi\zeta x \exp(-\zeta\pi x^2)$, where ζ is an arbitrary shape parameter. It is further assumed that the D2D enabled UE communicates in D2D mode only if the distance L is below a certain threshold r_{d2d}, otherwise cellular mode is selected. The probability of D2D mode selection is then given as

$$\mathbb{P}[L \leq r_{d2d}] = 1 - \exp(-\zeta\pi r_{d2d}^2).$$

Because of the independent thinning property of the HPPPs [32], the UEs operating in D2D mode constitute an HPPP $\Phi_d \in \mathbb{R}^2$ with intensity $\lambda_d = \varepsilon\lambda_u\,(1 - \exp(-\zeta\pi r_{d2d}^2))$ and the cellular UEs constitute a HPPP $\Phi_c \in \mathbb{R}^2$ with intensity $\lambda_c = (1-\varepsilon)\lambda_u \exp(-\zeta\pi r_{d2d}^2)$.

It is assumed that the cellular UEs are associated with the nearest BS. Notice

that the transmitting D2D UE and its intended receiver may not be present in the same cell due to the ad hoc nature of the D2D network. Without any loss of generality, the performance in cellular mode is measured at a typical BS in cellular mode and a typical D2D receiver in D2D mode. For the sake of analytical tractability, we exploit the stationarity property of HPPP. Therefore, in cellular mode, the typical node is assumed to be at the origin. A similar process can be repeated to position a typical D2D receiver at the origin by translating the PPP of the D2D receivers. Figure 4.1 displays the network spatial model under consideration.

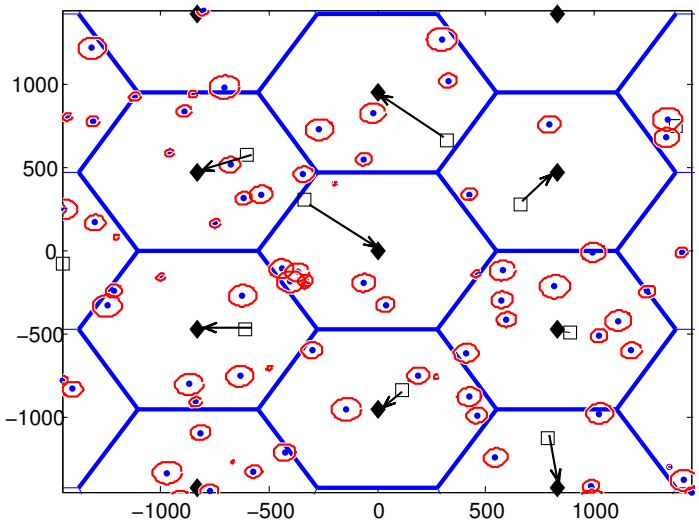


Figure 4.1.: Network model. Diamonds represent BSs placed in the center of a regular hexagonal grid. The cellular UEs operating on a single channel are shown by squares. All UEs operating in D2D mode are shown by dots. The circle centered at each D2D UE represents its distance to the intended receiver. The D2D receiver lies anywhere on the perimeter of this circle.

4.2.2. Propagation Model and Power Control

We consider that the radio signal experiences small-scale flat fading, which is complemented by the attenuation due to the large scale path loss. We assume a Rayleigh fading environment, where the channel power $h(\mathbf{x}_1, \mathbf{x}_2)$ between arbitrary locations $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^2$ is an i.i.d unit-mean exponential RV. Because of the i.i.d channel gains, we denote $h(\mathbf{x}_1, \mathbf{x}_2) = h$ in the rest of the analysis. We adopt the path loss model

specified by the 3GPP standard in [52]. Consequently, the path loss model for the UE located at a distance x from its corresponding BS[1] is given as

$$l_c(x) = \begin{cases} \underset{\text{with probability } Pr_c^{LOS}(x)}{A_{c,l}x^{-\alpha_{c,l}}} & \\ \underset{\text{with probability } 1-Pr_c^{LOS}(x),}{A_{c,n}x^{-\alpha_{c,n}}} & 0 \leq x \leq r_c, \\ \\ A_{c,n}x^{-\alpha_{c,n}} & x > r_c, \end{cases} \tag{4.1}$$

where $A_{c,l}$ and $\alpha_{c,l}$ are the cellular LoS reference path loss and path loss exponents respectively, $A_{c,n}$ and $\alpha_{c,n}$ are the cellular NLoS reference path loss and path loss exponents respectively, r_c is a constant based on practical measurements and $Pr_c^{LOS}(x)$ is the probability of having a LoS link of the transmitting UE with the BS at distance x. It is given as [49]

$$Pr_c^{LOS}(x) = \begin{cases} 1 - \frac{x}{r_c} & 0 \leq x \leq r_c, \\ 0 & \text{otherwise.} \end{cases} \tag{4.2}$$

The path loss model for the UE-UE link is slightly more involved and is given as

$$l_d(x) = \begin{cases} A_{fs}x^{-\alpha_{fs}}, & 0 \leq x \leq r_{fs}, \\ A_{d,l}x^{-\alpha_{d,l}}, & r_{fs} < x \leq r_D, \\ \\ \underset{\text{with probability } Pr_d^{LOS}(x)}{A_{d,l}x^{-\alpha_{d,l}}} & \\ \underset{\text{with probability } 1-Pr_d^{LOS}(x),}{A_{d,n}x^{-\alpha_{d,n}}} & x > r_D. \end{cases} \tag{4.3}$$

where A_{fs}, $A_{d,l}$ and $A_{d,n}$ are the free space, D2D LoS and D2D NLoS reference path losses respectively; α_{fs}, $\alpha_{d,l}$ and $\alpha_{d,n}$ are the free space, D2D LoS and D2D NLoS path loss exponents respectively; r_D is a constant based on measurements; $r_{fs} = q\, f_c$ is the free space distance which depends on the carrier frequency f_c, where $q = 2.56/c$ m/Hz is a constant depending on UE's antenna heights and the speed of light c; $Pr_d^{LOS}(x)$ is the probability of having a LoS link between the transmitting UE and

[1]Here we assume $x = \|\mathbf{x}_1 - \mathbf{x}_2\|$, $\mathbf{x}_2 = o$, $\mathbf{x}_1 \in \Phi_u$.

the UE at a distance x. It is given as

$$Pr_d^{LOS}(x) = \begin{cases} 1, & r_{fs} \leq x \leq r_D, \\ \frac{r_D}{x}\left(1 - \exp\left(-\frac{r_D}{x}\right)\right)\exp\left(-\frac{r_D}{x}\right), & x > r_D. \end{cases} \tag{4.4}$$

We use the following non-linear approximation to simplify the expression for $Pr_d^{LOS}(x)$ in (4.4):

$$Pr_d^{LOS}(x) \approx \begin{cases} 1, & r_{fs} \leq x \leq r_d, \\ \frac{r_d}{x}. & x > r_d, \end{cases} \tag{4.5}$$

where $r_d = r_D + v$ and v is a small displacement term.

4.2.2.1. Channel Inversion Power Control

The power received at a distance x in cellular or D2D mode (in the absence of noise) can be quantified as $P_r^{(i)} = P_i l_i(x)h,\ i = \{c, d\}$, where P_i is the UE transmit power in mode i. We adopt uplink channel inversion power control, where a transmitting UE inverts the path loss to serve the intended receiver. This implies

$$P_i = \rho_i l_i^{-1}(x),\ i = \{c, d\}. \tag{4.6}$$

Here, ρ_i is the sensitivity of the receiver in mode i. The small scale fading gain is not included in power control as it has little effect on the long term statistics and it removes the need to estimate h at every transmission slot. Furthermore, we assume that the links suffer from both co-channel interference and additive white Gaussian noise at receiver front end.

4.2.3. Spectrum Access Model

We assume that the available spectrum is divided between D2D and cellular networks. Thus, the D2D transmitters operate in an overlay mode in a disjoint spectrum partition. This enables network operator to suppress the inter-tier interference without sophisticated coordination mechanism. A fraction β of the bandwidth is allocated to the D2D UEs, while the remaining $1 - \beta$ is allocated to the cellular UEs. For cellular communication, there is no intra cell interference, i.e. only one UE is transmitting on a given channel in a cell at the particular time.

4.3. Transmit Power Analysis

Quantification of the average transmit power of UEs in cellular and D2D modes is central for further performance analysis. More specifically, both coverage and attainable rates are coupled with the average transmit power, which shapes the received signal strength and co-channel interference. To this end, we first derive the expected transmit power of the UEs in cellular mode.

Lemma 4.1. *The average power of a UE in cellular mode with channel inversion power control and 3GPP path loss model for UE-BS link is given as*

$$\mathbb{E}[P_c] \approx \left[\frac{\rho_c'}{A_{c,l}}\left(\frac{y^{a_{2,l}}}{a_{2,l}}-\frac{y^{a_{3,l}}}{a_{3,l}}\right)+\frac{\rho_c'}{A_{c,n}}\left(\frac{y^{a_{3,n}}}{a_{3,n}}-\frac{y^{a_{2,n}}}{a_{2,n}}\right)\right]+\frac{2\rho_c\left(A_{c,n}\right)^{-1}}{a_{2,n}\left(\pi\lambda_b\right)^{\frac{\alpha_{c,n}}{2}}} \quad (4.7)$$

where $\rho_c' = 2\pi\lambda_b\rho_c$, $y = \min(r_c, R)$ $a_{2,j} = (\alpha_{c,j}+2)$ *and* $a_{3,j} = (\alpha_{c,j}+3)$, $j = \{l, n\}$.

Proof. For tractability, we approximate the hexagonal cell with a circular cell of same area $1/\lambda_b$. The radius of the cells is then given as $R = (\pi\lambda_b)^{-\frac{1}{2}}$. Taking expectation of (4.6) over the distance gives $\mathbb{E}[P_c] = \int_0^R \rho_c l_c^{-1}(x)\, f_X(x)\, dx$, where $f_X(x)$ is the distribution of the distance of the UE from its BS. Since the tagged user is uniformly distributed in πR^2, $f_X(x) = \frac{2x}{R^2} = 2\pi\lambda_b x$. We get

$$\mathbb{E}[P_c] = 2\pi\lambda_b\rho_c\left[A_{c,l}^{-1}\int_0^y(1-\frac{x}{r_c})x^{\alpha_{c,l}+1}dx+\frac{A_{c,n}^{-1}}{r_c}\int_0^y x^{\alpha_{c,n}+2}dx+A_{c,n}^{-1}\int_y^R x^{\alpha_{c,n}+1}dx\right].$$

Solving the above integrals results in the expression in (4.7). □

The following Lemma gives the expected transmit power of the D2D UEs.

Lemma 4.2. *The average power of a UE in the D2D mode with channel inversion power control and 3GPP path loss model for the UE-UE link is given as*

$$\begin{aligned}\mathbb{E}[P_d] = K\Bigg[&\frac{\omega\left(y_{fs},0,b_1(\alpha_{fs})\right)}{A_{fs}z^{b_1(\alpha_{fs})}}+\frac{\omega\left(y_d,y_{fs},b_1(\alpha_{d,l})\right)}{A_{d,l}z^{b_1(\alpha_{d,l})}}+\frac{r_d\,\omega\left(r_{d2d},y_d,b_2(\alpha_{d,l})\right)}{A_{d,l}z^{b_2(\alpha_{d,l})}}\\&-\frac{r_d\,\omega\left(r_{d2d},y_d,b_2(\alpha_{d,n})\right)}{A_{d,n}z^{b_2(\alpha_{d,n})}}+\frac{\omega\left(r_{d2d},y_d,b_1(\alpha_{d,n})\right)}{A_{d,n}z^{b_1(\alpha_{d,n})}}\Bigg],\end{aligned} \quad (4.8)$$

where $y_{fs} = \min(r_{fs}, r_{d2d})$, $y_d = \min(r_d, r_{d2d})$, $z = \pi\zeta$, $K = z\rho_d/1-\exp(-zr_{d2d}^2)$, $b_1(a) = 1+a/2$, $b_2(a) = (1+a)/2$, $\omega(x_1,x_2,b) = \Gamma(zx_1^2,b)-\Gamma(zx_2^2,b)$ *and* $\Gamma(x,a) = \int_0^x t^{a-1}\exp(-t)\,dt$, *is the lower incomplete Gamma function.*

Proof. The proof is along similar lines as that for Lemma 4.1. The expected D2D transmit power can be represented as

$$\mathbb{E}[P_d] = \int_0^{r_{d2d}} \rho_d l_d^{-1}(x)\, f_{L|L<r_{d2d}}(x)\, dx, \tag{4.9}$$

where $f_{L|L<r_{d2d}}(x) = f_L(x)/1-\exp(-\zeta\pi r_{d2d}^2)$. Substituting (4.3) and (4.5) into (4.9) and evaluating the piece-wise integral, we obtain $\mathbb{E}[P_d]$ in (4.8). □

4.4. Analysis of Coverage and ASE

The SINR at the intended receiver is characterized as $SINR_i = \frac{\rho_i h}{I_i+\sigma^2}, i = \{c, d\}$, where σ^2 is the noise power and I_i is the interference power at the receiver. Due to the exponentially distributed channel power h, the probability that the SINR is greater than a certain modulation dependent threshold is expressed as

$$\mathcal{S}_i(\theta_i) \;=\; \mathbb{P}\left[SINR_i \geq \theta_i\right] = \exp\left(-s_i\sigma^2\right)\mathcal{L}_{I_i}\left(s_i\right), \tag{4.10}$$

where $s_i = \frac{\theta_i}{\rho_i}$ with $i = \{c, d\}$. $\mathcal{L}_{I_i}\left(s_i\right)$ is the Laplace transform of the interference. It is evident from (4.10) that in order to fully characterize the cellular and D2D coverage probabilities, we need to obtain expressions for $\mathcal{L}_{I_c}\left(s_c\right)$, and $\mathcal{L}_{I_d}\left(s_d\right)$. The following theorem gives the Laplace transform of interference in the cellular mode.

Theorem 4.1. *The Laplace transform of interference on the BS from the cellular UEs outside the cell using the 3GPP path loss model for UE-BS link and channel inversion power control is given as*

$$\begin{aligned}\mathcal{L}_{I_c}(s_c) \;=\; \exp\Bigg(-2\pi\lambda_b\Bigg[&\frac{\alpha_{c,n}^{-1}y^{2-\alpha_{c,n}}}{k_{c,n}(1-\delta_{c,n})}\xi_1\left(\alpha_{c,n}, k_{c,n}, y\right) + \frac{1}{3r_c}\left[y^3\chi_3(y) - R^3\chi_3(R)\right]\\ &+\frac{1}{2}\left[y^2\xi_2\left(\alpha_{c,l}, k_{c,l}, y\right) - R^2\xi_2\left(\alpha_{c,l}, k_{c,l}, R\right)\right]\Bigg]\Bigg),\end{aligned} \tag{4.11}$$

where $y = \max\left(r_c, R\right)$, $\delta_{c,n} = 2/\alpha_{c,n}$, $k_{c,j} = \left(s_c\mathbb{E}\left[P_c\right]A_{c,j}\rho_c\right)^{-1}, j = \{l, n\}$, $\xi_1\left(a, k, x\right) = {}_2F_1\left(1, 2/a; 1-2/a; -\left(kx^a\right)^{-1}\right)$, $\xi_2\left(a, k, x\right) = {}_2F_1\left(1, 2/a; 1+2/a; -kx^a\right)$ *and*

$$\chi_3(y) = \xi_3\left(\alpha_{c,l}, k_{c,l}, y\right) + \xi_3\left(\alpha_{c,n}, k_{c,n}, y\right),$$

where $\xi_3\left(a, k, x\right) = {}_2F_1\left(1, 3/a; 1+3/a; -kx^a\right)$ *and* ${}_2F_1\left(a, b; c; x\right)$ *is the generalized*

hypergeometric function.

Proof. The active interfering cellular users constitute a HPPP $\Phi_{c,a}$ with intensity λ_b as only one interfering user is present in a cell. The interference in this case is characterized as $I_c = \sum_{\mathbf{x}_m \in \Phi_{c,a} \setminus \mathbf{o}} P_{c_m} h_m l_c \left(\|\mathbf{x}_m\| \right)$. This implies

$$\begin{aligned} \mathcal{L}_{I_c}(s_c) &= \exp\left(-s_c \sum_{\mathbf{x}_m \in \Phi_{c,a} \setminus \mathbf{o}} P_{c_m} h_m l_c \left(\|\mathbf{x}_m\| \right) \right) \\ &\overset{(a)}{\approx} \exp\left(-2\pi\lambda_b \int_R^{\infty} \frac{x}{1 + \left(s_c \mathbb{E}\left[P_c\right] l_c(x) \right)^{-1}} dx \right), \end{aligned}$$

where (a) follows from the probability generating functional (PGFL) of PPP [32] and employing Jensen's inequality for the expectation of power and averaging with respect to the channel power. The lower limit of integration is the minimum separation distance between the typical BS and the nearest interfering user. Substituting (4.1), (4.2), and (4.7) and evaluating the piece-wise integral gives the Laplace transform in (4.11). □

Corollary 4.1. *For the realistic case of $R > r_c$, $y = R$ in (4.11) then $\mathcal{L}_{I_C}(s_c)$ reduces to*

$$\mathcal{L}_{I_c}(s_c) = \exp\left(-\frac{\delta_{c,n} \left(\pi\lambda_b\right)^{1/\delta_{c,n}} \xi_1\left(\alpha_{c,n}, k_{c,n}, \sqrt{\frac{1}{\pi\lambda_b}} \right)}{k_{c,n}\left(1 - \delta_{c,n}\right)} \right) \tag{4.12}$$

where $\delta_{c,n} = 2/\alpha_{c,n}$.

We observe from (4.12) that when $\alpha_{c,n} = \alpha_c$ and $A_{c,n} = 1$, $\mathcal{L}_{I_c}(s_c)$ reduces to the well-known expression for the Laplace transform of aggregate cellular interference in [27], which is given as

$$\mathcal{L}^{ref}_{I_c}(s_c) = \exp\left(-\frac{\delta_c \left(\pi\lambda_b\right)^{1/\delta_c}}{k_c\left(1 - \delta_c\right)} \xi_1\left(\alpha_c, k_c, \left(\pi\lambda_b\right)^{-1/2} \right) \right) \tag{4.13}$$

and $k_c = \left(s_c \mathbb{E}\left[P_c^{ref} \right] \rho_c \right)^{-1}$. The Laplace transform of aggregate interference for the D2D links is given in the following theorem.

Parameter	Value
$\lambda_b, \lambda_u, \zeta, \varepsilon, \beta$	$[1, 100, 15]/\pi 500^2, 0.5, 0.2$
$A_{c,l}, A_{c,n}$	$10^{-3.08}, 10^{-0.27}$
$A_{fs}, A_{d,l}, A_{dn}$	$10^{-3.302}, 10^{-3.08}, 10^{-0.27}$
$\alpha_c, \alpha_{c,l}, \alpha_{c,n}$	$3.5, 2.42, 4.28$
$\alpha_d, \alpha_{fs}, \alpha_{d,l}, \alpha_{d,n}$	$4, 2.27, 4, 4.375$
$r_c, r_{fs}, r_d, r_{d2d}$	300m, $q(2GHz)$m, 23m, 100m
ρ_c, ρ_d, σ^2	-70dBm,-70dBm,-100dBm

Table 4.1.: List of simulation parameters. The values of path loss exponents and the reference path loss are taken from [52].

Theorem 4.2. *The Laplace transform of interference on the typical D2D receiver from other UEs transmitting in D2D mode using the 3GPP path loss model for UE-UE link and channel inversion power control is given as*

$$\begin{aligned}\mathcal{L}_{I_d}(s_d) &= exp\Bigg(-\pi\lambda_d\Bigg[r_{fs}^2\chi_2(r_{fs}) + r_d^2\xi_2(\alpha_{d,l}, k_{d,l}, r_d) \\ &\quad +2r_d^2\chi_0(r_d) + \frac{r_d^{2-\alpha_{d,n}}}{k_{d,n}}\frac{\xi_1(\alpha_{d,n}, k_{d,n}, r_d)}{(\alpha_{d,n}-2)}\Bigg]\Bigg), \end{aligned} \tag{4.14}$$

where, $k_{d,j} = \left(s_d\mathbb{E}[P_d]A_{d,j}\rho_d\right)^{-1}, j = \{fs, l, n\}$, $\chi_2(r_{fs}) = \xi_2(\alpha_{fs}, k_{fs}, r_{fs}) - \xi_2(\alpha_{d,l}, k_{d,l}, r_{fs})$, $\chi_0(r_d) = \frac{\xi_0\left(\alpha_{d,l}, k_{d,l}, r_d\right)}{k_{d,l}\left(\alpha_{d,l}-1\right)r_d^{\alpha_{d,l}}} - \frac{\xi_0\left(\alpha_{d,n}, k_{d,n}, r_d\right)}{k_{d,n}\left(\alpha_{d,n}-1\right)r_d^{\alpha_{d,n}}}$ *and*

$$\xi_0(a, k, x) = {}_2F_1\left(1, 1/a; 1-1/a; -\left(kx^a\right)^{-1}\right).$$

Proof. The proof follows similar steps to the proof of Theorem 4.1 with the exception that the interfering UEs include all active D2D UEs $\mathbf{x}_m \in \Phi_d$ and the minimum separation distance between the typical receiver and the interfering UE is zero. □

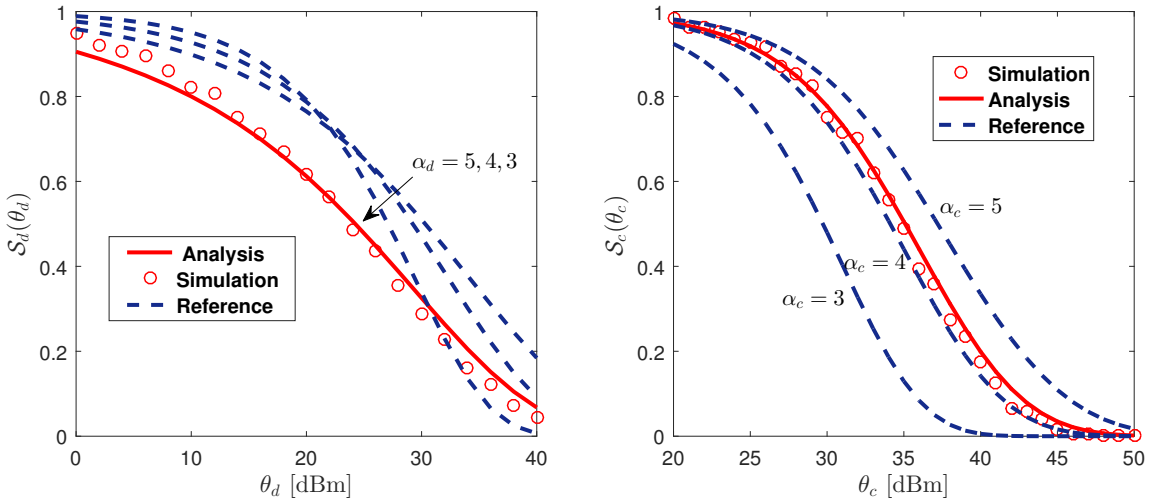
Similar to the analysis of Laplace transform of cellular interference, by setting $A_{d,j} = 1$ and $\alpha_{d,j} = \alpha_d, j = \{n, l, fs\}$ in (4.14), $\mathcal{L}_{I_d}(s_d)$ reduces to $\mathcal{L}_{I_d}^{ref}(s_d)$ in [27], where

$$\mathcal{L}_{I_d}^{ref}(s_d) = \exp\left(-\frac{\pi\lambda_d}{\mathrm{sinc}(\delta_d)}k_d^{\delta_d}\right) \tag{4.15}$$

and $k_d = s_d\,\mathbb{E}\left[P_d^{ref}\right]\rho_d$.

4.5. Results and Discussion

The first step is to validate our analysis for the D2D and cellular coverage probability using Theorem 4.1 and 4.2. We use the values listed in Table 4.1 unless stated otherwise. For the network simulations, we generate a hexagonal grid cellular network, where the area of each cell is $1/\lambda_b$. The users are distributed uniformly in each realization, where the number of users in each iteration is Poisson distributed with parameter λ_u. Figures 4.2a and 4.2b show that the simulation results for



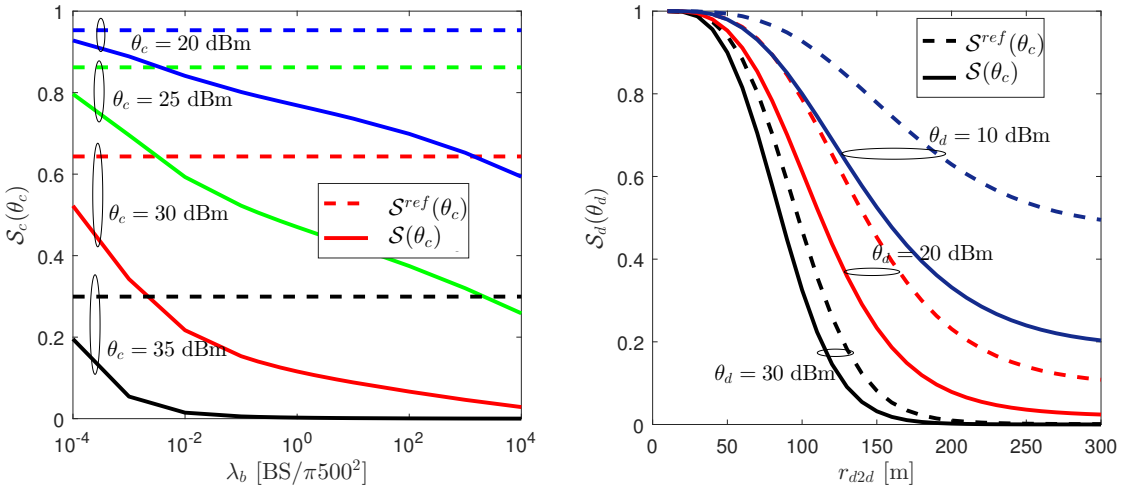
(a) D2D mode. **(b)** Cellular mode.

Figure 4.2.: Coverage probability of a generic UE in cellular and D2D modes.

D2D and cellular coverage closely match our theoretical analysis. We also compare our proposed model with the analysis in [27].

The plots reveal that the D2D coverage with the 3GPP path loss model significantly deviates from the simplistic approach in [27]. This is because of the nonlinear path loss model described in (4.5). The realistic cellular coverage in Figure 4.2b however, follows a similar trend as the cellular coverage in reference model in [27]. This is because the cellular path loss exponent remains fairly constant for all users for a small value of λ_b or alternatively a large macrocell. We can see from (4.1) that when $r_c > R$, $Pr_c^{LOS}(x) = 0$ for all the interfering links and the Laplace transform of cellular interference in (4.11) follows the same trend as (4.13).

The behavior of cellular coverage with the increasing BS density is studied with the help of Figure 4.3a. The reference cellular coverage is not affected by the change



(a) **(b)**

Figure 4.3.: Effect of variation of BS intensity λ_b on the cellular coverage probability (left) and variation of mode selection threshold r_{d2d} on the D2D coverage probability (right).

in λ_b. This is due to the channel inversion power control, as the cell size goes small, the interference power also decreases accordingly. This ideal behavior is not observed in reality with the 3GPP path loss model and we see that as λ_b grows, the chances of having LoS interference links also increases, which intensifies the interference power. Hence, the probability of cellular coverage drops with the increase in λ_b.

Figure 4.3b shows the effect of varying the mode selection threshold on the D2D coverage probability. The D2D coverage decays more steeply than the reference when the mode selection threshold is increased. This is because as r_{d2d} increases, the number of D2D transmitters increases. In the case of the proposed model, it also implies that the density of interfering UEs with free space path loss and LoS also increases. The results on the realistic D2D coverage probability may not be as optimistic as the previous findings however, we observe that with the suitable set of parameters, D2D communication still results in better coverage prospects compared to cellular communication.

4.5.1. Area Spectral Efficiency

The ASE of the network under discussion is the sum of the rates of all active links normalized over the transmission bandwidth and unit area. It is expressed as

$$\mathcal{T} = \lambda_c(1-\beta)R_c + \lambda_d\beta R_d, \text{ bps/Hz/m}^2 \tag{4.16}$$

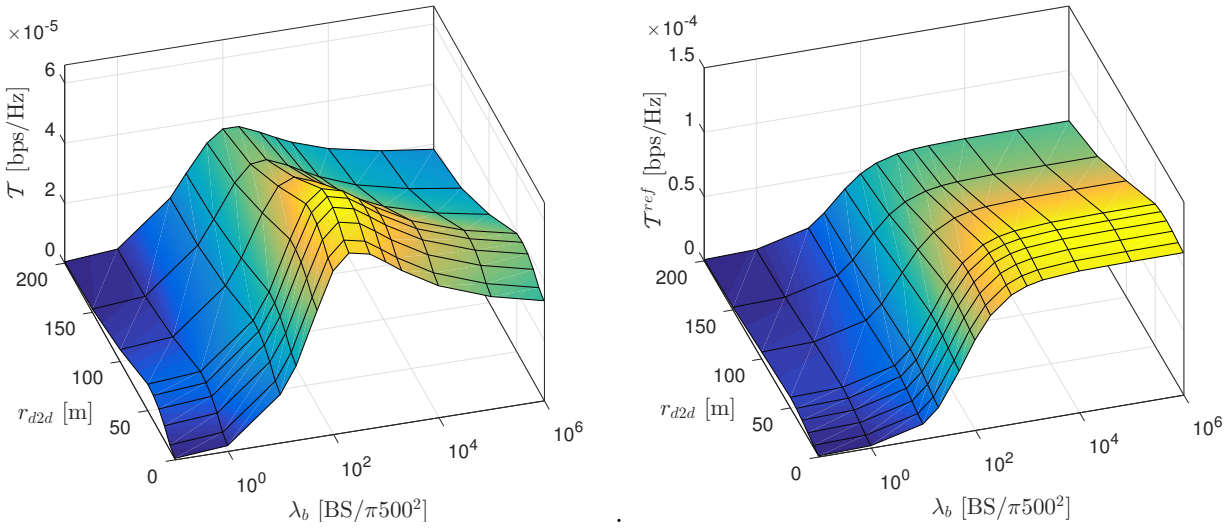
where R_c and R_d are the expected rates of the cellular and D2D links respectively and β is the spectrum resource partition factor. Using Shannon's capacity formulation, the cellular rate per unit bandwidth can easily be expressed as $R_c = \mathbb{E}\left[1/N\right] R_c^{'}$, where $R_c^{'} = \mathbb{E}\left[\log_2\left(1+SINR_c\right)\right]$ and $\mathbb{E}\left[1/N\right] = \lambda_b/\lambda_c\left(1-\exp\left(-\lambda_b/\lambda_c\right)\right)$ is the expectation taken over the number of active cellular users N inside the cell (including the tagged UE). The D2D rate is similarly expressed as $R_d = \mathbb{E}\left[\log_2\left(1+SINR_d\right)\right]$. The ASE is then given as

$$\mathcal{T} = \lambda_b(1-\beta)(1-\exp(-\lambda_c/\lambda_b))R_c^{'} + \lambda_d\beta R_d. \tag{4.17}$$

The expectation for $R_c^{'}$ and R_d are computed as

$$\begin{aligned}\mathbb{E}\left[\log_2\left(1+SINR_i\right)\right] &= \int\limits_{z>0} \mathbb{P}\left[\log_2\left(1+SINR_i\right) > z\right] dz \\ &= \int\limits_{z>0} \mathbb{P}\left[SINR_i > 2^z - 1\right] dz \\ &= \log(2) \int\limits_{z>0} (1+z)^{-1}\mathcal{S}_i(z)\, dz.\end{aligned}$$

We wish to see how the variation in λ_b and r_{d2d} impacts the ASE of our proposed model and the reference model as the ASE is a function of both the D2D and cellular coverage. Figure 4.4a shows that for the proposed coverage model, the ASE first increases with an increase in λ_b and r_{d2d} and attains a maximum value at a point (λ_b^*, r_{d2d}^*) after which it decays. The increase with respect to r_{d2d} is attributed to the fact that initially, the activation of more D2D users offloads cellular traffic and enables spatial frequency reuse. However, after a certain value of r_{d2d}, the interference due to further activation of D2D UEs becomes dominant and reduces the ASE. Recall from Figure 4.3a that the increase in λ_b results in a decrease in cellular coverage and hence $R_c^{'}$, but this decrease is initially overcome with the



(a) Existence of an optimal (r^*_{d2d}, λ^*_b) pair with 3GPP recommended path loss model.

(b) Reference model [27].

Figure 4.4.: ASE for various values of mode selection threshold and BS intensity.

increase in λ_b. But after a certain value of λ_b, the ASE beings to decrease. Any variation in r_{d2d} does not affect this optimal density λ^*_b. This is because, R_c is the only term in (4.17) which depends on λ_b and it is independent of r_{d2d}. This, however, is not the case for r_{d2d} as it appears in both terms in (4.17). The optimal point (λ^*_b, r^*_{d2d}) is obtained numerically and is equal to $(200/\pi 500^2, 40\text{m})$.

For the purpose of comparison, the ASE for the reference model is displayed in Figure 4.4. A similar trend is observed with the variation in r_{d2d} but a striking difference is seen for the variation in λ_b. This is because, due to a simplistic path loss assumption, the cellular coverage, and hence the cellular rate, does not change with respect to λ_b and as λ_b grows, (4.17) converges to $\lambda_c(1-\beta)R_c^{\prime ref} + \lambda_d \beta R_d^{ref}$ as

$$\lim_{\lambda_b \to \infty} \lambda_b(1 - \exp(-\lambda_c/\lambda_b)) = \lambda_c.$$

4.6. Conclusion

This chapter analyzed the cellular networks overlaid with D2D communication in the UL using path loss models recommended by the 3GPP and compared the coverage

and area spectral efficiency (ASE) with the baseline model in [27], which uses a simple power law path loss model and does not differentiate between the LoS, NLoS and free space regimes. The realistic path loss model significantly impacts the coverage and ASE results. A major difference is that our theoretical results confirm that as the density of the BSs grows, there is no perfect interference cancellation as suggested by the reference model in [27].

5 Offloading to D2D in a Multi-cell Scenario Based on Content Availability

In this chapter, we develop a comprehensive analytical framework for an approach similar to the EPC-level discovery defined by the 3GPP, where the BS is responsible for establishing and arbitrating D2D links based on the content availability in proximate devices. We consider two different helper selection schemes employed by the BS to establish D2D links. To model a D2D communication link, we derive the probability of the UE to be served in D2D mode and the distribution of the distance between an arbitrary user and its ith neighboring D2D helper within the macrocell. We fully characterize the overall rate coverage probability and the average rate experienced by an arbitrary user requesting a particular content. Our analysis reveals that significant performance gains can be achieved with the help of information-centric offloading compared to conventional cellular communication, especially when popular contents are requested. However, enhancing D2D opportunities for the users does not always result in better performance and the network parameters have to be carefully tuned to harness maximum gains.

5.1. Introduction

Ubiquitous devices such as smart phones and tablets have fueled the demand for data intensive applications, including ultra high-definition video streaming, social networking and e-gaming. This puts significant pressure on the traditional cellular networks, which are not designed to support such high data rates and reliability requirements. As a consequence, current research on 5G wireless networks is geared

towards developing intelligent ways of data dissemination by deviating from the traditional host centric network architecture to a more versatile information-centric architecture. Caching has already proved to be a promising way to reduce the congestion in client-server model of communication in the wired domain. However, only recently caching has been recognized as a key enabler of 5G networks in the wireless domain as it is expected to potentially reduce the backhaul access cost in terms of capacity, latency and energy consumption by turning memory into bandwidth [53]. Current statistics of wireless data traffic reveal that UE requests follow predictable patterns and a number of duplicate multimedia content requests occur from the same vicinity [54]. Therefore, UEs can leverage this trend to their advantage by accessing the information pre-downloaded by their neighboring UEs using D2D communication. Mobile UEs in close physical proximity can exchange popular files without the intervention of the base station [55]. This not only offloads the burden of duplicate transmissions from the BS, but it also provides higher rates due to short range D2D communication [56].

5.1.1. Contribution

In this chapter, we propose a new information-centric offloading mechanism, whereby the macro BS (MBS) offloads the UEs to D2D mode if its requested content is present in proximate devices. We focus on network-assisted D2D communication, where the MBS establishes, manages and arbitrates a D2D connection [19]. According to the 3GPP standard, centralized device discovery for D2D is defined as evolved packet core (EPC)-level discovery [5], where the UE locations and proximity information can be estimated at the mobility management entity of the core network through periodic updates and the GPS and location apps in smart phones. The network-assisted device discovery ensures a better QoS and a more reliable service because of the authentication and privacy mechanisms offered by the network operator. The MBS maintains a record of the previously downloaded contents by the UEs in its long-term coverage region and, based on this information, it schedules the appropriate D2D links between a UE and one of its neighboring D2D helpers. These D2D helpers can be considered as idle UEs, which are not receiving any data form the MBS in the current time slot. The D2D connection is subject to the content availability and the employed helper selection scheme. We consider two different helper selection schemes, namely, 1) nearest helper selection (NS): where the MBS selects the D2D helper closest to the UE possessing the requested content and 2)

uniform selection (US): where the MBS only checks for the content availability of a single D2D helper selected at random. Figure 5.1 displays a simple example of the scenario under consideration. The MBS examines its records for the arrived content requests and schedules possible D2D transmissions. Here, User#1 is served by its second nearest D2D helper, while User#2 is served by the MBS as none of its k neighboring helpers have the content.

The main contributions of this chapter are summarized as follows.

- For the information-centric offloading paradigm, we consider that the D2D helpers are equipped with caches and may serve requesting UEs in their close proximity. Even though we explicitly ignore the cost incurred in accessing content from the core network, the MBS still encourages D2D communication for the requesting UEs whenever feasible to reduce multiple duplicate requests and to provide better coverage at the UEs. We derive expressions for the probability that an arbitrary requesting UE is served in D2D mode by one of its k candidate D2D helpers. These expressions are based on a given content placement strategy at the D2D helper's cache, cache size, helper selection scheme (NS/US) and the requested content itself. We obtain bounds on this probability and study its behavior as the number of candidate helpers k grows.

- With the help of our stochastic geometry framework, we derive the distribution of distance between an arbitrary UE and its ith nearest D2D helper within the cell using a disk approximation for a Voronoi cell. We show that this approximation is fairly accurate for various values of i and compare it with the distribution of distance between the requesting UE and its ith nearest D2D helper not necessarily present inside the cell. We investigate the conditions in which our derived distribution reduces to the unconstrained case.

- We characterize the coverage probability for the cellular and individual D2D links. We also derive the probability that an arbitrary UE is in coverage for a certain desired rate and content under the NS and US schemes. Our analysis is validated with the help of network simulations.

- We explore the two important performance metrics, including the rate coverage probability and the average rate experienced by an arbitrary UE requesting a particular content c. We show the existence of an optimal number of candidate D2D helpers $k = k^*$ that maximizes the rate coverage and the average rate.

We also show that high performance gains (up to 9× the average throughput experienced by an arbitrary UE) can be harnessed compared to conventional cellular communication for most popular contents requests, skewed content popularity and dense deployment of D2D helpers.

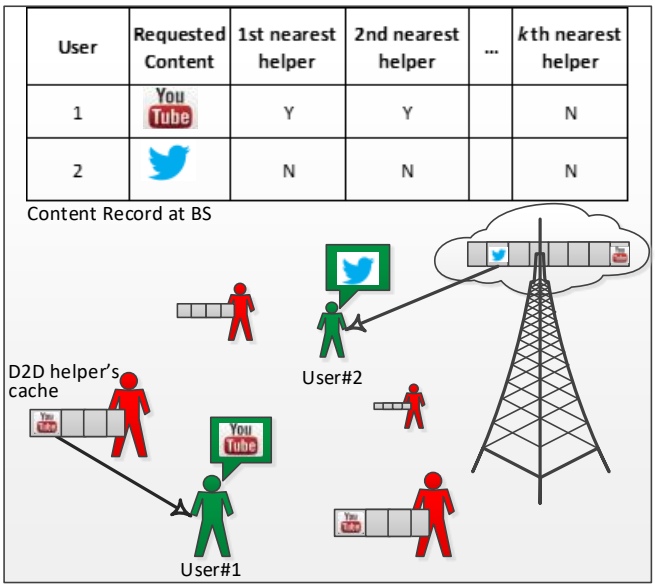


Figure 5.1.: Illustration of cache-enabled coordinated D2D network. The MBS pairs the requesting UEs with one of their k neighbors depending on the content availability and helper selection scheme. If none of the k neighbors have the content, the MBS serves the UE itself.

5.1.2. Related Work

Characterization of the performance of caching enabled cellular networks has been widely studied in [23, 57, 58] to name a few. However, all these works make use of the simplistic protocol model, where outage occurs if the intended receiver is at a distance greater than a fixed distance from the transmitter or there is another interfering transmitter present within the range of the receiver. The other approach, which makes use of the physical model, is where the outage occurs on the basis of the received signal-to-interference-and-noise ratio (SINR).

Stochastic geometry has been widely applied to analyze the physical layer metrics of the large scale wireless networks. For the case of cache enabled cellular

networks, the dynamics of content popularity, propagation conditions and spatial locations are employed in [59–66] to quantify the performance gains. The analysis of rate and energy efficiency for single-tier cache enabled cellular networks is carried out in [62] and [64]. An optimal content placement strategy is devised in [64], which maximizes the rate coverage and the energy efficiency of the single-tier cellular networks. The authors in [65] and [66] consider clustered D2D networks, which operate in isolation from the underlying cellular network. In [65], the authors consider clustering to mimic spatial content locality without explicitly considering content popularity and storage. It is assumed that for a UE in a given cluster, there will always be a device in that cluster with the requested content i.e., a D2D wireless link is always established. The authors in [66] consider D2D devices with caching and a Zipf type content popularity distribution. Here, clustering is considered so that there are finite transmissions within the cluster multiplexed in time as in TDMA and one link is active at a given time. In [63], the analysis is carried out with different D2D transmitter selection schemes, but this work makes use of the same clustered UEs model as in [65]. All these works do not take into account the coexistence of D2D communication with cellular networks and that if D2D communication is infeasible, the UEs can communicate with the MBS.

For the case of multi-tier analysis with caching, the authors in [60] and [61] consider distributed caching, where a UE can access data from the caches of multiple small BSs (SBSs) inside a cell. The UE is always served by its nearest SBS if the content is available in the cache of any one of the SBSs within the cell. The authors assume that the content transfer takes place among the SBSs. This is different from our case as we cannot expect such level of cooperation between D2D helpers and therefore need an explicit characterization of the distances of the individual D2D helpers from the arbitrary UE.

The selection of cellular and D2D modes is studied for the uplink in [27] and [22]. In [27], the decision to transmit in D2D mode is based on the distance to the receiver uniformly placed around the D2D transmitter, while in [22], it also depends on the distance from the BS. Both of these approaches ignore the aspects of content availability, popularity and storage.

Various content replacement policies and storage techniques are studied in [67–69]. Reference [68] shows how updating a cache by evicting the least recently used (LRU) content could provide performance gains. It is shown that least frequently used (LFU) policy outperforms LRU [70]. Reference [69] explores how caches could

be updated by exploiting social ties between UEs using transfer learning approach. Reference [67] proposes coded caching for delay sensitive applications. In [71] and [72], the authors explore the effect of geometric placement of caches to devise optimal content placement strategies. A simpler, fixed-caching approach is adopted in [62, 66, 73] and [61], where the cache is not updated and the stored files are simply considered to follow the popularity distribution.

5.1.3. Organization

The remainder of this chapter is organized as follows. Section 5.2 describes the spatial setup, signal propagation, content popularity and caching models. The information-centric offloading paradigm for both the NS and US schemes is explained in section 5.3. Section 5.4 provides the derivation of the distance between an arbitrary UE and its ith nearest D2D helper within the cell. The distribution of this distance is then used to characterize the overall coverage and the average rate for the NS and US schemes in Section 5.5. Section 5.6 discusses the results and validates our analysis with network simulations. Section 5.7 concludes the chapter.

5.2. System Model

We consider a cellular DL scenario of MBSs overlaid with D2D helpers. The MBS schedules a requesting UE with one of its neighboring D2D helpers inside the cell if the helper has the requested file. The network description and the key assumptions now follow.

5.2.1. Spatial Model

According to the theory of HPPPs, the number of points $\Phi(\mathcal{A})$ of the point process Φ inside a bounded Borel set $\mathcal{A}$, in $\mathbb{R}^2$, is given as [35]

$$\mathbb{P}[\Phi(\mathcal{A}) = n] = \frac{(\lambda\mu(\mathcal{A}))^n}{n!}\exp\left(-\lambda\mu(\mathcal{A})\right), \tag{5.1}$$

where λ is the intensity of the HPPP and $\mu(\mathcal{A}) = \int_{\mathcal{A}} dx$ is the Lebesgue measure on $\mathbb{R}^2$. For a disc of radius r in $\mathbb{R}^2$, $\mu(\mathcal{A}) = \pi r^2$. We consider that the MBSs, D2D helpers and the requesting UEs are distributed according to independent HPPPs Φ_m, Φ_d, and Φ_u with intensities λ_m, λ_d and λ_u respectively, where $\lambda_u, \lambda_d \gg \lambda_m$.

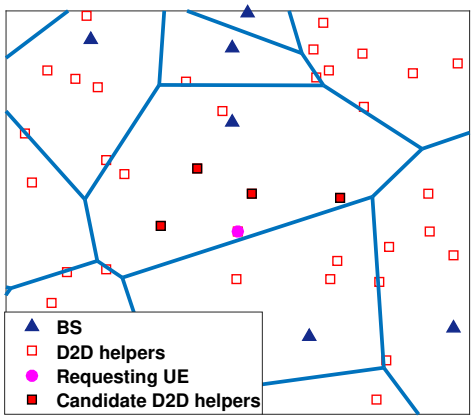


Figure 5.2.: Spatial model of the network: $k = 4$ candidate D2D helpers, $\lambda_d = 10\lambda_m$.

The requesting UEs are associated to the nearest MBS and the MBS association region is defined as [31]

$$\mathcal{C}_{v_i} \overset{def}{=} \{x \in \mathbb{R}^2 \, : \, \|\nu_i - x\| < \|\nu_j - x\|, \forall \, \nu_i, \nu_j \in \Phi_m, i \neq j\}, \tag{5.2}$$

where $\mathcal{C}_{v_i}$ represents a Voronoi cell of the MBS located at $\nu_i \in \Phi_m$[1]. Without any loss of generality, we measure performance at the requesting UE located at the origin. This follows from the palm distribution of HPPPs and Slivnyak's theorem [35]. The MBS selects one of the k closest D2D helpers inside its coverage region and establishes a D2D link. A realization of the spatial setup is shown in Figure 5.2. We note that the actual physically nearest D2D helper may lie in some other macrocell, but the D2D pairs are constrained to exist within the same macrocell.

5.2.2. Propagation Model and Spectrum Access

We assume that both the cellular and D2D links experience channel impairments including path loss and small-scale Rayleigh fading. The power received at the origin from the MBS/ D2D helper located at $z \in \Phi_n, n = \{m, d\}$ is given as $P_n h \, \|z\|^{-\alpha}$, where P_m and P_d are the transmit powers of the MBS and D2D helper respectively, α represents the path loss exponent usually greater than 2 and h is the channel power. We assume that h is a unit-mean exponential RV representing the squared-envelope

[1]We use the same notation to denote the the node itself and its location.

of Rayleigh fading.

We consider inband overlaid D2D communication, where the available downlink spectrum W is partitioned into orthogonal sub-bands W_m and W_d for cellular and D2D communication respectively. The main motivation behind this overlaid D2D communication is that no sophisticated interference management has to be employed. For multiple access, we assume that MBS employs TDMA scheduling for the UEs inside the cell. Therefore, at a given instant, there is at most one cellular and one D2D UE inside the cell.

5.2.3. Content Popularity and Caching Model

The performance of caching is crucially determined by the content popularity distribution. It has been observed that the popularity of data follows a Zipf popularity distribution where, the popularity of the cth content is proportional to the inverse of c^{ζ} for some real, positive, skewness parameter ζ. It is mathematically represented as [70]

$$pop(c) = \rho c^{-\zeta} \quad 1 \leq c \leq L, \tag{5.3}$$

where $\rho = \left(\sum_{l=1}^{L} l^{-\zeta}\right)^{-1}$ is the distribution normalizing factor and L is the file library size. $\zeta = 0$ corresponds to uniform popularity while a higher value of ζ results in a more skewed distribution. Empirical evidence shows that the value for ζ normally range between 0.6 to 0.8 for different content types including web, file sharing, UE generated content (UGC) and video on demand (VoD) [70]. We consider that the D2D helpers are equipped with caches of size C_d. All files are considered to have a unit size. Our analysis can easily be extended for variable file sizes as each memory slot will then contain a chunk of a file. We further assume that UE requests follow the independent reference model (IRM) as introduced in [70]. According to the IRM, the UE requests for a file in the library are independently generated following the popularity distribution and there is no spatio-temporal locality, i.e. identical contents have the same popularity in space and time [60].

Content Placement

When there is a set of candidate D2D helpers that can serve a single UE, the LFU placement for all D2D helpers is not optimal. Such a scenario requires a collaborative content placement strategy which takes into account the number and the locations

of the D2D helpers [71, 72]. Investigating the optimal content placement strategy for D2D helpers for this network setup is a research issue in itself and left for future work. We consider a sub-optimal but tractable content placement strategy for the D2D helpers to quantify the advantage of employing content-centric offloading on D2D mode. We consider that each D2D helper stores the content c in each memory slot independently according to the popularity distribution $pop(c)$. The D2D hit rate for content c, is then given as

$$\begin{aligned} h_d(c) &= 1 - \mathbb{P}[\text{Content } c \text{ not present in } C_d \text{ slots}] \\ &= 1 - [1 - \rho c^{-\zeta}]^{C_d}. \end{aligned} \tag{5.4}$$

5.2.4. Mode Selection

In the proposed framework, we consider that when a UE requests content c, the MBS employs the NS or the US scheme to establish an appropriate D2D connection for this UE. If the connection is not possible then the UE is served by the MBS. The description of the NS and US schemes are describes as follows.

Nearest selection (NS)

The MBS examines the record of pre-downloaded contents as depicted in Figure 5.1. Out of the k closest candidate D2D helpers inside the cell for the requesting UE, the MBS selects the nearest D2D helper that possesses the content c for D2D connection. If none of the k nearest helpers have the content, the MBS serves the UE.

Uniform selection (US)

The MBS randomly selects a D2D helper out of the k closest candidate D2D helpers inside the cell for the requesting UE. If this helper has the requested content then the D2D connection is established. Otherwise, the UE is served by the MBS. Even though this scheme is computationally less intensive, it will be clear in the following sections, that the performance is compromised with this random helper selection.

The motivation of considering the two schemes is as follows. Even though the instantaneous received power from the nearest D2D helper may be low, the effect of path loss outweighs the effect of small scale fading as it stays constant over a longer period of time as formally proven for the case of single tier networks in Lemma

1 of [74]. The assumption of nearest D2D helper providing maximum power in a long-term sense is therefore appropriate. In contrast, the US scheme depicts the average case helper selection because the link quality is completely neglected. We now describe the technical details and the analysis of the NS and US schemes in the following section.

5.3. Analysis of Information-Centric Offloading

The selection of the D2D helper for a requesting UE depends on the helper selection scheme, the number k of candidate D2D helpers and the popularity of the requested content itself. The first step is to derive the probability that a UE is served by its ith nearest D2D UE ($1 \leq i \leq k$) under the NS and US schemes. This probability is stated in the following proposition.

Proposition 5.1. *The probability that an arbitrary UE requesting content 'c' is served by its ith nearest D2D helper within the cell under NS and US schemes is given by*

$$p_{d,i,NS}(c) = \frac{3.5^{4.5}\Gamma(i+4.5)\,\eta_d^i\,{}_2F_1\left(1, i+4.5; i+1; \frac{\eta_d}{(\eta_d+3.5)}\right)}{i!\,\Gamma(4.5)\,(\eta_d+3.5)^{i+4.5}}\,(1-h_d(c))^{i-1}\,h_d(c), \tag{5.5}$$

and

$$p^{(k)}_{d,i,US}(c) = \frac{3.5^{4.5}}{\Gamma(4.5)}\left[\frac{\Gamma(k+5.5)\,\eta_d^{k+1}\,{}_2F_1\left(1, k+5.5; k+2; \frac{\eta_d}{(\eta_d+3.5)}\right)}{k(k+1)!\,(\eta_d+3.5)^{k+5.5}} + \sum_{j=0}^{k-i}\frac{\Gamma(i+j+4.5)\eta_d^{i+j}}{(i+j)\,(i+j)!\,(\eta_d+3.5)^{i+j+4.5}}\right]h_d(c) \tag{5.6}$$

respectively, where $\eta_d = \lambda_d/\lambda_m$, $i = \{1, ..k\}$, $\Gamma(a)$ is the complete Gamma function and ${}_2F_1\,(a, b; c; x)$ is the generalized hypergeometric function.

Proof. For the UE to be served by the ith nearest D2D helper, there must be at least i D2D helpers inside the cell. In the NS scheme, the UE is served by the ith helper if no closer helper has the requested content. This implies

$$p_{d,i,NS}(c) = \mathbb{P}\left[N_d \geqslant i\right](1-h_d(c))^{i-1}\,h_d(c), \tag{5.7}$$

where N_d is the number of D2D helpers in a Voronoi cell conditioned on covering

the origin. Even though the exact statistics of a Voronoi cell are still unknown, the probability mass function (PMF) of N_d can be well-approximated as [75]

$$\mathbb{P}\left[N_d = j\right] = \frac{3.5^{4.5}\Gamma(j+4.5)\eta_d^j}{\Gamma(4.5)\, j!\, (\eta_d + 3.5)^{j+4.5}}, \tag{5.8}$$

where $\eta_d = \lambda_d/\lambda_m$. We use the equality in ((5.8)) because of the tightness of approximation already established in [75]. Moving on, we obtain $\mathbb{P}\left[N_d \geqslant i\right] = 1 - \sum_{j=0}^{i-1} \mathbb{P}\left[N_d = j\right]$ and substitute this expression into (5.7) to get (5.5). For the US scheme, the UE is served by the ith helper if it is uniformly selected and has the requested content. This implies

$$p_{d,i,US}^{(k)}(c) = h_d(c) \left[\frac{1}{k}\mathbb{P}\left[N_d > k\right] + \sum_{j=0}^{k-i} \frac{1}{i+j}\mathbb{P}\left[N_d = i+j\right]\right]. \tag{5.9}$$

Substituting the expressions for $\mathbb{P}\left[N_d > k\right]$ and $\mathbb{P}\left[N_d = i+j\right]$ in ((5.9)) completes the proof. □

Corollary 5.1. *The probability for an arbitrary UE requesting content c to be served in D2D mode under the NS and US schemes is given as*

$$p_{d,NS}^{(k)}(c) = \sum_{i=1}^{k} p_{d,i,NS}(c), \tag{5.10}$$

and

$$p_{d,US}^{(k)}(c) = p_{d,US}^{(1)}(c) = p_{d,NS}^{(1)}(c) = h_d(c)\left[1 - (1 + 3.5^{-1}\eta_d)^{-4.5}\right]. \tag{5.11}$$

Proof. The probability of D2D mode under the NS scheme is a straightforward summation of $p_{d,i,NS}^{(k)}(c)$ in (5.5) over $i = \{1, 2.., k\}$. For the US scheme, we follow the complementary rule to obtain the overall probability of D2D mode. As the D2D helper is selected first and then checked for available content, the probability of D2D mode in the US scheme is essentially equal to $h_d(c)\left[1 - \mathbb{P}\left[N_d = 0\right]\right]$, which is the probability that there is at least one helper inside the cell and the uniformly selected helper has the requested content. We therefore obtain (5.11). □

Corollary 5.2. *The bounds on $p_{d,\Pi}^{(k)}(c), \Pi \in \{NS, US\}$ with respect to η_d are given as*

$$p_{d,NS}^{(k)}(c) \leq 1 - h_d(c) - (1 - h_d(c))^{k-1} \tag{5.12}$$

and

$$p_{d,US}^{(k)}(c) \leq h_d(c) \tag{5.13}$$

where the equalities hold when $\lambda_d \gg \lambda_m$ *and* $\eta_d \to \infty$.

Proof. For the NS scheme, $\mathbb{P}[N_d \geqslant i] \to 1$ when $\eta_d \to \infty$ in (5.7), i.e. there are definitely at least i D2D helpers within the cell and (5.7) reduces to $(1 - h_d(c))^{i-1} h_d(c)$. Similarly, for the US scheme, $\mathbb{P}[N_d \geqslant k] \to 1$ and $\mathbb{P}[N_d = i + j] \to 0$, where $i+j < k$ and $\eta_d \to \infty$. To show that the asymptotes also represent upper bounds, we prove that (5.10) and (5.11) are strictly increasing in η_d using the first derivative test. We have

$$\frac{\partial}{\partial \eta_d} p_{d,NS}^{(k)}(c) = \frac{3.5^{4.5}}{\Gamma(4.5)} h_d(c) \sum_{i=1}^{k} \frac{\Gamma(i+4.5)\, \eta_d^{i-1}}{\Gamma(i)\, (\eta_d + 3.5)^{i+4.5}} (1 - h_d(c))^{i-1} \tag{5.14}$$

and

$$\frac{\partial}{\partial \eta_d} p_{d,US}^{(k)}(c) = \frac{9}{7} \left(1 + 3.5^{-1} \eta_d\right)^{-5.5} h_d(c), \tag{5.15}$$

which are both positive in η_d. □

Before moving on to further analysis, we study the behavior of the D2D mode probability in Figures 5.3a and 5.3b. Figure 5.3a shows that $p_{d,NS}^{(k)}(c)$ sharply decreases as the requested content becomes less popular or the skewness parameter ζ decreases. Figure 5.3b reveals that $p_{d,NS}^{(k)}(c)$ initially increases rapidly with the increase in k, but diminishing gains are observed when k is further increased. On the other hand, it is interesting to note that $p_{d,US}^{(k)}(c)$ is independent of k as already established in (5.11). This is because increasing the candidate D2D helpers has no impact on the likelihood of finding the requested content and the cache hit still depends on the stored contents of one randomly selected helper. In Figure 5.3c, we plot the D2D mode probabilities $p_{d,II}^{(k)}(c)$ using the upper bounds from Corollary 5.2 and compare them for the actual values of $p_{d,II}^{(k)}(c)$ in (5.10) and (5.11). The deviation of $p_{d,II}^{(k)}(c)$ from the bounds for the NS scheme becomes large as the value of k increases, however, the bounds are fairly tight for $\eta_d \geq 10$.

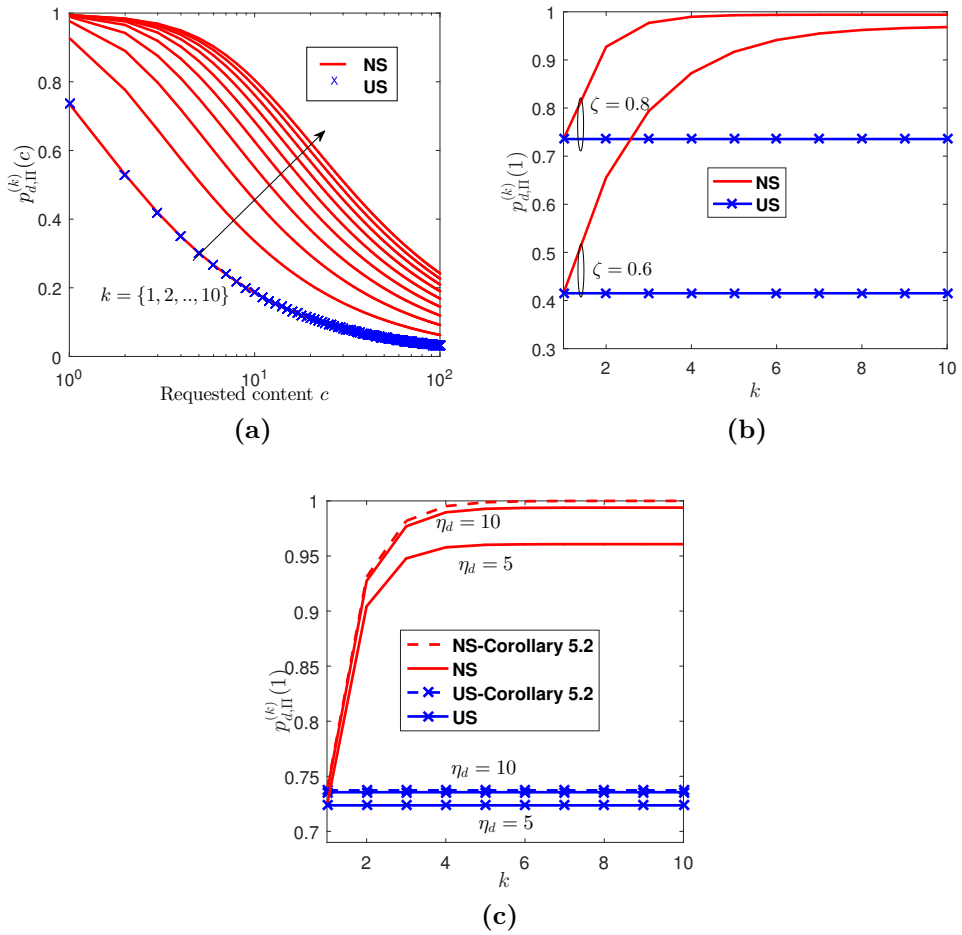


Figure 5.3.: Effect of helper selection schemes $\Pi \in \{NS, US\}$, requested content c, candidate D2D helpers k and popularity skewness parameter ζ on the probability of D2D mode.

5.4. Distance to the ith Nearest D2D Helper Within a Macrocell

One of the main contributions of this work is to characterize the distribution of the distance between the typical UE and the ith nearest D2D helper within the macro cell. It is a well-known fact that the distance between the nearest neighbors for a 2-D Poisson process is Rayleigh distributed and this has been widely adopted for the stochastic geometry analysis of cellular networks [22, 27], [31, 62]. In our case, however, the MBS only keeps a record of the files stored in the memory of D2D helpers within its coverage region. Therefore, it can only connect the requesting UE with the helpers within its cell. Figure 5.2 illustrates that in our spatial setup,

the ith nearest D2D helper is not always within the macrocell. Hence, this adds a layer of complexity to our model as the distance is no longer independent of the geometrical attributes of the cell, including its shape and size.

The distribution of the exact shape and size of a typical Voronoi cell in a 2-D Poisson Voronoi tessellation is still unknown. In their analysis of bivariate Poisson processes in [76], Foss and Zuyev make use of the maximal disk approximation for the Voronoi cell. The maximal disk, B_{max}, is the largest disk inscribing the Voronoi circle. The exact characterization of the distribution of the radius X of B_{max} is straight forward as it is the probability that there is no other BS at a distance $2x$ from the tagged BS and is expressed as $\mathbb{P}[X \geq 2x] = \exp\left(-4\lambda_m \pi x^2\right)$. This implies that

$$f_X(x) = 8\lambda_m \pi x \, \exp(-4\lambda_m \pi x^2). \tag{5.16}$$

We utilize the maximal disk approximation for the Voronoi cell to derive the distribution of the distance between the typical UE and its ith nearest D2D helper. It is important to mention that the maximal ball approximation is based on the properties of a typical Voronoi cell and not the on Voronoi cell covering the origin (typical UE) as considered in Prop. 5.1. For characterizing the distance distribution in this section, we will make use of the typical Voronoi cell assumption. As will be shown in the later sections, the impact of this assumption in negligible. The following Lemmas provide some useful preliminary results which are necessary conditions for the characterization of the distance distribution.

Lemma 5.1. *Given that a UE at a distance Y from the typical MBS lies inside B_{max}, the radius $\hat{X} = \{X : X \geq Y\}$ of the resulting B_{max} is distributed as*

$$f_{\hat{X}}(x) = \frac{8\lambda_m \pi}{p_{in}} x \left[\exp(-4\lambda_m \pi x^2) - \exp(-5\lambda_m \pi x^2)\right] \quad x > 0, \tag{5.17}$$

where $f_Y(y) = 2\lambda_m \pi y \exp(-\lambda_m \pi y^2)$ and $p_{in} = \mathbb{P}\left[X \geq Y\right]$ is the probability that the UE lies inside B_{max}. It is a constant and is given as

$$p_{in} = 1/5. \tag{5.18}$$

Proof. The CCDF of the radius $\hat{X}$ of B_{max} conditioned on having the UE at a

distance Y can be written as

$$\begin{aligned}\overline{F_{\hat{X}}}(x) &= \mathbb{P}\left[X \geq x | X \geq Y\right] \\ &= \frac{\mathbb{P}\left[X \geq x, X \geq Y\right]}{\mathbb{P}\left[X \geq Y\right]} = \frac{\mathbb{P}\left[X \geq \max(x, Y)\right]}{\mathbb{P}\left[X \geq Y\right]} \\ &= \frac{1}{\mathbb{P}\left[X \geq Y\right]} \left[\int_0^x \overline{F_X}(x) f_Y(y)\, dy + \int_x^\infty \overline{F_X}(y) f_Y(y)\, dy\right],\end{aligned}$$

where $\overline{F_X}(g) = \int_g^\infty f_X(x)\, dx = \exp\left(-4\lambda_u \pi g^2\right)$. The probability $p_{in} = \mathbb{P}\left[X \geq Y\right]$ that a UE lies inside B_{max} can be written as $\int_0^\infty \left[1 - F_X(y)\right] f_Y(y)\, dy$, where $F_X(y) = \int_0^y f_X(x)\, dx$. As we know that the cellular link distance Y is Rayleigh distributed [31], we can easily obtain (5.18). Differentiating $F_{\hat{X}}(x) = 1 - \overline{F_{\hat{X}}}(x)$ with respect to x gives (5.17). □

Lemma 5.2. *The probability that there are at least i D2D helpers inside B_{max} which also contains a UE at a distance Y is given as*

$$p_{N_d}^{(i)} = 5(1 + 4\eta_d^{-1})^{-i} - 4(1 + 5\eta_d^{-1})^{-i}. \tag{5.19}$$

Proof. Given a disk B_{max} with radius $\hat{X}$, the probability that there are at least i D2D helpers inside the disk is given as

$$\mathbb{P}\left[N_d \geq i | \hat{X}\right] \;=\; 1 - \sum_{j=0}^{i-1} \frac{\left(\lambda_d \pi x^2\right)^j}{j!} \exp\left(-\lambda_d \pi x^2\right) = 1 - \frac{\Gamma(i, \lambda_d \pi x^2)}{\Gamma(i)}.$$

Averaging over $\hat{X}$, we obtain

$$p_{N_d}^{(i)} = \int_0^\infty \left[1 - \frac{\Gamma(i, \lambda_d \pi x^2)}{\Gamma(i)}\right] f_{\hat{X}}(x)\, dx. \tag{5.20}$$

Simplification after evaluating the integral yields (5.19). □

We now give the distribution of distance in the following Theorem.

Theorem 5.1. *The distribution of the distance between the typical UE and its ith nearest D2D helper within the cell can be well approximated using the inscribed disk approximation for a Voronoi cell and is given as*

$$f_{R_i}(r) \approx \frac{f_{i,2}(r)\kappa(r) + \int_0^\infty \int_{a_1}^{a_2} f_{i,1}(r, y, x)\, f_X(x)\, dx\, f_Y(y)\, dy}{(1 + 4\eta_d^{-1})^{-i} - \frac{4}{5}(1 + 5\eta_d^{-1})^{-i}} \tag{5.21}$$

where,

$$f_{i,1}(r,y,x) = \frac{\lambda_d^i}{\Gamma(i)} \nabla(r,y,x)^{i-1} \nabla'(r,y,x) \exp(-\lambda_d \nabla(r,y,x)), \tag{5.22}$$

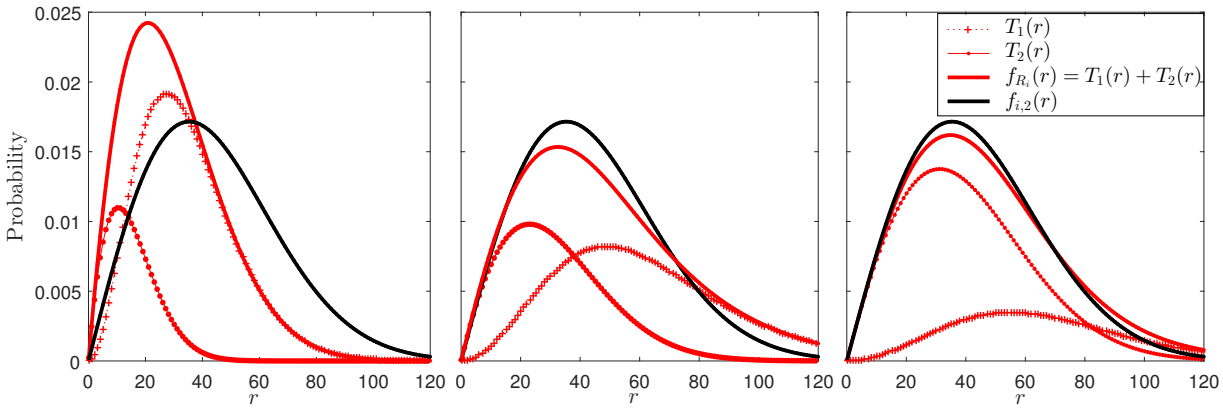
$$f_{i,2}(r) = 2\frac{(\lambda_d \pi)^i}{\Gamma(i)} r^{2i-1} \exp(-\lambda_d \pi r^2), \tag{5.23}$$

and

$$\kappa(r) \;=\; \frac{1}{5}\exp\left(-4\lambda_m \pi r^2\right) - \frac{4\sqrt{\pi}}{5}\beta r \exp\left(-4\left(\beta r\right)^2\right) \operatorname{erfc}\left(4\beta r\right), \tag{5.24}$$

where, $\nabla(r,y,x) = r^2 \arccos\left(\frac{\omega_1}{2y\,r}\right) + x^2 \arccos\left(\frac{\omega_2}{2y\,x}\right) - \frac{1}{2}\sqrt{4y^2x^2 - \omega_2^2}$, $\omega_1 = r^2 + y^2 - x^2$, $\omega_2 = x^2 + y^2 - r^2$, $\nabla'(r,y,x)$ *is the derivative of* $\nabla(r,y,x)$ *with respect to* r, $\beta = \sqrt{\frac{\lambda_m \pi}{5}}$ $a_1 = \max(y, r-y)$ *and* $a_2 = r + y$.

Proof. Please refer to Appendix A.1. □



(a) $\eta_d = 1$ **(b)** $\eta_d = 10$ **(c)** $\eta_d = 100$

Figure 5.4.: Effect of varying η_d on $T_1(r)$ and $T_2(r)$: $i = 1$.

The expression in (5.21) is validated with network simulations in Section 5.6. Before further analysis, we develop some insights on the derived distance distribution in (5.21). We can write $f_{R_i}(r) = T_1(r) + T_2(r)$, where $T_1(r) = \frac{\int_0^\infty \int_{a_1}^{a_2} f_{i,1}(r,y,x)\, f_X(x) f_Y(y)\, dx\, dy}{(1+4\eta_d^{-1})^{-i} - \frac{4}{5}(1+5\eta_d^{-1})^{-i}}$ and $T_2(r) = \frac{f_{i,2}(r)\kappa(r)}{(1+4\eta_d^{-1})^{-i} - \frac{4}{5}(1+5\eta_d^{-1})^{-i}}$. We wish to see how the density of MBSs impacts $T_1(r)$ and $T_2(r)$ and in turn $f_{R_i}(r)$.

Corollary 5.3. *For sparse networks, i.e. when* $\lambda_m \to 0 (\eta_d \to \infty)$, $f_{R_i}(r)$ *reduces to the distribution of distance to the unconstrained nearest D2D helper and is given by (5.23).*

Proof. Referring to Appendix A.1, we see that when $\lambda_m \to 0$, $x \gg r$ and $b(o,r)$ almost surely lies inside B_{max}. This in turn implies that $T_1(r) \to 0$. However, as $\lambda_m \to 0$, $T_2(r)$ reduces to $f_{i,2}(r)$ as $\kappa(r)/\left((1+4\eta_d^{-1})^{-i} - \frac{4}{5}(1+5\eta_d^{-1})^{-i}\right) \to 1$. □

Figure 5.4 reinforces the result in Cor. 5.3. We compare (5.21) with the unconstrained distribution of distance to the ith nearest neighbor $f_{i,2}(r)$ in (5.23). We see from Figure 5.4c that when the network is sparse, i.e. the cell size is large compared to the number of D2D helpers, then the term $T_2(r)$ dominates $f_{R_i}(r)$ and the distribution of the distance to the ith nearest neighbor essentially approaches that of the unconstrained case. This conclusion can be intuitively explained as we would expect that for very large cell sizes, the ith nearest D2D helper will reside in the same macrocell. However, as the η_d decreases, $T_1(r)$ begins to increase and cannot be ignored that is, when (5.21) begins to significantly deviate from (5.23) as obvious from Figure 5.4a.

5.5. Performance Analysis under NS and US schemes

To assess the performance of cellular networks enhanced with coordinated D2D communication, we define the following quality-of-service (QoS) parameters.

5.5.1. Rate Coverage Probability

The rate coverage probability is the probability that an arbitrary UE requesting content c experiences the desired rate R_{des} or higher in mode $n \in \{m, d\}$, where m and d correspond to cellular and D2D modes respectively. We formally define the overall coverage probability as follows.

Definition 5.1. The overall rate coverage probability $\mathcal{R}_{\Pi}^{(k)}(c)$ of an arbitrary UE

requesting content c under the $\Pi \in \{NS, US\}$ scheme is written as

$$\mathcal{R}_{\Pi}^{(k)}(c) = \underbrace{(1 - p_{d,\Pi}^{(k)}(c))}_{\text{Probability of cellular mode}} \cdot \overbrace{\mathcal{R}_{m,\Pi}^{(k)}}^{\text{Cellular rate coverage}} + \underbrace{p_{d,\Pi}^{(k)}(c)}_{\text{Probability of D2D mode}} \cdot \overbrace{\mathcal{R}_{d,\Pi}^{(k)}(c)}^{\text{D2D rate coverage}}, \tag{5.25}$$

where $\mathcal{R}_{n,\Pi}^{(k)}$ is the probability of coverage in mode $n \in \{m, d\}$ for a given value of k and $p_{d,\Pi}^{(k)}(c)$ and $\left(1 - p_{d,\Pi}^{(k)}(c)\right)$ are respectively the probabilities for a requesting UE to be served in D2D and cellular modes.

The description of rate coverage in cellular and D2D modes is given in the following subsections.

Cellular Mode

Using Shannon's capacity formulation, the rate coverage experienced by the UE in cellular mode can be characterized as

$$\mathcal{R}_{m,\Pi}^{(k)} = \mathbb{E}\left[\mathbb{P}\left[\frac{W_m}{N_{u_m,\Pi}} \log_2(1 + SINR_m) > R_{des}\right]\right], \tag{5.26}$$

where, $SINR_m = \frac{P_m h\, y}{\sigma^2 + I_m}$ is the singal-to-interference-and-noise ratio (SINR) experienced by the UE in cellular mode, P_m is the transmit power of the MBS, y is the distance separation between the UE and the associated MBS and $I_m = P_m \sum_{\nu_j \in \Phi_m^{act} \setminus \nu_0} g_j \, \|\nu_j\|^{-\alpha}$ is the aggregate inter-cell interference from all the active MBSs constituting Φ_m^{act} while excluding the MBS ν_0 inside the typical cell. $\|\nu_j\|$ is the distance between the typical UE and the interfering MBS ν_j and g_j is the power gain due to Rayleigh fading at the interfering link. Furthermore, W_m is the cellular bandwidth and $N_{u_m,\Pi}$ is the number of requesting UEs inside the macrocell in cellular mode under the $\Pi \in \{NS, US\}$ scheme including the typical UE. Without loss of generality, we assume the time slot of 1s. Because of TDMA operation, each UE in cellular mode gets $1/N_{u_m,\Pi}$ of the time slot for communication. Since the number of UEs inside the cell is Poisson distributed and mode selection is independent of the location of requesting UEs, the UEs in cellular mode also constitute a HPPP $\Phi_{m,\Pi}$ with intensity $\lambda_{u_m,\Pi} = \lambda_u(1 - p_{d,\Pi}^{(k)})$, where $p_{d,\Pi}^{(k)} = \sum_{c=1}^{L} p_{d,\Pi}^{(k)}(c) pop(c)$ is the probability of D2D mode in (5.10) averaged over all content requests. We adopt

the mean load approximation as described in [77] for tractability and replace the random variable $N_{u_m,\Pi}$ in (5.26) with its average value such that

$$\mathbb{E}\left[N_{u_m,\Pi}\right] \;=\; 1 + 1.28\eta_{u_m,\Pi}, \tag{5.27}$$

where $\eta_{u_m,\Pi} = {}^{\lambda_{u_m,\Pi}}/_{\lambda_m}$.

Rearranging the terms in (5.26), we get

$$\mathcal{R}^{(k)}_{m,\Pi} \;=\; \mathbb{E}\left[\mathbb{P}\left[SINR_m > \tau_m\right]\right], \tag{5.28}$$

where $\mathcal{R}^{(k)}_{m,\Pi}$ represents the SINR coverage probability for the threshold

$$\tau_m = 2^{R_{des}{}^{\mathbb{E}\left[N_{u_m,\Pi}\right]}/_{W_m}} - 1. \tag{5.29}$$

Before moving on to the analysis of rate coverage, we first derive the probability for a generic MBS to serve at least one UE inside its macrocell in the following Lemma. This result is particularly useful in analyzing the aggregate interference at the typical UE in cellular mode.

Lemma 5.3. *The probability that the MBS of an arbitrary macrocell serves at least one UE when $\Pi \in \{NS, US\}$ helper selection scheme is employed can be represented as*

$$p^{int}_m = 1 - \left(1 + 3.5^{-1}\eta_u\left(1 - p^{(k)}_{d,\Pi}\right)\right)^{-3.5}, \tag{5.30}$$

where $\eta_u = \lambda_u/\lambda_b$ and $p^{(k)}_{d,\Pi} = \sum_{c=1}^{L} pop(c)\, p^{(k)}_{d,\Pi}(c)$ is the probability of a UE in D2D mode.

Proof. For a macrocell with $N_u = j$ UEs, the probability that the MBS serves at least one out of j UEs inside the cell is equal to $1 - \mathbb{P}\left[\text{all } j \text{ UEs are in D2D mode}\right]$ because mode selection occurs independently for every UE. We therefore write

$$p^{int}_m|\left(N_u = j\right) = 1 - \left(p^{(k)}_{d,\Pi}\right)^j, \tag{5.31}$$

where $p^{(k)}_{d,\Pi}$ is obtained by averaging (5.10) and (5.11) with respect to c, which is distributed according to $pop(c)$. The PMF of the number of UEs in an arbitrary macrocell is given as $\mathbb{P}\left[N_u = j\right] = \frac{3.5^{3.5}\Gamma(j+3.5)\eta_u^j}{\Gamma(3.5)\, j!\, (\eta_u+3.5)^{j+3.5}}$. Averaging (5.31) with respect to N_u gives (5.30). □

It is interesting to note that when $p_{d,\Pi}^{(k)} = 0$ in (5.30), i.e. there is no D2D communication, $p_m^{int} = \left(1 + 3.5^{-1}\eta_u\right)^{-3.5}$, which is the traditional MBS activation probability. When the typical UE is served by the MBS in cellular mode, the probability of coverage is given by the following proposition.

Proposition 5.2. *Given that the typical UE is served in cellular mode, the rate coverage probability $\mathcal{R}_{m,\Pi}^{(k)}$ under the $\Pi \in \{NS, US\}$ scheme can be expressed as*

$$\mathcal{R}_{m,\Pi}^{(k)} = 2\pi\lambda_m \int_0^\infty \exp\left(-\pi\lambda_m \left[y^2 + 2p_m^{int}\delta\left(s_m, y, \alpha\right)\right] - \frac{s_m\sigma^2}{P_m}\right) y\, dy, \tag{5.32}$$

where, $s_m = \tau_m y^\alpha$, $\delta\left(s_m, y, \alpha\right) = {}^{s_m y^{2-\alpha}}/_{(2-\alpha)}\, {}_2F_1\left(1, \bar{\alpha}; 1+\bar{\alpha}; -s_m y^{-\alpha}\right)$ and $\bar{\alpha} = 1 - {}^2/_\alpha$.

Proof. The proof is the same as the downlink SINR coverage probability in [31], with the exception that the resulting process of active MBSs Φ_m^{act} is a HPPP thinned with probability p_m^{int} as mode selection process depends on the content availability and not on the location. The effective intensity of the active MBSs is then $\lambda_m^{int} = \lambda_m p_m^{int}$. This completes the proof. □

Corollary 5.4. *The rate coverage probability in cellular mode can be simplified for $\alpha = 4$ as*

$$\mathcal{R}_{m,\Pi}^{(k)} = \frac{\pi^{3/2}\lambda_m}{2\sqrt{\rho_2}} \exp\left(\frac{\rho_1^2}{4\rho_2}\right) \operatorname{erfc}\left(\frac{\rho_1}{2\sqrt{\rho_2}}\right), \tag{5.33}$$

where $\rho_1 = \pi\lambda_m\left[1 + p_m^{int}\sqrt{\tau_m}\tan^{-1}\left(\sqrt{\tau_m}\right)\right]$ and $\rho_2 = {}^{\tau_m\sigma^2}/_{P_m}$.

Proof. When $\alpha = 4$, $\delta_m\left(s_m, y, 4\right) = {}^{\sqrt{s_m}}/_2 \tan^{-1}\left(\sqrt{s_m}y^{-2}\right)$and (5.32) reduces to

$$\mathcal{R}_{m,\Pi}^{(k)} = 2\pi\lambda_m \int_0^\infty \exp\left(-\rho_1 y^2 - \rho_2 y^4\right) y\, dy.$$

Solving the integral gives (5.33). □

We notice from (5.33) that when $p_m^{int} \to 1$, (5.33) reduces to (14) in [31], which is the probability of coverage for the case when the UE is only served by the MBS and there is no D2D communication.

D2D Mode

Following the similar procedure, we express the rate coverage probability of a UE served by its ith nearest D2D helper in D2D mode as

$$\begin{aligned}\mathcal{R}_{d,i,\Pi}^{(k)} &= \mathbb{E}\left[\mathbb{P}\left[\frac{W_d}{N_{u_d,\Pi}}\log_2(1+SINR_{d,i}) > R_{des}\right]\right] \\ &= \mathbb{E}\left[\mathbb{P}\left[SINR_{d,i} \geq \tau_d\right]\right], \end{aligned} \tag{5.34}$$

where

$$\tau_d = 2^{R_{des}N_{u_d,\Pi}/W_d} - 1, \tag{5.35}$$

$SINR_{d,i} = \frac{P_d h\, r^{-\alpha}}{\sigma^2 + I_d}$, is the SINR experienced by the UE in D2D mode, P_d is the transmit power of the D2D helper, $R_i = r$ is the distance separation between the UE and the associated ith nearest D2D helper and $I_d = P_d \sum_{z_j \in \Phi_d^{int} \setminus z_i} g_j \, \|z_j\|^{-\alpha}$is the aggregate interference power from the other active D2D helpers, which is the sum of powers from the active D2D helpers constituting Φ_d^{int} excluding the D2D helper z_i serving the typical UE. Note that g_j is the channel power for the interfering link j. Furthermore, W_d is the bandwidth reserved for D2D communication and $N_{u_d,\Pi}$ is the number of UEs inside the macrocell in D2D mode, where according to the mean load approximation

$$\mathbb{E}\left[N_{u_d,\Pi}\right] = 1 + 1.28\eta_{u_d,\Pi} \tag{5.36}$$

and $\eta_{u_d,\Pi} = {}^{\lambda_{u_d,\Pi}}/_{\lambda_m}$ and $\lambda_{u_d,\Pi} = p_{d,\Pi}^{(k)}\lambda_u$ is the intensity of UEs in D2D mode.

We now move on to characterize the rate coverage probability experienced by a UE in D2D mode. But, before that, we present an important preliminary result in the following Lemma.

Lemma 5.4. *The probability that there is at least one UE being served in D2D mode under the $\Pi \in \{NS, US\}$ scheme can be written as*

$$p_d^{int} = 1 - \left(1 + 3.5^{-1}\eta_u p_{d,\Pi}^{(k)}\right)^{-3.5}. \tag{5.37}$$

Proof. The proof is similar to the proof of Lemma 5.3 with the exception that we now have $p_d^{int}|\left(N_u = j\right) = 1 - \left(1 - p_{d,\Pi}^{(k)}\right)^j$, which is the probability that there is at least one UE inside the macrocell operating in D2D mode. □

Proposition 5.3. *Given that the typical UE is served in D2D mode, the probability of rate coverage under the NS and US schemes can be expressed as*

$$\mathcal{R}_{d,\Pi}^{(k)}(c) = \frac{\sum_{i=1}^{k} p_{d,i,\Pi}^{(k)}(c)\mathcal{R}_{d,i,\Pi}^{(k)}}{p_{d,\Pi}^{(k)}(c)}, \tag{5.38}$$

where $\mathcal{R}_{d,i,\Pi}^{(k)}$ *is the rate coverage probability when the typical UE is served by the ith nearest D2D helper and is given as*

$$\mathcal{R}_{d,i,\Pi}^{(k)} \approx \int_{r=0}^{\infty} \exp\left(-2\pi\lambda_d^{int} s_d \mathbb{E}_Q\left[\delta(s_d, q, \alpha)\right] - \frac{s_d\sigma^2}{P_d}\right) f_{R_i}(r)\, dr, \tag{5.39}$$

where $s_d = \tau_d r^{\alpha}$, $f_Q(q) = 2\pi\lambda_d^{int} q \exp(-\lambda_d^{int}\pi q^2)$ *and* $\lambda_d^{int} = p_d^{int}\lambda_m$ *is the density of active D2D helpers.*

Proof. Please refer to Appendix A.2 □

Corollary 5.5. *For the case of* $\alpha = 4$*, the coverage probability for a UE served by its ith nearest D2D helper can be simplified as*

$$\mathcal{R}_{d,i,\Pi}^{(k)} \approx \int_{r=0}^{\infty} \exp\left(-\pi\lambda_d^{int}\sqrt{s_d}\tan^{-1}\left(4\lambda_d^{int}\sqrt{s_d}\right) - \frac{s_d\sigma^2}{P_d}\right) f_{R_i}(r)\, dr. \tag{5.40}$$

Proof. When $\alpha = 4$, $\delta\left(s_d, q, 4\right) = \sqrt{s_d}/2 \tan^{-1}\left(\sqrt{s_d}q^{-2}\right)$. To compute $\mathbb{E}_Q\left[\delta(s_d, q, 4)\right]$, we exploit the properties of the function $\tilde{\delta}\left(q\right) = \tan^{-1}\left(\sqrt{s_d}q^{-2}\right)$. This function is concave in q for the range $0 < q < q_0$ and convex for $q > q_0$. Here, $q_0 = \left(\frac{s_d}{3}\right)^{1/4}$ is the point of inflection for $\tilde{\delta}\left(q\right)$. To simplify the analysis, we make use of Jensen's inequality to shift the expectation inside $\tilde{\delta}\left(q\right)$ to obtain an approximation. We therefore get $\mathbb{E}_Q\left[\delta(s_d, q, 4)\right] \approx \sqrt{s_d}/2 \tan^{-1}\left(\sqrt{s_d}\mathbb{E}_Q\left[q\right]^{-2}\right)$, where $\mathbb{E}_Q\left[q\right] = 4\lambda_d^{int}$. Substituting this expression in (5.39) completes the proof. □

5.5.2. Average Rate

Using a similar exposition as in the previous subsection, we define the average rate as follows.

Definition 5.2. The average rate $\mathcal{T}_{\Pi}^{(k)}(c)$ experienced by an arbitrary UE requesting

content c under the NS and US schemes can be expressed as

$$\mathcal{T}_{\Pi}^{(k)}(c) = \underbrace{(1-p_{d,\Pi}^{(k)}(c))}_{\text{Probability of cellular mode}} \cdot \overbrace{\overline{\mathcal{T}_{m,\Pi}^{(k)}(c)}}^{\text{Average cellular rate}} + \underbrace{p_{d,\Pi}^{(k)}(c)}_{\text{Probability of D2D mode}} \cdot \overbrace{\overline{\mathcal{T}_{d,\Pi}^{(k)}(c)}}^{\text{Average D2D rate}} \text{ bps}, \tag{5.41}$$

where $p_{d,\Pi}^{(k)}(c)$ and $\left(1-p_{d,\Pi}^{(k)}(c)\right)$ are the probabilities for D2D and cellular communication respecitvely and $\overline{\mathcal{T}_{n,\Pi}^{(k)}}(c)$ is the average rate experienced by the UE in mode $n \in \{m,d\}$, which is written as

$$\overline{\mathcal{T}_{n,\Pi}^{(k)}}(c) = W_n \mathbb{E}\left[\frac{\log_2(1+SINR_n)}{N_{u_n,\Pi}}\right].$$

The average rate in cellular mode is characterized in the following Proposition.

Proposition 5.4. *The average rate experienced by an arbitrary UE requesting content 'c' in cellular mode under the $\Pi = \{NS, US\}$ scheme is expressed as*

$$\overline{\mathcal{T}_{m,\Pi}^{(k)}}(c) = W_m \left(1+1.28\eta_{u_m,\Pi}\right)^{-1} \mathcal{T}_{m,\Pi}^{(k)} \tag{5.42}$$

where

$$\mathcal{T}_{m,\Pi}^{(k)}(c) = \mathbb{E}\left[\log_2(1+SINR_m)\right] \tag{5.43}$$

is the average link spectral efficiency (LSE) measured in bps/Hz of the UE served in cellular mode.

Proof. Since the number of UEs in cellular mode does not affect the SINR of an individual UE, we can write

$$\overline{\mathcal{T}_{m,\Pi}^{(k)}}(c) = W_m \mathbb{E}\left[N_{u_m,\Pi}^{-1}\right] \mathcal{T}_{m,\Pi}^{(k)}, \tag{5.44}$$

where $\mathcal{T}_{m,\Pi}^{(k)} = \int_0^\infty \log(2)^{-1}(1+z)^{-1}\,\mathbb{P}\left[SINR_m \geq z\right] dz$ is the LSE of the UE, We use the mean load approximation $\mathbb{E}\left[N_{u_m,\Pi}\right]^{-1} \approx \mathbb{E}\left[N_{u_m,\Pi}^{-1}\right]$ to represent the average fraction of time slot available for a UE for cellular communication. Substituting these expressions in (5.44) results in (5.42) . □

The following proposition presents the average rate in D2D mode.

Parameter	Description	Value
α	Path loss exponent	4
$\lambda_m, \lambda_d, \lambda_u$	MBS, D2D helper and UE density	$[10, 100, 200]/\pi 500^2$
ζ	Popularity skewness parameter	0.8
c, L	Requested content, Library size	1, 10^3
C_d	D2D cache size	20
W_m, W_d	Cellular and D2D bandwidth	[7, 3] MHz
P_m, P_d	Cellular and D2D transmit power	[30, 23] dBm
R_{des}	Desired UE data rate	[0.1, 1] Mbps
σ^2	Noise power	-110 dBm

Table 5.1.: List of simulation parameters.

Proposition 5.5. *Given that the typical UE is served in D2D mode, the average rate experienced by the UE requesting content 'c' under the $\Pi = \{NS, US\}$ scheme can be expressed as*

$$\overline{\mathcal{T}_{d,\Pi}^{(k)}}(c) = W_d \left(1 + 1.28\eta_{u_{d,\Pi}}\right)^{-1} \mathcal{T}_{d,\Pi}^{(k)}(c), \tag{5.45}$$

where

$$\mathcal{T}_{d,\Pi}^{(k)}(c) = \frac{\sum_{i=1}^{k} p_{d,i,\Pi}^{(k)}(c) \mathcal{T}_{d,i,\Pi}^{(k)}}{p_{d,\Pi}^{(k)}(c)}, \tag{5.46}$$

is the average LSE of a UE in D2D mode requesting content c and $\mathcal{T}_{d,i,\Pi}^{(k)} = \mathbb{E}\left[\log_2(1 + SINR_{d,i})\right]$.

Proof. The proof follows the same procedure as that of Prop. 5.4, The average D2D rate can be written as

$$\overline{\mathcal{T}_{d,\Pi}^{(k)}}(c) = W_d \mathbb{E}\left[N_{u_d,\Pi}^{-1}\right] \frac{\sum_{i=1}^{k} p_{d,i,\Pi}^{(k)}(c) \mathcal{T}_{d,i,\Pi}^{(k)}}{p_{d,\Pi}^{(k)}(c)},$$

where

$$\mathcal{T}_{d,i,\Pi}^{(k)} = \int_0^\infty \log(2)^{-1} (1+z)^{-1} \mathbb{P}\left[SINR_{d,i} \geq z\right] dz$$

is the LSE of a UE in D2D mode served by its ith nearest D2D helper, $p_{d,i,\Pi}^{(k)}(c)$ is the probability of being served by the ith helper and denominator term $p_{d,\Pi}^{(k)}(c)$ is the probability that the UE is served in D2D mode. Employing the mean D2D load approximation $\mathbb{E}\left[N_{u_d,\Pi}\right]^{-1}$ completes the proof. □

5.6. Results and Discussion

In this section, we will give some key results and verify our analysis with Monte Carlo simulations. The values of the simulation parameters used in plotting the results are listed in Table 5.1 unless stated otherwise. For our simulation setup, the MBSs, D2D helpers and UEs are distributed according to HPPPs with intensities λ_m, λ_d and λ_u respectively in a circular area of radius 1 km. The performance is measured at the typical UE at the origin.

5.6.1. Validation with Monte Carlo simulations

We first validate the distribution of distance to the ith nearest D2D helper derived in Theorem 5.1 for various values of i. For the simulations, we ignore the realizations in which the number of D2D helpers is less than i in the typical cell. In case of the disk approximation, the realizations in which the typical UE lies outside B_{max}, or there are less than i D2D helpers inside B_{max}, are all ignored. Figure 5.5 shows that the disk approximation accurately captures the distance distribution for all values of i, while the unconstrained nearest neighbor distribution in (5.23) does not encapsulate the behavior of the distance distribution and the deviations from the actual distribution become large as the value of i increases.

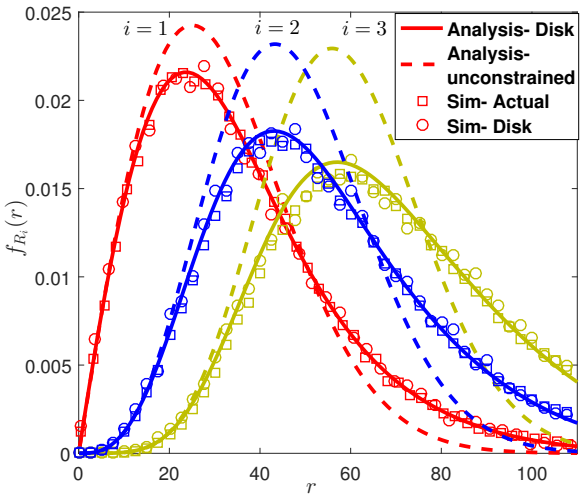
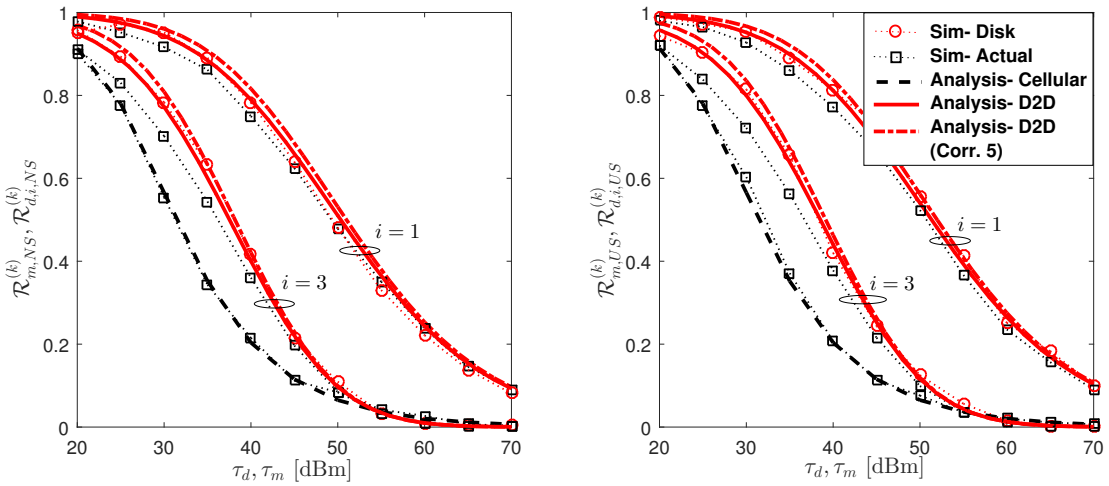


Figure 5.5.: Distribution of the distance to the ith nearest D2D helper from the tagged UE within the Voronoi cell, where $\lambda_m = 20/\pi 500^2$, and $\lambda_d = 200/\pi 500^2$.

In Figure 5.6, we validate our analysis for the link SINR coverage probability

in cellular and D2D modes by plotting $\mathcal{R}^{(k)}_{m,\Pi}$ and $\mathcal{R}^{(k)}_{d,i,\Pi}$ from (5.33) and (5.39) when τ_m and τ_d take arbitrary values. It is evident from the figure that the cellular coverage $\mathcal{R}^{(k)}_{m,\Pi}$ accurately matches the simulations. For the case of D2D links, we observe that the analysis using disk approximation is in good agreement with the simulations of the actual coverage for various values of i. The small difference between the analysis and simulations using the disk approximation is due to the equi-dense HPPP approximation for the D2D interferers used in the analysis. Expectedly, we observe that $\mathcal{R}^{(k)}_{d,i,\Pi}$ drops with the increase in i for all values of τ_d. This is because, as i increases, the distance between the transmitting D2D helper and the typical UE increases, thereby aggravating the path loss. This suggests that selecting a farther helper may result in poorer coverage compared to the cellular coverage and therefore, helper selection plays a crucial role in determining the performance of the system. It may not be obvious at first, but it is interesting to note that $\mathcal{R}^{(k)}_{d,i,US}$ is slightly higher than $\mathcal{R}^{(k)}_{d,i,NS}$ for any given τ_d. This is because, when the desired link for both the NS and US schemes is fixed to the ith nearest helper, the only difference arises in the interference characteristics pertaining to the two schemes. We see from (5.11) that the probability of an arbitrary UE to operate in D2D mode is capped at $p^{(k)}_{d,US}(c) = p^{(1)}_{d,NS}(c)$. This implies that the fraction of active D2D helpers p^{int}_d in (5.37) and in turn the density of active D2D helpers λ^{int}_d will always be small for the US scheme compared to the NS scheme for any $k > 1$. Hence, the link coverage improves due to reduced interference in the US scheme. We also plot the simplified rate coverage expression in (5.40) presented in Corr. 5.5 and observe only a small deviation compared to the analysis in (5.39).

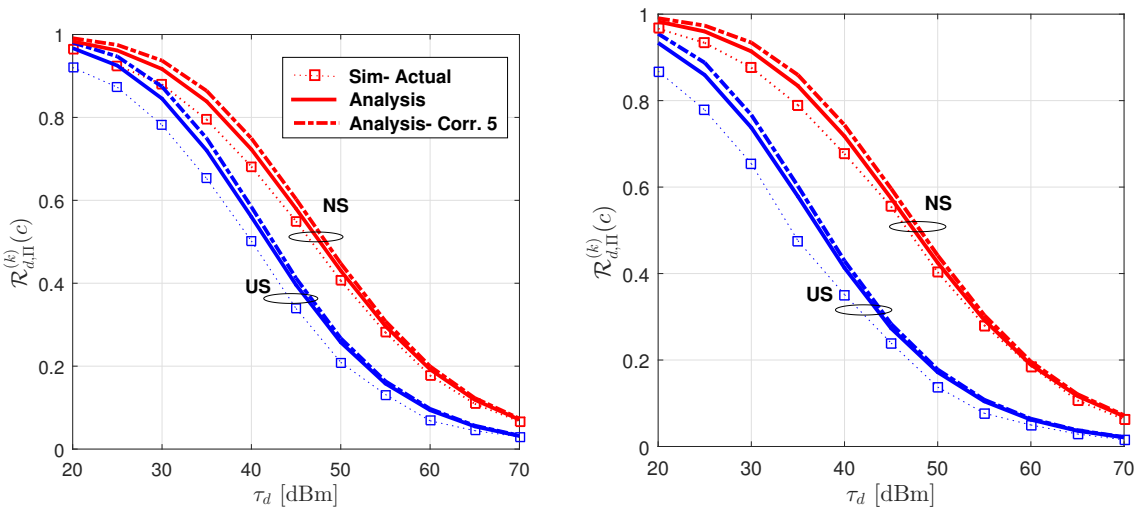
Figure 5.7 validates the overall coverage probability for a UE in D2D mode $\mathcal{R}^{(k)}_{d,\Pi}(c), \Pi \in \{NS, US\}$ given in (5.38) for arbitrary values of τ_d. For each simulation trial, k closest D2D helpers are first checked for content availability. The content availability ($c = 1$ in this case) is a Bernoulli event with probability $h_d(c)$. Out of the candidate D2D helpers (if there are any), the helper is selected either uniformly (US scheme) or closest to the typical UE (NS scheme). We observe that even though the individual link coverage is better in the US scheme compared to the NS scheme, the overall coverage for the US scheme is poor compared to the NS scheme. While $\mathcal{R}^{(k)}_{d,NS}(c)$ is robust to changes in k, $\mathcal{R}^{(k)}_{d,US}(c)$ drops significantly with the increase in k. This is because, the probability of helper selection $p_{d,i,NS}(c)$ under the NS scheme is independent of k and the closest helper is always given priority whereas, all helpers are treated equally in the US scheme.



(a) NS scheme **(b)** US scheme

Figure 5.6.: Probability of coverage when a typical UE is served by the MBS or the ith nearest D2D helper: $k = 4$.

The final step is to verify the average LSE of the typical UE in cellular and D2D modes in (5.43) and (5.46). Figure 5.8 illustrates that the analytical results match closely with the actual results for various values of k. We observe that the NS scheme outperforms the US scheme for all values of k and also results in an LSE gain of up to 2-4 times that of the LSE in cellular mode. This gain is due to the better coverage provided by the proximate D2D helpers compared with the BS. We note that $\mathcal{T}^{(k)}_{d,NS}(c)$ is more robust to the changes in k as compared to $\mathcal{T}^{(k)}_{d,US}(c)$ which decays rapidly with the increase in k. This is because the closest helpers in the NS scheme are given the same preference even when k increases as $p_{d,i,NS}(c)$ is independent of k. On the other hand, the probability of selecting the nearest helper $p^{(k)}_{d,1,US}(c)$ decreases steadily with the increase in k. We also observe that $\mathcal{T}^{(k)}_{m,US}(c)$ is not affected by the increase in k. This is because $p^{(k)}_{d,US} = p^{(1)}_{d,NS}$ in (5.30) and the resulting density of active MBSs remains unchanged. However, there is only a slight increase (not visible in the figure) in $\mathcal{T}^{(k)}_{m,NS}(c)$ with the increase in k because $p^{(k)}_{d,NS}$ rises with k causing p^{int}_m to drop, thereby improving the cellular SINR under the NS scheme.



(a) $k = 4$ **(b)** $k = 8$

Figure 5.7.: Coverage probability in D2D mode under the NS and US schemes: $c = 1$.

5.6.2. Performance Evaluation

After validating our analysis for the SINR coverage and the average LSE, we study the performance metrics defined in Sec. 5.5 with respect to the the key parameters to obtain useful design insights. For the rest of the discussion, we will focus only on the NS scheme as it has already been shown to outperform the US scheme.

5.6.2.1. Rate coverage probability

We begin with the discussion on the overall rate coverage probability $\mathcal{R}^{(k)}_{NS}(c)$ given in (5.25) using the thresholds τ_m and τ_d from (5.29) and (5.35) respectively. Figure 5.9 shows the behavior of $\mathcal{R}^{(k)}_{NS}(c)$ with changes in the number of candidate D2D helpers k, desired rate R_{des}, requested content c and the popularity skewness parameter ζ. We compare our results with the rate coverage experienced by a UE in traditional cellular network without D2D communication, which is given by (5.32) with $p_m^{int} = 1$ and $\tau_m = 2^{R_{des}\mathbb{E}[N_u]/(W_m+W_d)} - 1$. Notice that in the cellular only scenario, all of the bandwidth $W_d + W_c$ is used for cellular communication and all the N_u UEs inside the cell contend for the cellular resources.

We see that for a small R_{des} of 100kbps, coverage of almost 98% can be achieved in the case of popular content requests with the help of D2D communication. The

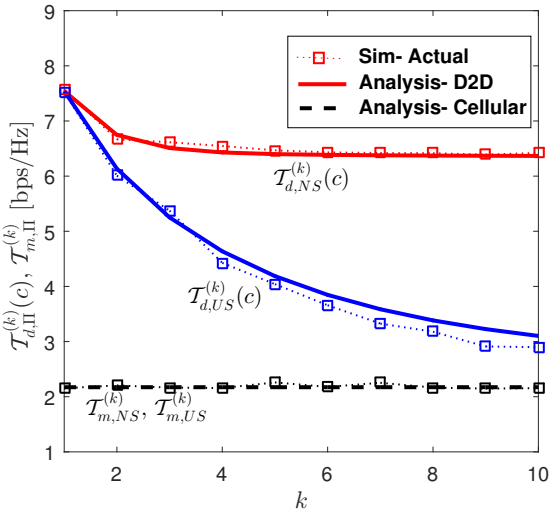
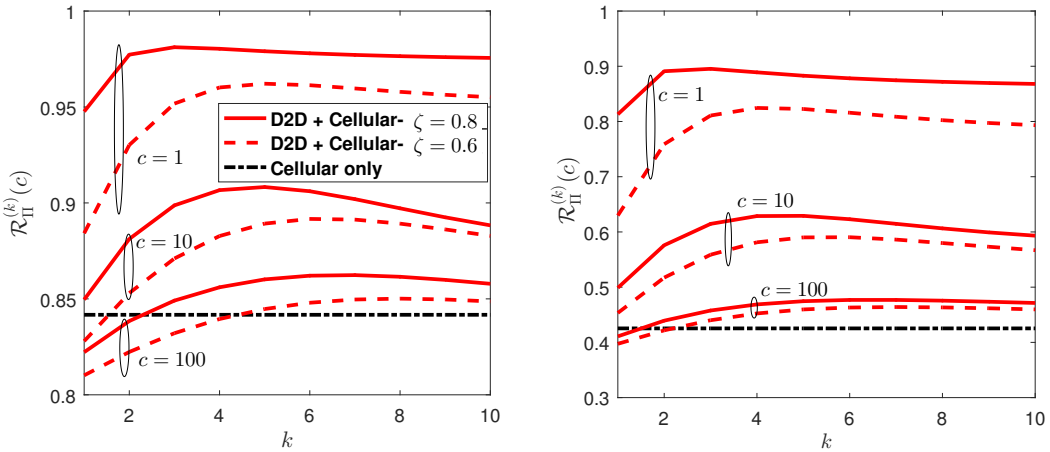


Figure 5.8.: Average link spectral efficiency experienced by the typical UE in cellular and D2D modes under the NS and US schemes.

cellular only scenario also gives a comparable performance. On the other hand, for a higher R_{des} of 1Mbps, the rate coverage of cellular only scenario drops significantly, while the drop in $\mathcal{R}^{(k)}_{NS}(c)$ is less severe, especially for popular content requests. The proposed NS scheme in this case improves the rate coverage by a factor of 2.

Another important observation is that a careful selection of the number of candidate D2D helpers k plays a crucial role in maximizing $\mathcal{R}^{(k)}_{NS}(c)$. We see that the performance in D2D mode when $k = 1$ is sub-optimal and it is required to further increase k to harness the maximum rate coverage in D2D mode. This is because it is not just the nearest D2D helper suitable for D2D communication, but even the farther helpers may provide better coverage than the MBS. As k increases, $p^{(k)}_{d,NS}(c)$ in (5.10) rises resulting in an overall increase in D2D opportunities for the UE. The plots in Figure 5.3a also indicate the existence of an optimal value of $k = k^*$ maximizing the rate coverage for a fixed set of parameters. We see that the rate coverage decreases ever so slightly when $k > k^*$. The reason for this robustness to the increase in k is that the probability of being served by the closest helper $p_{d,1,NS}(c)$ is independent of k. The slight decrease is due to assigning a smaller but finite selection probability to much farther helpers instead of switching to cellular mode. This also implies that the overall probability of a UE to be served in D2D mode $p^{(k)}_{d,NS}$ increases resulting in a higher activation probability p^{int}_d and in turn aggravated interference in D2D mode.

We also observe from the plots that the rate coverage increases as popular contents are requested or the popularity distribution is skewed. This is because, as ζ increases or c decreases, the D2D cache hit rate $h_d(c)$ increases. Hence, there is a higher likelihood to connect with the closest D2D helper because $p^{(k)}_{d,1,NS}(c)$ in (5.12) increases with $h_d(c)$. It is interesting to note that a higher $k = k^*$ is required for lower ζ or higher c to maximize the rate coverage.



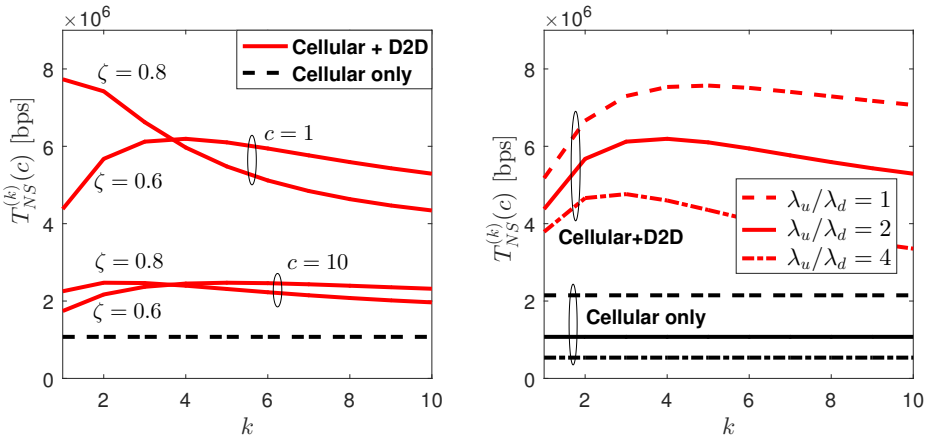
(a) $R_{des} = 100$kbps. **(b)** $R_{des} = 1$Mbps.

Figure 5.9.: Effect of increasing k on the rate coverage probability for various content requests and popularity distributions.

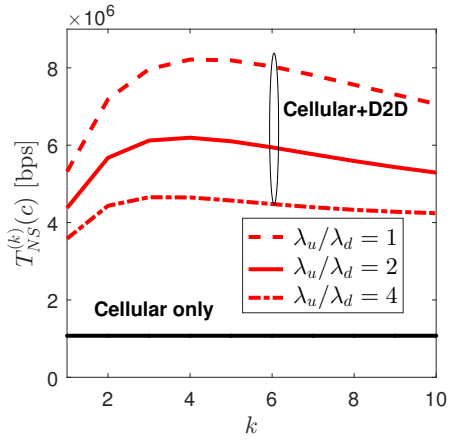
5.6.2.2. Average throughput

We will now focus on the analysis of average rate experienced by the UE in the NS scheme. Figure 5.10 displays $\mathcal{T}^{(k)}_{NS}(c)$ plotted against k. We compare $\mathcal{T}^{(k)}_{NS}(c)$ with the average rate experienced by the UE when no D2D communication is employed. The rate experienced by the UE in this case is given as $\mathcal{T}^{old}_m = (W_m + W_d)\,(1 + 1.28\eta_u)^{-1}\,\mathcal{T}_m$, where $\eta_u = \lambda_u/\lambda_b$, $\mathcal{T}_m = \mathcal{T}^{(k)}_{m,NS}$ from (5.43) with $p^{int}_m = 1$.

As a general overview of Figure 5.10, we observe that coordinated D2D communication with information-centric mode selection greatly enhances data rates compared to cellular only scenario providing more than 8× gains in the average throughput based on the set of network paramters. Yet again, there exists an obvious trade off in the selection of the number of neighboring D2D helpers k which maximizes $\mathcal{T}^{(k)}_{NS}(c)$. The existence of the optimal value of $k = k^*$ maximizing $\mathcal{T}^{(k)}_{NS}(c)$ follows a similar reasoning as the rate coverage $\mathcal{R}^{(k)}_{NS}(c)$, as with the increase in k, $p^{(k)}_{d,NS}(c)$ increases



(a) $\lambda_u = 200/A, \lambda_d = 100/A$. **(b)** $\zeta = 0.6, \lambda_d = 100/A$.



(c) $\zeta = 0.6, \lambda_u = 200/A$.

Figure 5.10.: Effect of increasing k on the average rate experienced by an arbitrary UE: $L = 10^4, A = \pi 500^2$.

causing the average number of UEs in D2D mode $\mathbb{E}[N_{u_d,\Pi}]$ to rise. However, as $k > k^*$, $\mathcal{T}_{NS}^{(k)}(c)$ begins to drop because too many UEs offloaded in D2D mode try to access the reserved bandwidth W_d. The performance is also degraded because a finite probability is assigned to associate with the kth nearest helper.

We see from Figure 5.10a that the results are extremely favorable when popular contents are requested or when the popularity distribution is skewed. For the less popular contents, the average rate experienced by the requesting UE is degraded because the cache hit probability $h_d(c)$ becomes small and more UEs are pushed to communicate in cellular mode. But this rate is still better than the case without D2D communication.

Figures 5.10b and 5.10c illustrate the effect of varying the density of UEs and D2D helpers on $\mathcal{T}_{NS}^{(k)}(c)$. It is evident that as λ_u increases keeping λ_d fixed, $\eta_{u_m,\Pi}$ in (5.27) and $\eta_{u_d,\Pi}$ in (5.36) also increase. The per UE rate $\mathcal{T}_{NS}^{(k)}(c)$ drops because the fraction of available time for communication in both cellular and D2D modes decreases. The optimal value of $k = k^*$ decreases to admit lesser UEs in D2D mode to maximize $\mathcal{T}_{NS}^{(k)}(c)$. The increase in λ_u also negatively affects the cellular only rate $\mathcal{T}_m$. Alternatively, when λ_d increases keeping λ_u fixed, the maximum value of $\mathcal{T}_{NS}^{(k)}(c)$ increases because of the smaller average distance separation between the UE and D2D helpers. An important conclusion from this figure is that the effect of a more uniform popularity distribution can be compensated by tweaking the user traffic and the number of D2D helpers.

5.7. Conclusion

We presented a novel framework for the analysis of cache-enabled cellular networks with coordinated D2D communication. The arbitrary UE requesting a particular content is offloaded to communicate with one of its k neighboring D2D helpers within the cell based on the content availability and helper selection schemes. We derived the distribution of the distance between the UE and its ith nearest D2D helper within the cell using disk cell approximation, which is shown to be fairly accurate. We obtained the probabilities for being served in cellular and D2D modes and the coverage and data rates experienced by the UE in both these modes. With the help of our analysis, we showed that the content-centric offloading with coordinated D2D can result in performance gains if highly popular contents are requested as they are more likely to be cached among the closest D2D helpers.

6 Using D2D Interfaces to Aggregate M2M Data in a Cellular Network

With billions of machine-type devices (MTDs) deployed in the near future, massive access of cellular resources from MTDs will cause congestion at the access network and disrupt the quality of experience of the existing users. There is a growin g need to devise ways to control and aggregate the generated data to ensure reliable communication while maintaining connectivity. In this chapter, we develop a statistical framework to model the aggregation of M2M data via D2D links. We consider a Poisson hard sphere model for UE coverage regions instead of the commonly used Voronoi tessellations. The main motivation of this model is twofold: i) It helps us ensure that an MTD is associated to at most one UE and, more importantly, ii) It is a realistic model for UE coverage as a UE can gather data from low-power MTDs located only in its close proximity. We analyze and compare the effectiveness of both coordinated and uncoordinated transmission strategies for the MTDs and also account for the high signaling and scheduling overhead for the coordination scheme. We explore the inherent trade-off between the time reserved for aggregation and successful trunking of data to the BS and compare our results with the baseline case where no aggregation mechanism is used. Our results reveal that while the baseline case of connecting a bulk of MTDs directly with the BS is prohibitive, M2M data can be efficiently relayed using the MAT scheme in a distributed manner.

6.1. Introduction

The large volumes of these machine-type-devices (MTDs) could potentially make up 90% of the total IoT market value [78]. One of the main concerns, however, from both the industry and academia, is that the current M2M applications are being developed as vertical silos and the *last 100m MTDs* are still mostly unconnected due to different standards and technologies [78]. Thus, there is a dire need to unify and connect these smart devices to analyze trends and extract meaningful information from the plethora of devices. Hence, the ultimate goal is to establish an interface connecting the short range MTDs with the cloud to fully realize the potential of IoTs [79].

According to a recent report on spectrum access for the internet of things by the GSM association [80], the use of licensed spectrum is vital to deliver the most reliable, secure and high quality IoT services. Because of their extensive global coverage and accessibility, cellular networks represent a viable solution to cater for massive access from the MTDs [81]. For mobile devices such as vehicles, fitness trackers and health monitors, cellular spectrum is the only choice offering low-risk, reliable communication. The existing cellular infrastructure, however, is optimized for the Quality-of-Service (QoS) requirements of human-to-human (H2H) communication, which is based on fewer and longer sessions with the main focus on providing higher data rates and better user experience. Conversely, MTDs are low power devices sending small amounts of data sporadically. Connecting a sheer bulk of MTDs with the cellular network will cause congestion at the access network. This poses a number of challenges on cellular networks and necessitates efficient resource management and clustering techniques with minimal signaling overhead [82].

6.1.1. Motivation and Related Work

A number of recent studies have proposed random access for MTDs over random access channel (RACH) in the cellular long-term evolution (LTE). In [83], the authors study various transmissions strategies over RACH and devise an optimal power allocation scheme for the MTDs. Reference [84] explores whether it is better to transmit the complete payload over RACH or only contend with the other MTDs for cellular resources depending on the payload size and packet arrival rate. These techniques however, are not scalable for ultra dense scenarios due to increased collisions on RACH. In [85], the authors survey the techniques used to reduce collision in ran-

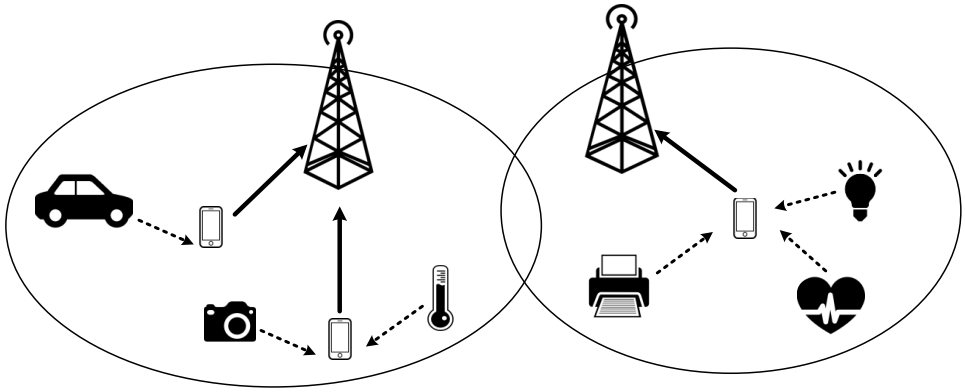


Figure 6.1.: Simplified illustration: UEs collect data from the MTDs and transmit it to the nearby BS along with their own data in uplink.

dom access. Energy efficient power allocation and cognitive underlay transmission to satisfy the QoS requirements of H2H communication has been studied respectively in [86] and [87].

Techniques like clustering of MTDs and aggregation of MTD data have proved to significantly mitigate the congestion problems [88]. Aggregation of M2M traffic with the help of machine type communication (MTC) gateways can not only reduce the burden on the base stations (BSs) of scheduling and signaling, but this multi-hop relaying of MTDs significantly reduces their transmit power and improves the overall energy efficiency [89]. Device-to-Device (D2D) communication has been identified as a promising solution to connect the MTDs with the aggregators [90]. D2D proximity service (ProSe) is an integral part of 4G and 5G networks as it enables short-range, low power communication [55]. Cellular UEs can serve as ideal candidates for D2D enabled aggregators due to their abundance and high computation capabilities. In [91], random access for UEs aggregating M2M packets is explored. The UEs contend with the other UEs on RACH after aggregating a certain number of packets. Here, no physical model is considered for the MTDs and UEs. Hierarchical data aggregation in the cellular uplink has been analyzed using stochastic geometry in [92] for half-duplex and full duplex transmission schemes. But, at every stage of aggregation, the aggregators are assumed to cover the devices distributed in the entire Euclidean space. This approximation, though valid for a BS-UE link in a cellular network, is not suitable for the UE-MTD link due to the short range and low transmit power of MTDs. A single cell, single UE framework for the aggregation and trunking of

M2M traffic via D2D links with the UEs is provided in [93]. The cellular uplink (UL) slot is divided in three slots for reservation, aggregation and trunking. The MTDs contend for UE resources in the reservation phase and the successfully reserved MTDs transmit in the aggregation phase. The UE then aggregates the M2M data and transmits it along with its own data to the BS in the trunking phase. This work in [93] is the closest to our approach, but it presents two major shortcomings, i) it does not take into account the physical locations of the MTDs and the UEs and, ii) it also does not consider the impact of interference from MTDs transmitting to other UEs in the aggregation phase.

Motivated by the above literature, we develop a large scale analytical framework for the aggregation of M2M data with the help of user equipments (UEs). The UEs collect M2M data by establishing D2D links with the nearby MTDs and then pass this information to the base station (BS) with their own data in the uplink slot.

6.1.2. Contributions

The contributions of this chapter are highlighted as follows.

- We propose a M2M data Aggregation and Trunking (MAT) scheme, where the UEs establish D2D links with the MTDs to collect M2M data and forward it along with their own data in cellular uplink (UL) radio resources as illustrated in Fig. 6.1. We consider a Poisson hard sphere model to represent the coverage regions of the UEs [94]. For the Euclidean space in $\mathbb{R}^2$, this implies that the coverage regions of the UEs are non-overlapping disks centered at the UEs. The main motivation of this model is to be able to analyze an individual arbitrarily sized cluster centered at a UE containing an arbitrary number of MTDs which can transmit to this UE. More details of this model are presented in the subsequent sections. To the best of our knowledge, this type of cluster process has not been studied extensively before in the context of cellular networks.

- For the aggregation phase, we analyze and compare two separate transmission schemes for the MTDs: 1) Multiplexed (MX), where the MTDs employ time division multiple access (TDMA) or frequency division multiple access (FDMA) and 2) Combined, where the MTDs access the channel at the same time. The transmissions for the multiplexed case do not interfere with each other, but are coordinated by the UEs and require intensive signaling and control from

the UEs. Conversely, in the case of uncoordinated combined transmissions, there is significant interference but minimal signaling involved. In this case, the UEs employ successive interference cancellation (SIC) to decode the M2M data. For the trunking phase, we assume that the UEs employ constrained uplink channel inversion power control, whereby the UE inverts the long term path loss. We obtain tight approximations of rate coverage for aggregation and trunking phases. The aggregation coverage is shown to be independent of the MTD and UE densities (λ_m and λ_u) as long as the ratio λ_m/λ_u is kept constant.

- We explore the inherent trade-off between the time reserved for aggregation by the UEs and the achieved rate coverage in both aggregation and trunking phases. In fact, a higher time for aggregation results in better aggregation opportunities but degrades trunking performance. Finally, we compare the effectiveness of MX and SIC schemes and their effect on the optimal aggregation time.

6.2. The MAT Scheme

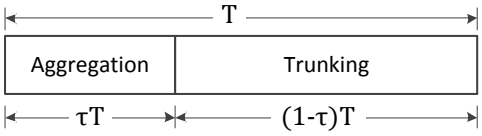


Figure 6.2.: Division of the uplink time slot.

In this work, we focus on the cellular UL scenario, where the BSs, UEs and the MTDs are distributed independently in $\mathbb{R}^2$ according to homogeneous Poisson point processes (HPPPs) Φ_b, Φ_u, and Φ_m with intensities λ_b, λ_u and λ_m respectively. The independent deployment of MTDs for dense networks is in line with practical wireless networks. We consider that the UL time slot is further divided into two slots as shown in Fig. 6.2. In the first slot, the UEs establish D2D links with the nearby MTDs to aggregate the M2M data. In the second slot, regular UL transmission takes place whereby UEs employ power control to transmit their own data as well as the collected M2M data to the BS. The UEs associate to the nearest BS, which implies

that the macrocells form a Voronoi tessellation in $\mathbb{R}^2$ [35]. After the preliminary description on the spatial setup, the key stages involved in the MAT scheme are discussed as follows.

6.2.1. Poisson Hard Sphere Model for UE coverage

The first step is to determine how a UE collects data from its nearby MTDs and how we can ensure that each MTD is associated to at most one UE. According to the Poisson hard sphere (PHS) model in 2-d, the interiors of the disks centered at the points of the parent process do not overlap almost surely (a.s.). These models have been extensively used in chemical and physical sciences to account for the interaction between particles [95]. More specifically, the nearest neighbor model (NNM) proposed by Stienen in [96] finds its use in studying disk packing and percolation [94]. In the case of NNM, the diameter of the disk centered at a given particle is the distance to its nearest neighbor of the same process. To understand this better in a wireless networks sense, consider a UE z_j, where $z_j \in \Phi_u$[1]. The radius of the Stienen disk of z_j is then given as

$$X = \min_{z_l} \varepsilon \|z_j - z_l\|, \forall z_l, z_j \in \Phi_u, l \neq j, \tag{6.1}$$

where the Stienen cell (S-cell) $\mathcal{B}_{z_j} = \{y \in b(z_j, X),\ y \in \mathbb{R}^2\}$ is such that $b(z_j, X)$ is a disk of radius X centered at z_j and $\varepsilon = 1/2$ for the NNM. In this chapter, we modify the Stienen's model to a more general case, where the scalar ε in (6.1) may take any value from the range $0 < \varepsilon \leq 1/2$. This gives us control on accurately modeling the coverage regions of UEs especially when the area is sparsely populated with UEs. The resulting PHS model may also be considered as a snapshot of the Lilypond model described in [97], where the disks expand dynamically as ε starts with 0 initially and increases over time eventually stopping at $\varepsilon = 1/2$. Notice that for any value of $0 < \varepsilon \leq 1/2$, the UE S-cells form disjoint sets such that $\mathcal{B}_{z_l} \cap \mathcal{B}_{z_j} = \emptyset, \forall z_j \neq z_l$. This implies that the MTDs inside a UE's S-cell are only associated to it. The coverage model will enter the soft sphere regime if ε is increased beyond 1/2. This is because the disks begin to overlap complicating the MTD association process, the study of which is beyond the scope of this work.

[1]Throughout this thesis the same notation is used for the UE and its location.

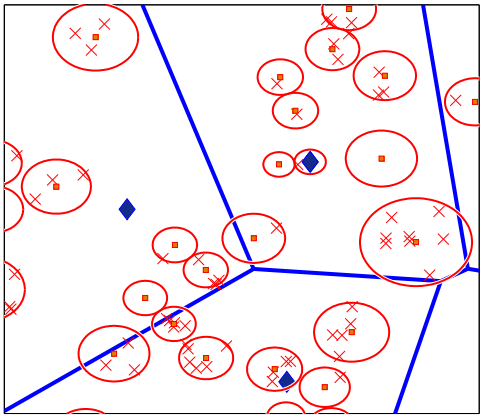


Figure 6.3.: Realization of the spatial setup. BSs are represented by blue, filled diamonds, UEs by orange filled squares and MTDs by cross marks. The circles indicate the S-cell boundaries: $\lambda_m = 10\lambda_u = 100\lambda_b, \varepsilon = 1/2$.

Analysis of Modified Stienen's Model

We now present some fundamental results pertaining to the Stienen's model. The distribution $f_X(x)$ can be quantified using the concept of the void probability of a Poisson processe. The probability that the radius of S-cell X exceeds a certain threshold x is the probability that there is no UE at a distance $\varepsilon^{-1}x$ from the given UE. It can be expressed as $\mathbb{P}[X \geq x] = \exp\left(-\lambda_u \pi \varepsilon^{-2} x^2\right)$. This implies

$$f_X(x) = 2\lambda_u \pi \varepsilon^{-2} x \exp\left(-\lambda_u \pi \varepsilon^{-2} x^2\right). \tag{6.2}$$

The resulting process of MTDs inside the S-cells constitute a modified Matern cluster process [98], where the radius of the disks is random and is distributed according to (6.2). Using this distribution of the radius of the S-cell, the probability mass function of the number of MTDs inside an arbitrary S-cell [99] can be represented as follows

$$\mathbb{P}[N_m = n] \overset{(a)}{=} \mathbb{E}_X \left[\frac{(\lambda_m \pi x^2)^n}{n!} \exp\left(-\lambda_m \pi x^2\right)\right] = \frac{\eta}{(1+\eta)^{n+1}},\ n \in \mathbb{Z}_{\geq 0}, \tag{6.3}$$

where (a) follows from the fact that MTDs are Poisson distributed with mean measure $\lambda_m \pi x^2$, and $\eta = \lambda_u \varepsilon^{-2}/\lambda_m$. The probability that there are at least i MTDs

inside a S-cell can then be written as

$$\mathbb{P}\left[N_m \geq i\right] = \frac{\eta}{(1+\eta)} \sum_{n=i}^{\infty} (1+\eta)^{-n} = (1+\eta)^{-i}. \tag{6.4}$$

Similarly, the average number of MTDs inside a given S-cell is expressed as follows.

$$N_m^{avg} = \frac{\eta}{(1+\eta)} \sum_{n=0}^{\infty} n\,(1+\eta)^{-n} = \eta^{-1}. \tag{6.5}$$

6.2.2. Transmission Schemes

The transmissions in both aggregation and trunking stages suffer from channel impairments including small scale Rayleigh fading and path loss. As a consequence, the channel power gain $h \sim \exp(1)$ is a unit mean exponential random variable. Throughout this chapter, we assume a simple power law path loss function $r^{-\alpha}$ for a distance separation r where α is the path loss exponent. We consider the same value of α to account for MTD-UE and UE-BS links, however, the presented framework can be easily extended to account for various propagation environments. We further assume interference-limited communication between these links. The following subsections provide a detailed description of the transmission schemes for the aggregation and trunking phases.

6.2.2.1. Aggregation

We consider that the MTDs transmit at a fixed power P_m under various medium selection schemes for communicating with the UEs. Details are given as follows.

- *Multiplexed*: Under this scheme, the UEs schedule MTDs such that the transmissions are multiplexed in either frequency or time. The MTDs employ frequency division multiple access (FDMA) or time division multiple access (TDMA). If k MTDs are scheduled for transmission to a UE, then i) under FDMA, each MTD uses $1/k$ of the available bandwidth (BW) W_m in the aggregation time τT, and ii) under TDMA, each MTD utilizes the total BW W_m for the kth fraction of the aggregation time τT. Details of the channel assignment techniques are out of the scope of this work. Both TDMA and FDMA result in the same average data rate so we do not make further distinction between the two in the rest of the chapter. To account for the extra signaling for multiplexed transmission, we consider that the available aggregation time

$T_{eff} = \tau T - kT_{sig}$ is effectively smaller than T by $k \times T_{sig}$, where T_{sig} is the time reserved for signaling and resource allocation of MTDs.

- *Combined*: In case of combined transmission, the MTDs concurrently transmit their data to the UE and each MTD's transmission spans the available BW W_m and time τT. Even though the channel utilization is maximized for every MTD, there is a need for sophisticated signal processing techniques for joint decoding or SIC. Because of the powerful computation capabilities in smart phones, we consider that the UEs successively cancel interference starting with decoding the signal with the strongest power and subtracting it from the received signal to improve the signal-to-interference and noise ratio (SINR). The process is repeated until the signal of interest (SoI) is decoded. The received power at the UE is ordered as $\{P_m h_i r_i^{-\alpha}, P_m h_j r_j^{-\alpha}, ...\}$ such that $h_i r_i^{-\alpha} \geq h_j r_j^{-\alpha}$. Since the order in which the MTDs are decoded is based on both path loss and fading, the probability of successfully decoding the SoI depends on the sum of order statistics. Characterization of the sum of order statistics for SIC has been handled in [100] and [101] to name a few. However, in [102] the authors prove that the received power is dominated by path loss. Therefore, for simplification in our analysis, we assume that the received power is ordered with respect to the distance between the MTDs and the UE. The UEs successively decode the data from MTDs starting from the nearest until the MTD with the SoI is decoded.

6.2.2.2. Trunking

In this phase, the UEs transmit their own data along with the collected M2M data to their respective serving BSs. We consider that the cellular bandwidth W_u is equally divided among the active UEs located within the BS coverage and there in no intracell interference. For energy efficient operation, the UEs employ UL channel inversion power control. The transmit power under the truncated UL channel inversion power control is written as

$$P_u = \min\left(P_u^{max}, \rho_0 l(y)^{-1}\right). \tag{6.6}$$

Notice that that the transmit power is constrained by the upper limit P_u^{max}, which is the maximum transmit power of a UE and $l(y) = y^{-\alpha}$ is the path loss when the UE and the BS are separated by a distance y. The distribution of the distance between

an arbitrary UE and its associated BS follows a well known Rayleigh distribution and is given as [31] $f_Y(y) = 2\pi\lambda_b y \exp\left(-\lambda_b \pi y^2\right)$. In (6.6), the term ρ_0 is the normalizing factor depending on the receiver sensitivity of the BS. We can see from (6.6) that the UEs only at a certain distance $R_{max} = \left[\frac{P_u^{max}}{\rho_0}\right]^{1/\alpha}$ can successfully invert the path loss. The UEs outside a disk of radius R_{max} will transmit at maximum power. Unlike the truncated channel inversion power control presented in [22], where the UEs farther from R_{max} are forced to go into outage, we present a more realistic power control scheme as the disadvantaged UEs still get a chance to transmit.

6.3. Probability of Successful Aggregation and Trunking

For cellular downlink scenarios, network operators are interested in load balancing and maximizing the rate experienced by a UE as well as the overall area spectral efficiency of the network [77]. On the contrary, the performance metrics are quite different for cellular UL and M2M applications, where ensuring reliability and enhancing connectivity is the primary focus. Based on these criteria, we define a key performance determining metric for the analysis of the aggregation and trunking communication framework described above.

Definition 6.1. PSAT: The probability that a UE is able to successfully aggregate M2M data from k MTDs in time τT and can trunk it along with its own data in time $(1-\tau)T$ can be expressed as the product of rate coverage in aggregation and trunking phases. It is given as

$$\mathcal{P}_k^{\Upsilon} = \mathcal{R}_a^{\Upsilon} \times \mathcal{R}_t, \tag{6.7}$$

where $\Upsilon \in \{MX, SIC\}$ represents the transmission scheme used by the MTDs in the aggregation phase, $\mathcal{R}_a^{\Upsilon}$ is the rate coverage probability of the MTDs in the aggregation phase and $\mathcal{R}_t$ is the rate coverage probability of a UE in the trunking phase. The description and derivation of the rate coverage for each phase is given in the following subsections.

6.3.1. Aggregation Phase

In this phase, the UE collects data from the MTDs within its S-cell that are transmitted to it. We assume that the UEs aggregate data from k MTDs in a slot of bandwidth W_m and time τT. Each MTD has a fixed data rate requirement of R_m bits/s. The probability that the rate requirement is satisfied, also known as the rate coverage for MX and combined transmission schemes, is characterized using Shannon's capacity formulation and is discussed as follows.

6.3.1.1. Multiplexed transmission

The rate coverage for the multiplexed case can be represented as

$$\begin{aligned}\mathcal{R}_a^{MX} &= \left(\mathbb{P}\left[\frac{W_m}{k}\log_2(1+SIR_a^{MX}) \geq \frac{D_m}{T_{eff}}\right]\right)^k \\ &= \left(\mathbb{P}\left[SIR_a^{MX} \geq \theta_a^{MX}\right]\right)^k, \end{aligned} \tag{6.8}$$

where $D_m = R_m T$ is the number of data bits per MTD, $T_{eff} = \tau T - T_{sig}$ is the effective time available for MTD transmission after the signaling and channel reservation for MTDs and $\theta_a^{MX} = 2^{\frac{kD_m}{W_m T_{eff}}} - 1$. The SIR of the received signal at the UE from an arbitrary MTD inside the S-cell at a distance $R_{arb} = r$ is given as $SIR_a^{MX} = \frac{h\, r^{-\alpha}}{I_m}$, where $I_m = \sum_{\omega_j \in \Phi_m^{act} \setminus o} h_j ||\omega_j||^{-\alpha}$ is the aggregate interference power experienced by the UE from MTDs in the other S-cells and Φ_m^{act} is the set of active MTDs ω_j. The distribution of the distance R_{arb} is given by the following Lemma.

Lemma 6.1. *If an arbitrarily selected MTD is present inside the coverage region of a UE, the distribution of the distance between the UE and the MTD can be approximated as*

$$f_{R_{arb}}(r) \approx 2\pi r \lambda_{arb} \exp\left(-\pi r^2 \lambda_{arb}\right), \tag{6.9}$$

where $\lambda_{arb} = \varsigma \lambda_u$ *and* $\varsigma = (1 + \varepsilon^{-2})$.

Remark 6.1. The approximation in Lemma 6.1 is due to the fact that when the S-cell is conditioned to contain a point, the distribution of the radius of the resulting S-cell is no longer the same as that of (6.1). This is because there is a finite probability $\mathbb{P}[N_m = 0]$ that there is no MTD inside an arbitrary S-cell. However, as will be clear from the subsequent sections that the effect of this is negligible.

Proof. The unconstrained distribution of the distance between an arbitrary MTD and the nearest UE is Rayleigh distributed and is given as

$$f_{R_{uncon}}(r) = 2\pi r\,\lambda_u \exp\left(-\lambda_u \pi r^2\right). \tag{6.10}$$

However, in this case, the MTD must also lie in the S-cell of the nearest UE. The CDF of distance with the condition that the S-cell encapsulates the arbitrary MTD can be written as

$$\begin{aligned} F_{R_{arb}}(r) &\approx \mathbb{P}\left[R_{uncon} \leq r | R_{uncon} \leq X\right], \\ &= \frac{\mathbb{P}\left[R_{uncon} \leq r, R_{uncon} \leq X\right]}{\mathbb{P}\left[R_{uncon} \leq X\right]} = \frac{\mathbb{P}\left[R_{uncon} \leq \min\left(r, X\right)\right]}{\mathbb{P}\left[R_{uncon} \leq X\right]} \\ &= \frac{\int_0^r \left[1 - \exp\left(-\lambda_u \pi x^2\right)\right] f_X(x)\,dx + \int_r^\infty \left[1 - \exp\left(-\lambda_u \pi r^2\right)\right] f_X(x)\,dx}{\mathbb{P}\left[R_{uncon} \leq X\right]} \end{aligned} \tag{6.11}$$

where $\mathbb{P}\left[R_{uncon} \leq g\right] = \int_0^g f_{R_{uncon}}(r)dr = 1 - \exp\left(-\lambda_u \pi g^2\right)$ and

$$\mathbb{P}\left[R_{uncon} \leq X\right] = \int_0^\infty \left(1 - F_X(t)\right) f_{R_{uncon}}(t)\,dt = \varsigma^{-1}.$$

Evaluating the integrals in (6.11) and differentiating with respect to r gives (6.9). □

For brevity and conciseness, we separately highlight an important result used for analysis in the rest of the chapter.

Lemma 6.2. *The solution of the integral of the type* $\mathcal{C}(\alpha, \beta, d) = \int_d^\infty \frac{\omega}{1+\beta^{-1}\omega^\alpha} d\omega$ *is given as*

$$\mathcal{C}(\alpha, \beta, d) = \frac{\beta d^{(2-\alpha)}}{(\alpha - 2)} \mathcal{F}\left(\alpha, \beta d^{-\alpha}\right), \tag{6.12}$$

where $\mathcal{F}\left(\alpha, \beta d^{-\alpha}\right) = {}_2F_1\left(1, 1 - \frac{2}{\alpha}; 2 - \frac{2}{\alpha}; -\beta d^{-\alpha}\right)$ *and* ${}_2F_1\left(.,.;.;.\right)$ *is the generalized hypergeometric function. For the special case of* $\alpha = 4$*, we obtain*

$$\mathcal{C}(4, \beta, d) = \frac{1}{2}\sqrt{\beta}\tan^{-1}\left(\sqrt{\beta} d^{-2}\right). \tag{6.13}$$

The following theorem gives the coverage probability for the aggregation phase with MX transmission.

Theorem 6.1. *For a given SIR threshold θ, the probability that the UE successfully decodes the data from an arbitrary MTD within its S-cell is given as*

$$\mathcal{S}_a^{MX}(\theta) \approx \int_0^\infty \exp\left(-2\pi\lambda_m^{act}\mathbb{E}_Q\left[\mathcal{C}\left(\alpha, s_a, q\right)\right]\right) f_{R_{arb}}(r)dr, \tag{6.14}$$

where Q is distributed according to $f_Q(q) = 2\pi\lambda_m^{act} q \exp\left(-\lambda_m^{act}\pi q^2\right)$, $\lambda_m^{act} = \lambda_u(1+\eta)^{-1}$ and $s_a = \theta r^\alpha$.

Proof. Please refer to Appendix B.1. □

Conjecture 6.1. *The aggregation coverage in MX case $\mathcal{S}_a^{MX}(\theta)$ does not vary with respect to λ_m or λ_u as long as the ratio λ_m/λ_u is kept constant.*

Proof. As no closed form expression exists for $\mathcal{S}_a^{MX}(\theta)$, it is not possible to obtain a concrete proof. However, we can obtain an approximation for coverage for $\alpha = 4$ to prove the validity of this claim. Using (6.13), (6.14) can be simplified as

$$\mathcal{S}_a^{MX}(\theta) \approx \int_0^\infty \exp\left(-\pi\lambda_m^{act}\sqrt{s_a}\mathbb{E}_Q\left[\tan^{-1}\left(q^{-2}\sqrt{s_a}\right)\right]\right) f_{R_{arb}}(r)dr.$$

The function $\tan^{-1}\left(\sqrt{s_a}q^{-2}\right)$ is concave in q until the inflection point $q = q_0 = \left(\frac{s_a}{3}\right)^{1/4}$ and convex for $q > q_0$. To simplify the analysis, we apply Jensen's inequality to shift the expectation with respect to q inside the function. As $\mathbb{E}_Q[q] = \frac{1}{2\sqrt{\lambda_m^{act}}}$, we obtain

$$\mathcal{S}_a^{MX}(\theta) \approx \int_0^\infty \exp\left(-\pi\lambda_m^{act}\sqrt{\theta}r^2 \tan^{-1}\left(4\lambda_m^{act}\sqrt{\theta}r^2\right)\right) f_{R_{arb}}(r)\, dr.$$

Solving the integral with the arctan term inside the exponential is tricky. We therefore approximate $\tan^{-1}\left(4\lambda_m^{act}\sqrt{\theta}r^2\right)$ with a simple pulse function such that $\tan^{-1}\left(4\lambda_m^{act}\sqrt{\theta}r^2\right) = \begin{cases} 0 & r < r_0 \\ \pi/2 & r \geq r_0 \end{cases}$, where $r_0 = \sqrt{\frac{\gamma_{th}}{4\lambda_m^{act}\sqrt{\theta}}}$ is the point of transition

and γ_{th} controls the value of this point. We can now write the MX coverage as

$$\begin{aligned}\mathcal{S}_a^{MX}(\theta) &\approx \int_0^{r_0} f_{R_{arb}}(r)\,dr + \int_{r_0}^{\infty} \exp\left(-\frac{\pi^2}{2}\lambda_m^{act}\sqrt{\theta}r^2\right) f_{R_{arb}}(r)\,dr \\ &\approx 1-\exp\left(-\lambda_{arb}\pi r_0^2\right) + \left(1+\frac{\pi}{2}\frac{\lambda_m^{act}\sqrt{\theta}}{\lambda_{arb}}\right)^{-1} \exp\left(-\left(\lambda_{arb}+\frac{\pi}{2}\lambda_m^{act}\sqrt{\theta}\right)\pi r_0^2\right) \\ &\approx 1-\exp\left(-\frac{\pi\varsigma\gamma_{th}}{4\sqrt{\theta}}(1+\eta)\right) + \frac{\exp\left(-\pi\gamma_{th}\left(\frac{\varsigma}{4\sqrt{\theta}}(1+\eta)+\frac{\pi}{8}\right)\right)}{\left(1+\frac{\pi}{2}\frac{\sqrt{\theta}}{\varsigma(1+\eta)}\right)}. \qquad (6.15)\end{aligned}$$

It is evident from (6.15) that $\mathcal{S}_a^{MX}(\theta)$ depends on the ratio λ_m/λ_u as $\eta = \lambda_u\varepsilon^{-2}/\lambda_m$ and remains unchanged as long as λ_m/λ_u is kept constant. □

6.3.1.2. Combined transmission

The rate coverage in this case can be written as

$$\begin{aligned}\mathcal{R}_a^{SIC} &= \prod_{i=1}^{k} \mathbb{P}\left[W_m \log_2(1+SIR_{a,i}^{SIC}) \geq \frac{D_m}{\tau T}\right] \\ &= \prod_{i=1}^{k} \mathbb{P}\left[SIR_{a,i}^{SIC} \geq \theta_a^{SIC}\right], \qquad (6.16)\end{aligned}$$

where $\theta_a^{SIC} = 2^{\frac{Dm}{W_m\tau T}} - 1$ and $SIR_{a,i}^{SIC}$ is the received SIR at a typical UE when the SoI is from the ith nearest MTD ($i \in \{1,..k\}$) given that the UE has successfully decoded the data from all $(i-1)$ MTDs. The SIR of the received signal at the UE while decoding the ith nearest MTD at a distance $R_i = r$ can be written as $SIR_{a,i}^{SIC} = \frac{h\,r^{-\alpha}}{I_{in}+I_{out}}$, where $I_{in} = \sum_{j=i+1}^{N_m} h_j||u_j||^{-\alpha}$ is the aggregate interference power from the transmitting MTDs located in the same S-cell which have yet to be decoded and $u_j \in \Phi_m^{in} = \Phi_m \cap \mathcal{B}_o$. Notice that the summation starts from $i+1$ as at this stage, it is considered that the signal from the closest $(i-1)$ MTDs has been previously decoded and does not contribute to the aggregate interference. The term $I_{out} = \sum_{v_j \in \Phi_m^{out}} h_j||v_j||^{-\alpha}$ is the aggregate interference from the active MTDs outside the S-cell under consideration, where $\Phi_m^{out} = \Phi_m \cap \mathcal{B}_{z_j} \backslash \mathcal{B}_o : z_j \in \Phi_u$ constitutes the MTDs inside other S-cells. The following Lemma gives the distribution of the distance between the UE and its ith nearest MTD.

Lemma 6.3. *The distribution of distance between the typical UE and its ith nearest*

MTD inside its S-cell is given as

$$f_{R_i}(r) = 2\left(\pi\lambda_{eff}\right)^i r^{2i-1}\exp\left(-\lambda_{eff}\pi r^2\right), \tag{6.17}$$

where $\lambda_{eff} = \lambda_m + \lambda_u\varepsilon^{-2}$.

Proof. For the unconstrained case, the distribution of distance to the ith nearest MTD is given as [103]

$$f_{R_i^{uncon}}(r) = \frac{2}{\Gamma(i)}\left(\lambda_m\pi\right)^i r^{2i-1}\exp\left(-\lambda_m\pi r^2\right). \tag{6.18}$$

The derivation of distance distribution for the ith nearest MTD inside the S-cell follows the same procedure as the proof of Lemma 6.1. We first evaluate $F_{R_i}(r) = \mathbb{P}\left[R_i^{uncon} \leq r | R_i^{uncon} \leq X\right]$, which is the conditional CDF of distance ensuring that there are at least i MTDs inside the S-cell. It can be calculated as

$$F_{R_i}(r) = \frac{\int_0^r \left[1 - \frac{\Gamma\left(k,\lambda_u\pi x^2\right)}{\Gamma(k)}\right] f_X(x)\,dx + \int_r^\infty \left[1 - \frac{\Gamma\left(k,\lambda_u\pi r^2\right)}{\Gamma(k)}\right] f_X(x)\,dx}{\mathbb{P}\left[R_i^{uncon} \leq X\right]}, \tag{6.19}$$

where

$$\mathbb{P}\left[R_i^{uncon} < g\right] = \int_0^g f_{R_i^{uncon}}(r)dr = 1 - \frac{\Gamma\left(k,\lambda_u\pi g^2\right)}{\Gamma(k)}$$

and

$$\mathbb{P}\left[R_i^{uncon} < X\right] = \mathbb{P}\left[N_m \geq i\right] = \left(1+\mu\right)^{-i}.$$

Substituting these expressions into (6.19), evaluating the integrals and differentiating with respect to r yields (6.17). □

The following theorem gives the probability of coverage when a UE decodes its k nearest MTDs using SIC.

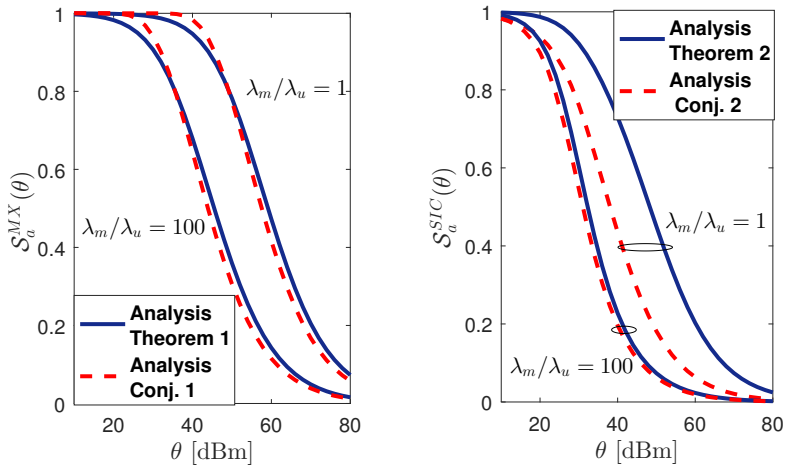
Theorem 6.2. *Given that the MTDs inside a UE's S-cell transmit simultaneously and the SIR decoding threshold is* θ*, the probability that the UE employs SIC and*

successfully decodes the data from k MTDs can be represented as

$$\mathcal{S}_a^{SIC}(\theta) \approx \prod_{i=1}^{k} \mathbb{E}_{R_i}\Bigg[\exp\Bigg(-2\pi\lambda_m\left[\mathcal{C}\left(\alpha, s_a, r\right) - \mathbb{E}_{X|X>r}\left[\mathcal{C}\left(\alpha, s_a, x\right)\right]\right)\Bigg) \exp\Bigg(-2\pi\lambda_u \mathbb{E}_{R_{uncon}}\left[\mathcal{C}\left(\alpha, N_m^{avg} s_a, w\right)\right]\Bigg)\Bigg], \tag{6.20}$$

where $s_a = \theta r^{\alpha}$, $f_{R_{uncon}}(r) = 2\pi r\, \lambda_u \exp\left(-\lambda_u \pi r^2\right)$ *and* $f_{X|X>r}(x) = \dfrac{f_X(x)}{\exp\left(-\lambda_u \varepsilon^{-2} \pi r^2\right)}$.

Proof. Please refer to Appendix B.2. □



(a) Multiplexed transmission: $\gamma_{th} = \pi/6$.

(b) Combined transmission.

Figure 6.4.: Comparison of the exact analysis with results from Conjectures 6.1 and 6.2: $k = 1$.

Conjecture 6.2. *The aggregation coverage in case of combined transmission* $\mathcal{S}_a^{SIC}(\theta)$ *does not vary with respect to* λ_m *or* λ_u *as long as the ratio* λ_m/λ_u *is kept constant.*

Proof. The proof follows similar steps as that of Conjecture 6.1. For the case of $\alpha = 4$, we have for the ith nearest MTD

$$\mathcal{S}_{a,i}^{SIC}(\theta) \approx \mathbb{E}_{R_i}\Bigg[\exp\Bigg(-\pi\lambda_m\sqrt{\theta}r^2\left[\tan^{-1}\left(\sqrt{\theta}\right) - \mathbb{E}_{X|X>r}\left[\tan^{-1}\left(\sqrt{\theta}r^2x^{-2}\right)\right]\right]\Bigg) \exp\Bigg(-2\pi\lambda_u\sqrt{N_m^{avg}\theta}r^2\mathbb{E}_{R_{uncon}}\left[\tan^{-1}\left(\sqrt{N_m^{avg}\theta}r^2w^{-2}\right)\right]\Bigg)\Bigg].$$

To simplify the analysis, we obtain a loose lower bound by removing the second term in the first exponential and setting $\tan^{-1}\left(\sqrt{N_m^{avg}\theta}r^2w^{-2}\right) = \frac{\pi}{2}$. We therefore get

$$\begin{aligned}\mathcal{S}_{a,i}^{SIC}(\theta) &\geq \int_0^{\infty} \exp\left(-\pi\theta^{1/2}\left(\frac{\pi}{2}\lambda_u\left(N_m^{avg}\right)^{1/2} + \lambda_m \tan^{-1}\left(\theta^{1/2}\right)\right)r^2\right) f_{R_i}(r)dr \\ &\geq \left(1 + \frac{\pi\left(\theta\eta\right)^{1/2}\varepsilon^2}{2\left(1+\eta\right)} + \frac{\left(\theta\right)^{1/2}\tan^{-1}\left(\theta^{1/2}\right)}{\left(1+\eta\right)}\right)^{-i}. \end{aligned} \tag{6.21}$$

It is clear from (6.21) that $\mathcal{S}_{a,i}^{SIC}(\theta)$ also depends on the ratio λ_m/λ_u. This completes the proof. □

The derived results in Conjectures 6.1 and 6.2 are compared with the respective analysis in (6.14) and (6.20) in Figures 6.4. The approximation for the MX case is quite close to the actual analysis. The lower bound in (6.20) is not very accurate especially when $\lambda_m/\lambda_u = 1$. However, it still follows a similar trend as the original analysis and becomes very tight for higher values of λ_m/λ_u. Because of the dependence of coverage on η alone, the following important conclusion can be derived. For both MX and combined transmission schemes, the effect of interference intensifies with the increase in both λ_m and λ_u. However, for a fixed λ_m/λ_u, this effect is perfectly canceled with the improved desired signal strength. In the case of MX transmission, the intensity of active interfering MTDs λ_m^{act} increases but the average distance to the arbitrary MTD inside the S-cell decreases. Similarly, in the case of combined transmission, the average number of active MTDs inside an S-cell N_m^{avg} increases causing the interference power to aggravate, but this effect is also perfectly balanced by the reduction in path loss of the desired link.

6.3.2. Trunking Phase

In this phase, the UEs transmit to the BS the data collected from k MTDs along with their own data. The rate coverage in trunking phase can then be written as

$$\begin{aligned}\mathcal{R}_t &= \mathbb{E}_{N_u}\left[\mathbb{P}\left[\frac{W_u}{n}\log_2(1+SIR_t) \geq \frac{D_u + kD_m}{(1-\tau)T}\right]\right], \\ &= \mathbb{E}_{N_u}\left[\mathbb{P}\left[SIR_t \geq \theta_t\right]\right], \end{aligned} \tag{6.22}$$

where D_u satisfies the UL rate requirement for the UE such that $R_u = D_u/T$, $\theta_t = 2^{\frac{n(D_u+kD_m)}{W_u(1-\tau)T}} - 1$ and N_u is the number of UEs attached to a typical cell. The

distribution of N_u is given as in [75] $f_{N_u}(n) = \frac{3.5^{4.5}\Gamma(n+4.5)(\lambda_u/\lambda_b)^n}{\Gamma(4.5)\,n!(\lambda_u/\lambda_b+3.5)^{n+4.5}}$. To simplify things, we adopt the mean-load approximation as in [77]. The average number of UEs inside a macrocell is given by $N_u^{avg} = 1 + 1.28\,\lambda_u/\lambda_b$. Therefore, the rate coverage simplifies to

$$\mathcal{R}_t = \mathbb{P}\left[\frac{W_u}{N_u^{avg}}\log_2(1+SIR_t) \geq \frac{D_u + kD_m}{(1-\tau)T}\right], \tag{6.23}$$

and $\theta_t = 2^{\frac{N_u^{avg}(D_u+kD_m)}{W_u(1-\tau)T}} - 1$. We will make use of this approximation throughout the course of this chapter. The SIR at the BS in trunking phase at a given distance separation $Y = y$ can be represented as

$$SIR_t = \frac{P_u h\, y^{-\alpha}}{I_u}, \tag{6.24}$$

where P_u is the variable transmit power given in (6.6) depending on the distance y between the UE and the BS it is associated with and $I_u = \sum_{z_j \in \Phi_u^{act}\setminus z_o} P_{u,j} h_j\, ||z_j||^{-\alpha}$ is the interference power from the set of active UEs $z_j \in \Phi_u^{act}$ in other macrocells as we assume that there is no intra-cell interference. In the following Lemma, we obtain the average power transmitted by a UE which will help characterize the coverage in the trunking phase.

Lemma 6.4. *Under constrained UL channel inversion power control, the average power transmitted by a UE is given as*

$$P_u^{avg} = \frac{\rho_0 \Gamma(\delta)}{(\lambda_b \pi)^{\alpha/2}}\,\gamma\left(\lambda_b \pi R_{max}^2, \delta\right) + p_{max} P_u^{max}, \tag{6.25}$$

where $\delta = 1 + \alpha/2$, $R_{max} = \left[\frac{P_u^{max}}{\rho_0}\right]^{1/\alpha}$, $p_{max} = \exp\left(-\lambda_b \pi R_{max}^2\right)$ *and* $\gamma(b,a) = 1/\Gamma(a)\int_0^b t^{a-1}\exp(-t)\,dt$ *is the normalized lower incomplete Gamma function.*

Proof. The average transmit power is calculated by taking expectation of (6.6) over y. It is given as

$$\begin{aligned}
\mathbb{E}[P_u] &= \rho_0 \int_0^{R_{max}} y^{\alpha} f_Y(y)\,dy + P_u^{max}\int_{R_{max}}^{\infty} f_Y(y)\,dy \\
&= 2\pi\lambda_b\rho_0 \int_0^{R_{max}} y^{\alpha+1}\exp\left(-\lambda_b \pi y^2\right)\,dy + P_u^{max}\exp\left(-\lambda_b \pi R_{max}^2\right).
\end{aligned} \tag{6.26}$$

Parameter	Description	Value
α, ε	Path loss exponent, HS coefficient	4, 1/2
$\lambda_b, \lambda_u, \lambda_m$	MTD, UE and BS density	$[1, 20, 200]/\left(A = \pi 500^2\right)$
P_m, P_u^{max}, ρ_0	MTD and UE max. transmit power, BS receiver sensitivity	$[-18, 23, -80]$ dBm
W_m, W_u	MTD and UE bandwidth	$k \times 180$ kHz, 10 MHz
R_m, R_u	Desired M2M and UE data rates	$[4, 20]$ kbps
τ	Fraction of UL slot for aggregation	0.2
T, T_{sig}	UL slot time period, Signaling time per MTD	1 ms, 50 µs

Table 6.1.: List of simulation parameters.

Solving the integral in (6.26), we then obtain (6.25). □

The coverage probability for a generic UL with constrained channel inversion power control is given by the following theorem.

Theorem 6.3. *When a generic user transmits to the nearest BS by employing constrained channel inversion power control, the probability that the BS can successfully decode this signal can be expressed as*

$$\mathcal{S}_t(\theta) \;=\; (1 - p_{max})\, \mathcal{L}_{I_u}(s_1) + \int_{R_{max}}^{\infty} \mathcal{L}_{I_u}(s_2)\, f_Y(y) dy \tag{6.27}$$

where $s_1 = \theta/\rho_0$, $s_2 = \theta y^{\alpha}/P_u^{max}$ *and*

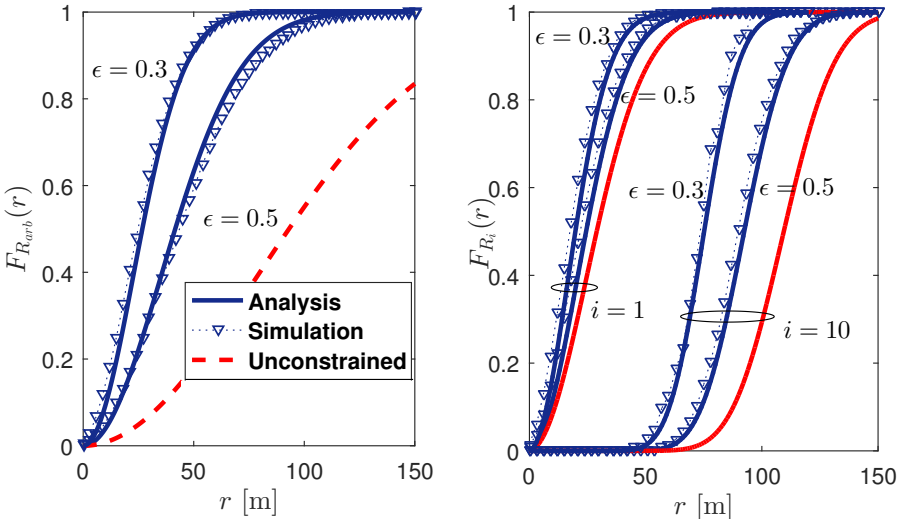
$$\begin{aligned}\mathcal{L}_{I_u}(s_t) =& \exp\left(-2\left(1 - p_{max}\left(1 + \lambda_b \pi R_{max}^2\right)\right)\mathcal{C}(\alpha, s_t \rho_0, 1)\right) \\ & .\exp\left(-2\lambda_b \pi p_{max} \int_{R_{max}}^{\infty} u^2 \mathcal{C}(\alpha, s_t P_u^{max} u^{-\alpha}, 1) f_Y(u) du\right). \end{aligned} \tag{6.28}$$

Proof. Please refer to Appendix B.3. □

6.4. Results and Discussion

In this section, we verify our analysis using Monte-Carlo simulations and provide some useful design insights for the aggregation and trunking framework. To compute

the distribution of distance and coverage, we conducted 10^4 iterations. In each iteration, the BSs, UEs and MTDs are distributed independently according to HPPPs with densities λ_b, λ_u and λ_m respectively in a simulation area of radius 3 km. The radius of each S-cell is calculated for each UE with respect to its neighboring UEs according to (6.1). The list of simulation parameters and their description is given in table 6.1 unless otherwise stated.

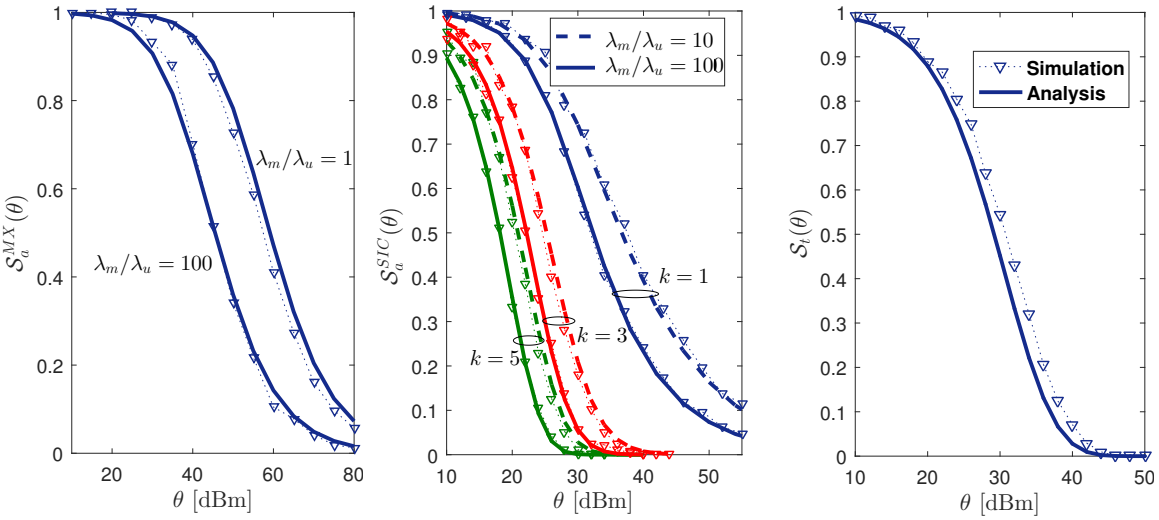


(a) Arbitrarily distributed MTD. **(b)** ith nearest MTD.

Figure 6.5.: CDF of the distance between the UE and MTD inside the S-cell: $\lambda_m = 500/A$.

We begin with the verification of Lemmas 6.1 and 6.3. For the distance between the UE and an arbitrarily distributed MTD, we generate the S-cells and fix the location of the MTD at the origin. The distance between the UE and an MTD is recorded if the MTD lies inside the nearest UE's S-cell. The realizations where the MTD lies outside the S-cell are ignored. For the distance to the ith nearest MTD, the location of a UE is fixed to the origin and the S-cells are generated. The distance of the ith nearest MTD from the origin is recorded if the MTD lies inside the S-cell of the UE. For a clear comparison, we obtain the cumulative distribution function (CDF) of the distances as $F_{R_{arb}}(r) \approx \int_0^r f_{R_{arb}}(\nu)d\nu = 1 - \exp\left(-\lambda_{arb}\pi r^2\right)$, and $F_{R_i}(r) = \int_0^r f_{R_i}(\nu)d\nu = 1 - \Gamma(i, \lambda_{eff}\pi r^2)$, where $\Gamma\left(a, b\right) = \int_b^\infty t^{a-1}\exp(-t)\,dt$ is the upper incomplete Gamma function. As shown in Fig. 6.5, the simulation accurately matches our analytical results for various values of ε and i. As ε decreases, the size of S-cell also decreases and therefore, the distance between the UE and

MTDs inside the macrocell also decreases. Compared with the unconstrained case, where $F_{R_{uncon}}(r) = 1 - \exp(-\lambda_u \pi r^2)$ and $F_{R_i^{uncon}}(r) = 1 - \Gamma(i, \lambda_m \pi r^2)$, we see that the corresponding distributions with S-cell restrictions differ in the densities only. As $\lambda_{arb} \geq \lambda_u$ and $\lambda_{eff} \geq \lambda_m$, the average distance between the UE and the MTDs is smaller. This demonstrates that the current network with S-cell boundary restrictions can be translated into a denser unconstrained HPPP network.

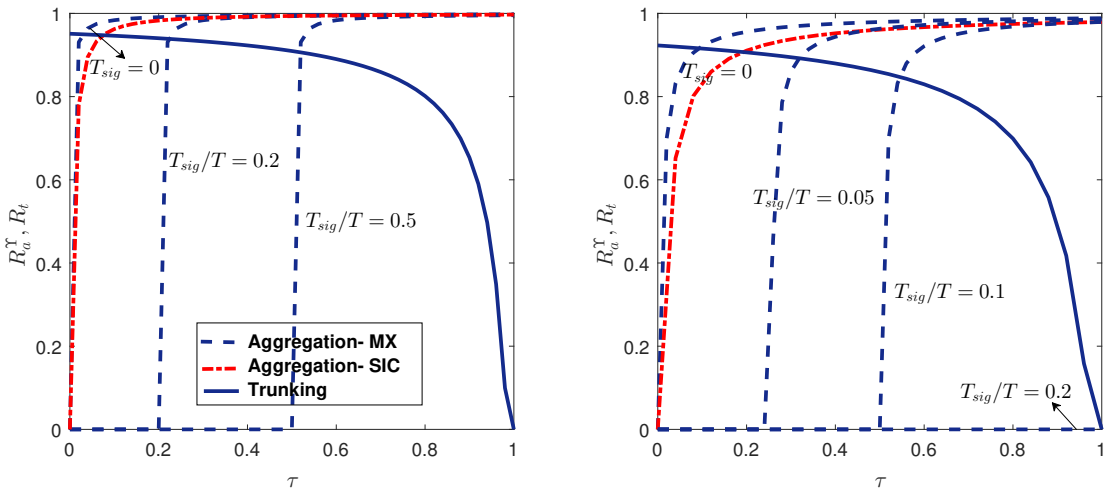


(a) Aggregation: Multiplexed. **(b)** Aggregation: Combined. **(c)** Trunking: UE-BS uplink.

Figure 6.6.: SIR coverage probability.

Moving on, we validate the expressions for SIR coverage for the aggregation and trunking phases derived in theorems 6.1, 6.2 and 6.3 in Fig. 6.6. The plots demonstrate a strong agreement between the simulations and the derived analytical results. We observe in Fig. 6.6b that when a UE employs SIC to decode the signals, $\mathcal{S}_a^{SIC}(\theta)$ sharply drops with the increase in k. This is because, in order to decode the kth strongest signal, a UE has to ensure that all $1, .., k-1$ signals are successfully decoded. Figures 6.6a and 6.6b further reinforce Conjectures 6.1 and 6.2 by showing that the coverage probability in aggregation phase $\mathcal{S}_a^{\Upsilon}(\theta), \Upsilon = \{MX, SIC\}$ remains unchanged as long as the ratio λ_m/λ_u is kept constant, while it decreases when λ_m/λ_u increases. The drop in $\mathcal{S}_a^{MX}$ is attributed to the increase in the interferer intensity $\lambda_m^{act} = \lambda_u(1+\eta)^{-1}$ as $\lambda_m/\lambda_u \to \infty$, $\lambda_m^{act} = \lambda_u$. This implies that each S-cell has at least one MTD transmitting to its UE. In the case of combined transmission, the increase in λ_m/λ_u also results in aggravated interference at the UE. The is because the number of interfering MTDs transmitting within the S-cell increases as

$\mathbb{P}[N_m \geq i] = (1+\eta)^{-i} \to 1, \forall i$ as $\lambda_m/\lambda_u \to \infty$. Moreover, the average interference power from MTDs outside the typical S-cell also increases as $N_m^{avg} P_m = \eta^{-1} P_m$ increases with λ_m/λ_u. On the contrary, the trunking coverage $\mathcal{S}_t(\theta)$ does not depend on λ_u or λ_m as is evident from (6.27).



(a) $k = 1$. **(b)** $k = 5$.

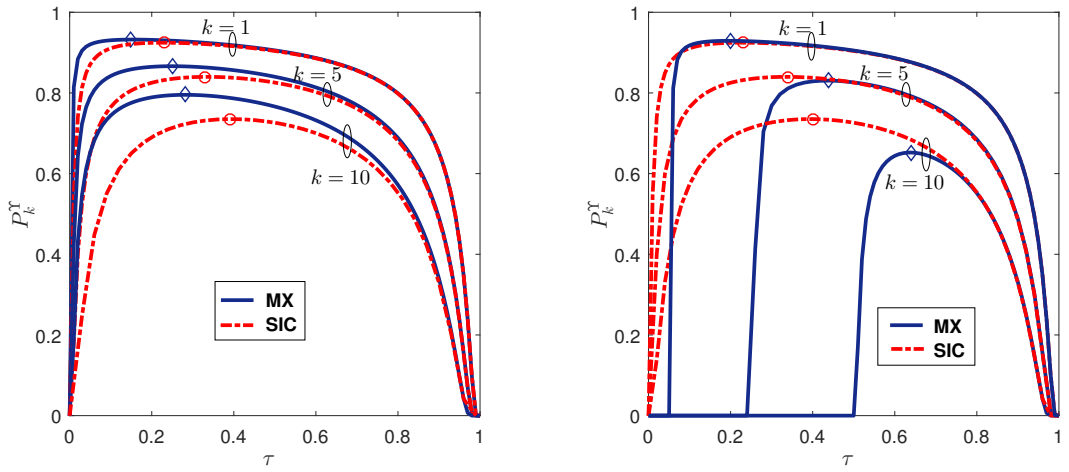
Figure 6.7.: Rate coverage probability in aggregation and trunking phases.

Fig. 6.6c reveals that the PPP assumption for the interfering UEs in UL transmission in the trunking phase is quite accurate even though the interfering UEs constitute a Poisson perturbed lattice process.

6.4.1. Effect of design parameters on Probability of Successful Aggregation and Trunking (PSAT)

After validation of the preliminary results using network simulations, we investigate in detail the factors affecting PSAT and the scenarios where aggregation and trunking is feasible. Fig. 6.7 explores the effect of the number of MTDs per S-cell k and the fraction τ of the UL time slot reserved for aggregation on the aggregation and trunking rate coverage performance. The results are intuitive as the increase in k causes both $\mathcal{R}_t$ and $\mathcal{R}_a^{\Upsilon}$ to degrade. However, for a given k, an interesting trade off in τ is revealed. As we increase τ, the trunking rate coverage $\mathcal{R}_t$ drops, while the aggregation rate coverage $\mathcal{R}_a^{\Upsilon}$ increases. This is because a higher τ corresponds to

a better aggregation opportunity for k MTDs as θ_a^{Υ} decreases whereas, it results in a degraded trunking performance as θ_t increases because lesser time is available for trunking UE and M2M data from k MTDs $(D_u + kD_m)$. Hence, there must exist an optimal $\tau = \tau_{\Upsilon}^*$ which maximizes $\mathcal{P}_k^{\Upsilon}$. Another important factor to take into consideration while deciding the transmission scheme in aggregation phase is the signaling overhead in the MX case. Even though the MX transmission is more robust and less



(a) Multiplexed transmission: $T_{sig} = 0$. **(b)** Combined transmission: $T_{sig}/T = 0.05$.

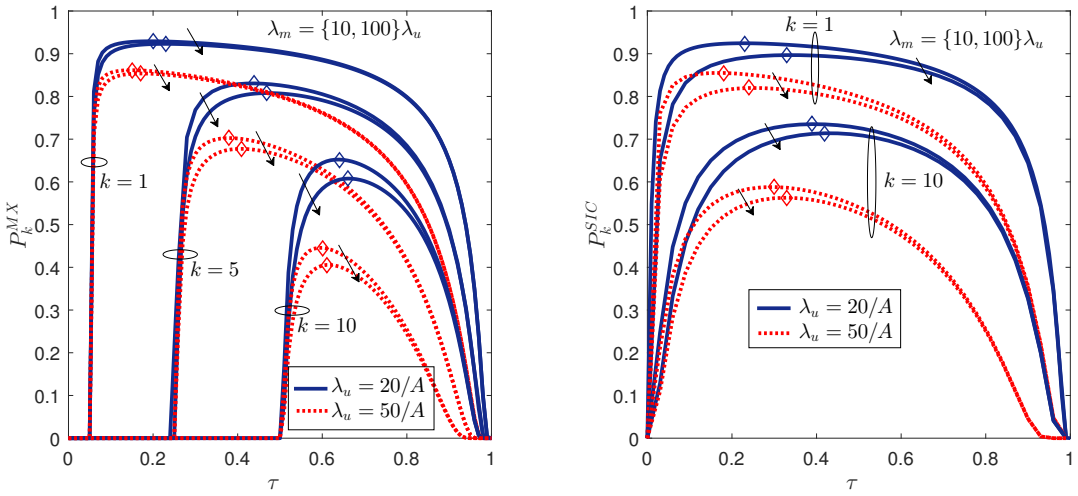
Figure 6.8.: Variation of PSAT with respect to τ and k. Blue diamonds and red circles represent the maximum value of $\mathcal{P}_k^{MX}$ and $\mathcal{P}_k^{SIC}$ at τ_{MX}^* and τ_{SIC}^* respectively.

prone to error, we see that even a slightly higher signaling and scheduling time T_{sig} results in complete outage for small values of τ. This is because there is no time for aggregation as $\tau T \leq kT_{sig}$. This problem is avoided in the combined transmission as all MTDs inside the S-cell are transmitting in the same time and no coordination by the UE is required.

Design of optimal aggregation time fraction

We explore the effect of various parameters on the optimal aggregation time fraction τ_{Υ}^* which maximizes PSAT. Fig. 6.8 shows the variation of $\mathcal{P}_k^{\Upsilon}$ with respect to τ and k for both transmission schemes. A common observation is that as k increases, the maximum achievable PSAT decreases as both $\mathcal{R}_t$ and $\mathcal{R}_a^{\Upsilon}$ decrease. The increase

in k also causes τ^*_Υ to increase and the optimal point shifts further right. This implies that the degradation in $\mathcal{R}^\Upsilon_a$ is higher than in $\mathcal{R}_t$. We also note that when $T_{sig} = 0$ and the MTDs have been pre-assigned the resources, it is better to adopt multiplexed transmission. However, when the MTDs have to be scheduled, T_{sig}



(a) Multiplexed transmission of MTDs. **(b)** Combined transmission of MTDs.

Figure 6.9.: Effect of MTD and UE densities on PSAT. The direction of arrows indicates increasing λ_m.

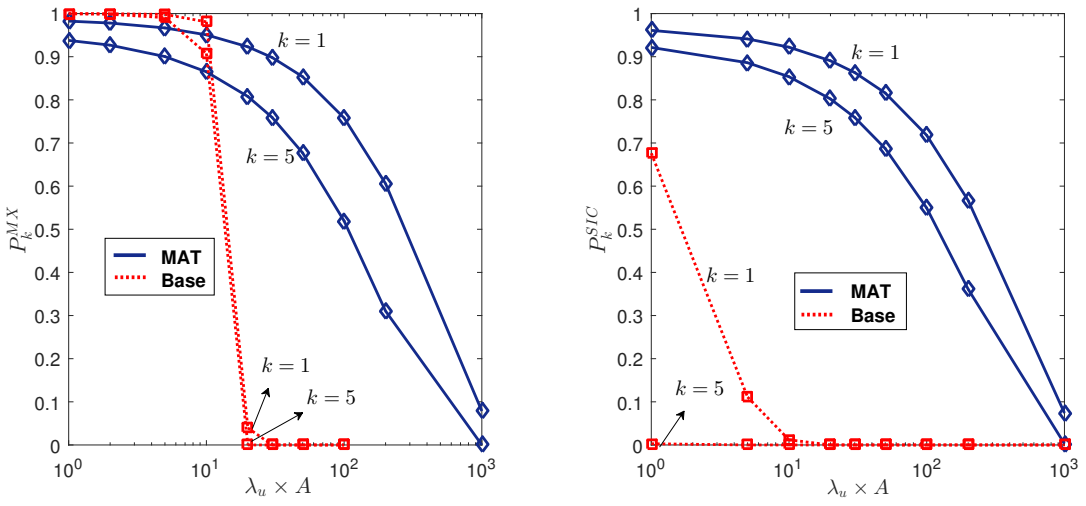
cannot be ignored. The signaling overhead adversely affects the available resources for data transmission causing $\mathcal{P}^{MX}_k$ to degrade. Therefore, combined transmission of MTDs gives much better performance. Another important point to note is that for higher values of k, the MX scheme requires very careful tuning of τ to maximize $\mathcal{P}^{MX}_k$ as seen in Fig. 6.8a. Even changing $\tau = \tau^*_{MX}$ by just10% could result in severe degradation in $\mathcal{P}^{MX}_k$. On the other hand, combined transmission is more resilient to changes in τ as we observe flatter curves for $\mathcal{P}^{SIC}_k$ in Figure 6.8b.

The effect of MTD and UE densities on $\mathcal{P}^\Upsilon_k$ is studied in Figure 6.9. We see that unlike $\mathcal{S}^\Upsilon_a(\theta)$ and $\mathcal{S}_t(\theta)$, the PSAT $\mathcal{P}^\Upsilon_k$ is no longer independent of the individual densities. For the case of fixed λ_u, as λ_m increases, $\mathcal{P}^\Upsilon_k$ drops and τ^*_Υ also increases. This drop in $\mathcal{P}^\Upsilon_k$ is solely attributed to the drop in $\mathcal{R}^\Upsilon_a$ as when λ_u is fixed, $\mathcal{R}_t$ remains unchanged. Alternatively, when λ_u is increased keeping λ_m/λ_u constant, there is a drop in $\mathcal{P}^\Upsilon_k$, which is particularly significant for larger values of k. This is because the trunking rate coverage $\mathcal{R}_t$ depends on λ_u and not λ_m as the trunked M2M data

is fixed to k MTDs. The value of θ_t increases with λ_u as the average number of UEs inside a macrocell increases (see N_u^{avg} in (6.23)) and lesser bandwidth is available for transmission for a UE. The value of τ_Υ^* is also observed to decrease slightly. We see that the trunking performance plays a critical role in maintaining a high PSAT for roughly the same τ_Υ^*. Improving UL coverage by increasing λ_b or introducing more BSs to have lesser UEs per BS could be a possible solution to improve $\mathcal{R}_t$ and in turn $\mathcal{P}_k^\Upsilon$ for $\tau = \tau_\Upsilon^*$.

6.4.2. Comparison with Baseline

We evaluate the effectiveness of the proposed MAT scheme with the baseline scenario, where the MTDs transmit to the BS directly and there is no hierarchy involved. For a realistic comparison, we consider that the BS gives priority to the cellular UEs and the available bandwidth is equally divided among the cellular users and all the active MTDs are treated as a single user in this division. The analysis for baseline scenario for the two transmission strategies is given as follows.



(a) Multiplexed transmission of MTDs. **(b)** Combined transmission of MTDs.

Figure 6.10.: Performance of the MAT scheme compared to baseline for various number of cellular UEs: $\lambda_m/\lambda_u = 100$, $A = \pi 500^2$.

Multiplexed transmission

In this case, both the cellular UEs and MTDs contend for cellular resources and no two devices can use the same channel inside a macrocell at a given time. In a snapshot of such a network, where the MTD is transmitting to the typical BS, the interference at the typical BS is coming from MTDs in other macrocells . The probability of successful aggregation in this case can be easily obtained using standard coverage probability expressions for downlink given in [31] as no channel inversion power control is assumed for MTDs. The probability that the BS is able to successfully decode the data from k MTDs can then be expressed as

$$\mathcal{P}_{a,B}^{MX} \approx \mathbb{E}_Y \left[\exp\left(-2\pi\lambda_b \mathcal{C}\left(\alpha, \theta_{a,B}^{MX} y^{\alpha}, y\right)\right)\right]^k, \tag{6.29}$$

where $\theta_{a,B}^{MX} = 2^{\frac{(N_u^{avg}+1)Dm}{W_u T}} - 1$.

Combined transmission

As there is no S-cell boundary restriction in this scenario, the rate coverage for the case of combined transmission of MTDs is given by the following theorem.

Theorem 6.4. *When the MTDs simultaneously transmit to the BS, the probability that the BS successfully decodes the data from k MTDs using SIC can be represented as*

$$\mathcal{P}_{a,B}^{SIC} \approx \prod_{i=1}^{k} \mathbb{E}_{R_i^{uncon}} \left[\exp\left(-2\pi N_m^{avg} \lambda_u \mathcal{C}\left(\alpha, \theta_{a,i,B}^{SIC} r^{\alpha}, r\right)\right)\right], \tag{6.30}$$

where $\theta_{a,i,B}^{SIC} = 2^{\frac{(N_u^{avg}+1)Dm}{W_n T}} - 1$.

Proof. The proof is based on the same lines as that of (6.20), but simplified as there is no S-cell boundary restriction. We only present the sketch of the proof to avoid repetition. The distribution of the distance to the ith nearest MTD R_i^{uncon} is given by (6.18). At the ith decoding stage and $R_i^{uncon} = r$, the interference power comes from all active MTDs farther than the ith nearest MTD. The Laplace transform of the resulting interference is written as $\mathcal{L}_{I_{out}}(s_{a,B}) \approx \exp\left(-2\pi N_m^{avg} \lambda_u \mathcal{C}\left(\alpha, \theta_{a,i,B}^{SIC} r^{\alpha}, r\right)\right)$, where the minimum distance separation to the nearest interfering MTD is equal to r. Averaging the above equation with respect to r completes the proof. □

Figure 6.10 compares the PSAT for the MAT scheme and the baseline scenario

for both multiplexed and combined transmission strategies. We plot the maximum achievable PSAT at $\tau = \tau^*_\Upsilon$ and compare it with baseline probabilities for various values of UE density λ_u keeping a fixed ratio $\lambda_m/\lambda_u = 100$. It is evident that introducing hierarchical aggregation with the MAT scheme significantly enhances the aggregation performance, especially for higher cell loads and combined transmission of MTDs. This is because with the MAT scheme, only proximate MTDs within the S-cell are allowed to transmit. This type of clustering helps generate less interference compared to the baseline scenario where all MTDs transmit simultaneously.

6.4.3. Maximum Allowable k

To realize the full potential of the MAT scheme from a design perspective, we answer

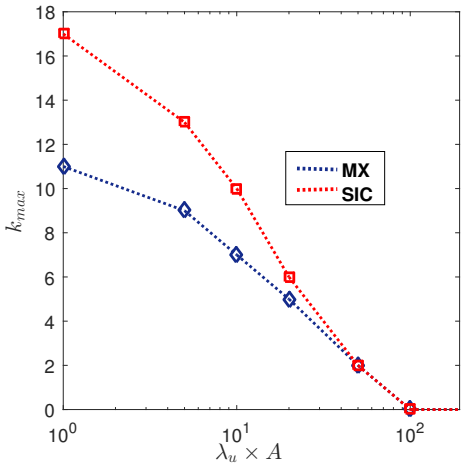


Figure 6.11.: Maximum number of MTDs that can be served by a UE under the MAT scheme: $\mathcal{P}^{\Upsilon}_{k_{max}} = 0.8$, $\lambda_m/\lambda_u = 100$.

the following question. What is the maximum number of MTDs k_{max} that can be served by a UE provided that the QoS requirements of both the UE and MTDs are satisfied? Figure 6.11 displays how k_{max} varies with λ_u when the ratio λ_m/λ_u is kept fixed. We see that with combined transmission of MTDs, a higher number of MTDs can be supported as compared to the multiplexed transmission of MTDs. A common trend is that k_{max} decreases steadily with the increase in λ_u to achieve the same PSAT. However, even for higher cell loads $\sim 50/A$, each UE can still aggregate data from up to 2 MTDs without compromising the QoS requirements.

6.5. Conclusion

This work proposes a novel stochastic geometric framework for M2M data aggregation and trunking (MAT) scheme, where the UEs establish D2D links with the MTDs and transmit this data to the BS. A hard sphere model is used to characterize the coverage regions of the UEs. In addition, various MTD transmission schemes are considered, i) multiplexed transmission, where a time/frequency resource is dedicated for an MTD in close proximity to the UE, and ii) combined transmission, where all the MTDs in close proximity of a UE transmit simultaneously and the UE employs SIC to decode and differentiate between the data from different MTDs. We explore the design space of this system and compare the performance of both transmission schemes with their corresponding baselines where no hierarchical aggregation is involved. Our results reveal that for a fixed number of MTDs, there exists an optimal fraction of the UL time slot that should be reserved for data aggregation to maximize the probability of successful data delivery. The proposed MAT scheme is particularly useful compared to baseline when the UEs are densely populated. An important conclusion of this study is that a UE can easily aggregate and data from several MTDs without significantly sacrificing its QoS requirements. The proposed framework can easily be extended to account for multi-tier, multi-antenna scenarios.

7 Cellular Network Underlaid with Cognitive M2M Network

In this chapter, we revisit the definition of cognition in the context of M2M networks and argue that both the energy efficiency and the spectrum efficiency are key design constraints. We define a new performance metric called the 'overall link success probability' which encapsulates these constraints. The overall link success probability is characterized by both the self-sustainability of the link through energy harvesting and the availability of spectrum for transmissions. We consider a large scale deployment of a cognitive machine-to-machine (CM2M) network empowered by solar energy harvesting at the MTDs, where the MTDs transmit to a nearby CM2M aggregator. We assume that the MTDs in a CM2M network share spectrum with the cellular BS subject to a certain co-existence constraint. Both the self-sustainability of the link, and the availability of transmission opportunities, are coupled through a common parameter, i.e., the transmit power at the MTDs. We demonstrate through our stochastic geometry analysis that the overall link level success probability can be maximized by optimizing the transmit power of the MTDs.

7.1. Introduction

A recent survey from the Economist Intelligence Unit (EiU) [104] indicated that around 75% of businesses are either actively considering to use, or are currently employing M2M solutions. It is reasonable to assume that cellular spectrum will be widely used by the MTDs because of the performance it can guarantee in terms of delay, security, coverage, and the fact that they already exist and are widespread so

can be used anywhere and at any time to deploy a smart environment. Examples of M2M applications currently using cellular spectrum include, but are not limited to, video surveilance, asset tracking, vehicle to infrastructure communication and health and fitness tracking. The proliferation of MTDs operating in the cellular spectrum would lead to congestion at the access network and a consequent loss of performance of the smart system. The internet of things (IoT) paradigm will soon become heavy on 'things' while it would struggle on the connectivity frontier.

A quick glance at the frequency allocation charts provided by the regulatory bodies reveals that most of the cellular spectrum is already assigned and the margin for accommodating emerging M2M applications is low. It seems natural to think of the spectrum scarcity as a real challenge posed due to the high utilization of the Hertzian medium. However, a reality check on the usage patterns of the available spectral resources reveals that in a nutshell the spectrum scarcity is nothing but artificial. Spectrum occupancy measurements [105, 106] have revealed that these licensed bands are highly under-utilized across space and time. From 13% to 87% of the radio spectrum remains unused across spatio-temporal domains. This sporadic utilization of scarce electromagnetic spectrum creates an artificial scarcity which in turn poses the connectivity challenge for M2M communication. Regulatory bodies such as the FCC (in the USA) and Ofcom (in the UK) have already noticed that such under-utilization of the spectrum can be avoided by more flexible and dynamic spectrum access (DSA) mechanisms [107]. Radio spectrum is a multi-dimensional entity, i.e., frequency is not the only parameter/dimension which characterizes the spectral opportunity. Space, time, transmission power, polarization, medium access and interference collectively shape the radio environment. The DSA mechanism employs one or more of these parameters to break the shackles of rigidity imposed by the command and control mechanism. Cognitive radios (CRs) are envisioned to be the key enablers for provisioning DSA. CRs are based on opportunistic exploitation of radio spectrum across one or more dimensions. Nevertheless, while the CR platform renders itself as a promising solution for improving connectivity, its suitability in the context of M2M communication is limited for two main reasons:

1. High cost: CRs employ sophisticated hardware to derive operational environment awareness and so naturally the radio platforms costs are higher when compared with dumb radio terminals. The radio platforms of MTDs in an M2M network will be embedded inside objects requiring both additional cost and form factors. Thus the radio platforms should be as simple as possible,

ideally comprising of a single chip on which a radio transceiver is integrated with the micro-controller unit (MCU). Manufacturers such as Texas Instrument, Nordic Semiconductor, Maxim and CSR are already providing such simple solutions.

2. Energy consumption and life-time: CR terminals often pay the cost of opportunism in terms of their higher energy consumption. More specifically, the operational environment awareness is derived from the inference process which consumes more energy as compared to simple radio platforms. For the wireless access applications, energy consumption is not considered as a design constraint due to the constant supply of power from the grid. Nevertheless, for M2M based applications energy-consumption of MTDs is of utmost importance. As discussed earlier, the radio platform is part of a variety of objects, most of them having no/limited access to the power running only on coin cell batteries. In this context, the cost of opportunism may be incurred in terms of the reduced operational life-time of these objects.

While object life-time is a critical aspect of design, the issue of so called 'green design' is further brought into play due to the predicted high volume of 'smart things'. Specifically, as predicted in a recent report by Ericsson [108], the CO_2 emissions due to a growing number of internet connected devices will increase from 800 Mtonnes to 1200 Mtonnes by 2020. In terms of net emissions, ICT will continue to maintain its 2% contribution to the global carbon foot-print. Nevertheless, according to the Intergovernmental Panel on Climate Change (IPCC), current emission trends are far from sustainable, requiring an exponential reduction to meet a 2°C rise in global temperature. In a recent survey by Cable News Network (CNN) it was estimated that a 2°C rise in global temperature will result in a 100 billion US dollar expense rise for addressing various challenges due to climate change. In summary, like all other sectors, ICT should exponentially reduce its energy consumption to operate in an eco-friendly manner. Thus in summary, for future deployment of 500 billion MTDs, a clean slate design is necessary to address both energy and spectral efficiency issues.

7.1.1. Design Attributes and Proposed Architecture

The grand challenges posed in the context of the cognitive M2M (CM2M) networks can be easily translated into design attributes/constraints. To summarize, the radio platform employed in the MTDs in a CM2M network should be: (i) simple yet

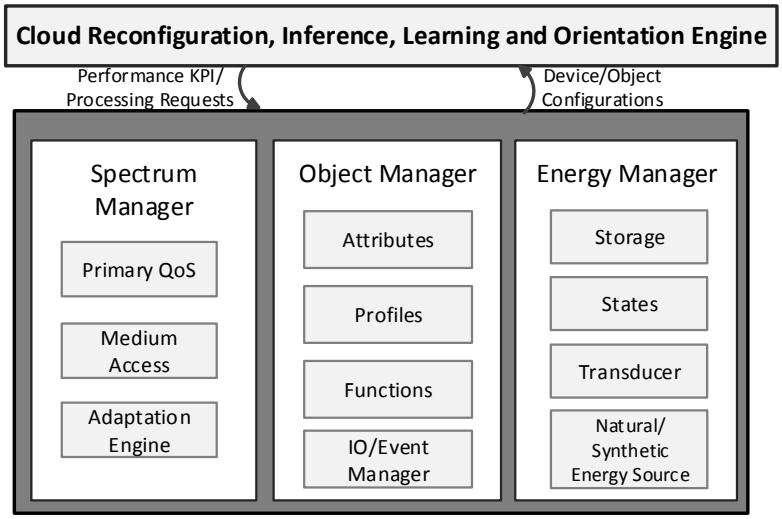


Figure 7.1.: Proposed Architecture for the Cognitive M2M Networks.

agile; (ii) spectrally efficient and (iii) low power with a miniaturized form factor. To satisfy these design attributes, the definition of cognition in the context of the M2M networks must be revisited. In particular, not only is spectral agility of prime importance, but power consumption awareness should also be embedded into the cognitive engine. We advocate that the cognitive engine must be equipped with a potential to harvest energy from ambient sources and in some cases from the objects themselves. For instance, consider smart door locks installed in modern houses. The radio transceivers on these locks can be powered using solar panels harvesting indoor ambient light from both natural and synthetic sources. Moreover, these locks can also harvest power from the mechanical motion of door itself. As smart objects have a very low-duty cycle, harvested energy provides a significant potential for designing self-sustainable so called 'zero-energy consumption' CM2M networks.

In this chapter, we propose a cloud enabled CM2M platform as depicted in Figure 7.1 to address the aforementioned challenges. From an object oriented programming approach it is well known that an object can be adequately described by its attributes and functionalities. These functionalists and attributes can be linked to external stimuli characterizing events. The behavior of the object in response to an external stimulus is defined by the device profile. External and internal stimuli may trigger interrupts which should be handled in accordance with device profile and current state. We propose that this object related functionality should be implemented in the so called 'object manager' which forms the central part of the CM2M

engine. The object engine coordinates with both the energy and spectrum managers to provide context awareness and indicate required quality-of-service (QoS) or quality-of-information (QoI) constraints. The object management life cycle can be simplified as most of the inference can be moved up to the centralized cloud processor. Thus objects can be made simpler by implementing basic look-up tables which map events, stimulus, attributes and functionality. Notice that the cloud based architecture provides flexibility of re-configuring the object management engine on the fly.

Spectrum and energy management engines are responsible for maximizing the spectral and energy efficiencies of a CM2M network. Unlike data-intensive applications such as cellular networks, where optimizing area spectral efficiency and load balancing are the critical tasks [109], the main purpose of an M2M network is to provide reliable interconnections between smart things. A number of MTDs need to communicate with a central aggregator for inferences, decisions and processing. We advocate the use of a cognitive underlay based spectrum access which requires only transmit power/medium access probability adaptation at the CM2M platforms. The interested reader is referred to [110] for a detailed discussion on exploitation of different degrees of freedom in cognitive underlay networks. In the case of a cognitive underlay mode of operation, it is important to know the spectrum availability and the probability of successful reception when the spectrum is utilized. Moreover, catering for the energy demands of the increased number of CM2M nodes is yet another important issue. The intrinsic advantage of the proposed spectrum access is that its implementation is simple and does not require additional sophisticated hardware. Based on the dynamics of the primary network, the cloud reconfiguration engine can reconfigure the spectrum access probability and the transmit power to guarantee that the QoS of the legacy network is not violated. This is to ensure a robust co-existence framework between the primary users (such as mobile users in cellular network) and the CM2M devices. The practical implementation of such a spectrum access would require a simple look up table at each device so that CM2M platforms do not lose their cost-effectiveness or the form factor by implementing the proposed cognitive access strategy. In the subsequent discussion, we develop an analytical framework for quantifying the performance of the large scale CM2M network by considering a reference scenario under the proposed architecture.

7.1.2. Outline

The outline of the rest of the chapter is as follows: In section 6.2, we introduce our reference scenario and mathematical preliminaries which are employed throughout the remainder of this chapter. Additionally, we introduce the proposed energy harvesting model and detail its dynamics. In section 7.3, we define two CM2M performance metrics called (i) the energy success probability and (ii) spectral success probability for the considered CM2M network. We consider a realistic model to compute the harvested power and develop a stochastic model for the energy success probability which can be treated as a proxy for self-sustainability of a CM2M interconnect. Considering, the proposed spectrum management engine, the co-existence constraint is enforced on a CM2M network to provision spectrum access in underlay mode. The co-existence constraint is defined in terms of the average outage probability for an arbitrary user equipment (UE) scheduled in the downlink[1] of the primary network. Notice that in context of the considered primary network, i.e., cellular network, the outage probability is the complement of the coverage probability. Moreover, the averaging is performed over the location of the UE since it is unknown a priori (see [31]). Consequently, the definition of outage for the primary network differs from the one employed in [110]. With the outage analysis of the primary network, we obtain the maximum allowable spectrum access probability for the considered CM2M network under Slotted-ALOHA type protocol[2]. The maximum allowable access probability ensures that the primary user's enforced co-existence constraint is not violated. Section 7.4 combines the aforementioned performance metrics and introduces a new metric called the overall success probability. The introduced metric provides interesting insight for optimizing the transmit power employed by MTDs to strike a balance between the spectrum access and the energy usage. Section 6.5 concludes the chapter and summarizes the important future directions.

[1]Notice that the analysis is general and is not affected by considering the uplink of the primary cellular network.

[2]In this chapter, we consider the Slotted-ALOHA type access strategy for a CM2M network. However, the spectral access probability computed here, can be effectively mapped to the carrier sensing threshold for a CSMA/CA type protocol.

7.2. System Model

We consider a large scale energy harvesting CM2M network co-existing with a primary cellular network as shown in Figure 7.2. We consider that the MTDs are furnished with solar panels to harvest ambient energy. The MTDs with sufficient harvested energy, employ the cognitive underlay approach and utilize the same resource blocks as the primary base stations (BSs) to communicate with the central CM2M aggregator. We assume that full frequency reuse is employed in the primary cellular network for maximizing the area spectral efficiency. The cognition employed at the MTDs shapes transmission parameters such that the aggregate interference at the primary UE remains below a certain threshold to satisfy the desired QoS requirement. Consequently, only a fraction of MTDs are activated during each transmission sub-frame. From the context of the CM2M network, the accumulation of interference from the primary BSs and other MTDs plays a critical role in determining whether the transmissions can be successfully decoded at the corresponding aggregator. Thus both the opportunity for the transmission and the probability of success when provided with such opportunity collectively define the performance of the CM2M network for the considered reference scenario.

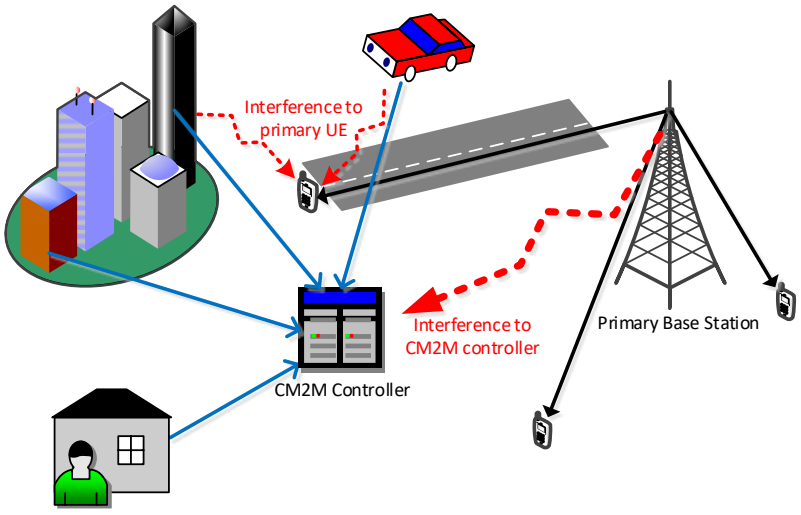


Figure 7.2.: Top-level diagram showing the coexistence of a CM2M network with the primary cellular network in spectrum underlay mode.

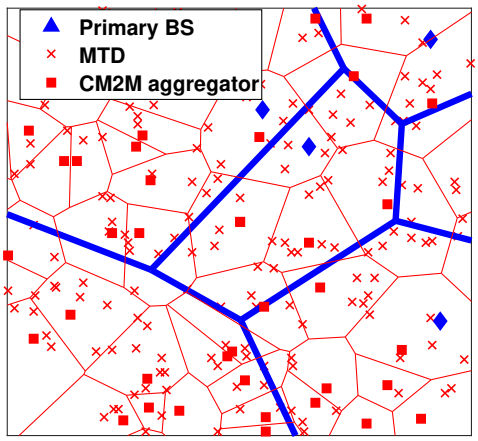


Figure 7.3.: Realization of the CM2M network in the underlay.

7.2.1. Spatial Model

The spatial distribution of the primary BSs CM2M aggregators and MTDs is captured by the independent homogeneous Poisson point processes (HPPPs) Φ_b, Φ_c and Φ_m respectively, with λ_i, $i \in [b, c, m]$ being the average number of these entities in a unit area. The selection of HPPPs to model the location of the primary BSs and UEs is widely studied in the literature [31]. In the context of CM2M network, the HPPP assumption for the MTDs is quite reasonable as the MTDs are expected to be densely deployed with considerable irregularity. Figure 7.3 depicts a realization of the network under these considerations. The primary UEs (not shown in the figure), as we average the performance metric over all possible locations of the UE in each cell. Each UE associates itself with the nearest primary BS. The coverage areas of BSs and CM2M aggregators result in separate Voronoi tessellations on $\mathbb{R}^2$ [111]. Without any loss of generality a probe CM2M aggregator can be placed at the origin[3] and the signal from an arbitrary active MTD is considered as the intended signal. All the other active MTDs and BSs are considered as interferers for the transmission received on the probe CM2M aggregator.

7.2.2. Spectrum Access Strategy

It is assumed that, the MTDs employ a random access strategy similar to the Slotted ALOHA MAC protocol to schedule their transmissions over a shared medium

[3]This follows from the Slivnyak's theorem and the palm distribution of HPPPs [36].

for communicating with the corresponding aggregators. More specifically, at an arbitrary time instant, the MTDs can be classified into two distinct groups, i.e., nodes which are granted spectrum access and those whose transmissions are deferred. If p_m denotes the spectrum access probability (SAP) for an arbitrary MTD $\boldsymbol{x} \in {\Phi_m}^4$, then the set of active users under the considered spectrum access paradigm also forms a HPPP $\Phi_m^{\{TX\}} = \{\, \boldsymbol{x} \in \Phi_m : \mathbb{1}(\boldsymbol{x}) = 1\}$ with density $\lambda_m p_m$, where $\mathbb{1}(\boldsymbol{x})$ denotes a Bernoulli random variable and is independent of Φ_m.

7.2.3. Channel Model

We adopt a standard power loss propagation model with the environment dependent path loss exponent. All wireless communication links assume Rayleigh flat fading. So, the overall channel power gain is represented as $hr^{-\alpha}$, where $h \sim \exp(1)$ follows a unit-mean exponential distribution representing the received power gain under Rayleigh fading and r is the distance between the transmitter and the receiver. We focus on the interference limited scenario and assume that the effects of the thermal noise on the network's performance are negligible.

7.2.4. Energy Harvesting

Harvesting energy from natural (solar, wind or vibration) and synthesized (microwave power transfer) sources is envisioned as a key enabler for realizing green wireless networks. The choice of a suitable renewable source depends on a number of factors including, the availability of a particular source, ease of harvesting from it and most importantly the energy demand of the application. A detailed comparison of the power densities from various sources is provided in [112]. As solar energy is currently the most scalable renewable resource, from powering small indoor sensors to huge buildings, it naturally becomes a suitable candidate to cater for the contrasting power requirements of things operating under the umbrella of a CM2M network. Consequently, in the considered reference scenario, we assume that MTDs are furnished with solar panels to harvest ambient energy. We begin our discussion on the solar energy harvesting in a outdoor setting[5].

[4] With a slight abuse of notation, $\boldsymbol{x} \in \mathbb{R}^2$ is employed to refer to the node's location as well as the node itself.

[5] Notice that the model remains same for the indoor setting except for the fact that the output power is attenuated by a factor of 10-100. This is because, indoor panels cannot harvest the direct component of a solar energy field and must rely on the diffuse component.

Solar Energy Harvesting

The solar power density measured just above the earth's atmosphere is about 1367 W/m^2 and it is commonly referred to as extra-terrestrial (ET) irradiance. The power density at the surface of the solar panel, also known as insolation, is much less than the ET irradiance and it depends on a number of geometrical and astronomical aspects such as the earth's elliptical orbit around the sun, earth's tilted axis of rotation, sun's position from directly overhead the panel, the panel's location and its angle of tilt[6]. The aforementioned factors are deterministic and can be quantified. However, the effect of both cloud cover and atmosphere induces randomness in the observed insolation energy field.

Hourly okta	Distribution of clearness index, $f_{\kappa_c}(.)$
0	$\mathcal{N}(0.99, 0.08)$
≤ 6	$\mathcal{N}(0.6784, 0.2046)$
7	$\mathcal{W}(0.5577, 2.4061)$
8	$\mathcal{G}(3.5624, 0.0867)$

Table 7.1.: Distribution of clearness index for various levels of cloud cover at 12:00 pm.

While the atmospheric profile of temperature, concentration of water vapor, ozone, aerosol gases, other particles and the surface albedo (measure of reflectivity) remains fairly constant during the day, the cloud cover may fluctuate considerably. Meteorologists classify the cloud density in terms of oktas, which are increasing levels of cloud cover with values from $1-8$. The extreme scenarios of the absence of clouds on a clear day and completely overcast conditions or night time are represented by okta 0 and okta 9 respectively. For a particular value of okta or a range of oktas, the atmospheric transmission of the solar energy is commonly parametrized by the clearness index (κ_c), which takes the form

$$\kappa_c = \frac{E}{E_c}, \tag{7.1}$$

where, E_c is the theoretical clear-sky (cloud free) insolation and E is the instantaneous insolation observed at the panel. This clearness index is a random variable encapsulating the effect of atmospheric variations. The authors in [113] obtained

[6]For medium to average size cities, variations in longitude and latitude are not significant. Thus neglecting the environmental randomness, the ET irradiance does not vary significantly over the spatial scale of neighborhood.

the distributions for κ_c for a range of oktas as described in Table 7.1. They use the data provided by the UK Met Office Integrated Archive System (MIDAS) for hourly values of E, and the corresponding cloud okta for the year 2012. For the computation of realistic values of E_c, they used the DISTORT radiative transfer solver [114] in the libRadtran package [115] with the actual data for the atmospheric conditions. Maximum likelihood estimation (MLE) is used to fit the distributions. To this end, we employ these distributions in our solar panel model to observe the effect of the cloud intensity on the harvested energy at noon. This time of the day is chosen as maximum energy can be harvested at noon in clear sky conditions.

7.3. Performance Analysis of an Energy Harvesting Empowered CM2M Network

Before moving on to the mathematical analysis of the performance determining variables, we define the two key metrics as follows:

Definition 7.1. Assuming that the harvested energy is utilized for scheduling the transmissions, we define 'energy success probability' as a metric to measure the availability of energy for a certain desired transmit power P_m. The dynamics of the energy harvester and thus management engine of a CM2M platform can be completely characterized in terms of energy success probability given as [116]

$$\mathcal{S}^{\{e\}}(P_m) = \mathbb{P}\left[P_{PV} \geq P_m\right], \tag{7.2}$$

where P_{PV} is the power harvested by the photovoltaic (PV) solar panel. Notice that the definition can be easily modified to cater for the case where energy storage of finite size is present. However, in this thesis, we do not consider storage and thus harvested energy is employed for transmission scheduling in an instantaneous manner. Such consideration results in a lower bound on actual performance which can be attained by exploiting the energy storage. Furthermore, we assume that transmit energy is the dominant factor in terms of energy consumption and that the power consumption foot-print of the transceiver circuitry is negligible. The circuit power consumption can be accommodated by performing an affine transformation on P_m.

We define below the term 'spectrum success probability' as the probability that an MTD is able to access the spectrum and subsequently successfully communicate

with the CM2M aggregator while satisfying the UE's desired QoS constraint.

$$\mathcal{S}^{\{s\}}(P_m) = \mathbb{E}_{R_m}\left[\mathbb{P}\left[SIR_m(p_m^{max}, r) \geq \theta_m \mid R_m\right]\right], \tag{7.3}$$

where R_m is the random distance between an MTD and its nearest aggregator and θ_m is the modulation dependent decoding threshold which is selected to satisfy a certain desired frame error rate. Notice that the MTDs employ the maximum SAP p_m^{max} to satisfy the UE's QoS requirements.

7.3.1. Energy Success Probability in Harvesting Empowered CM2M

We calculate the harvested power using a well-known single diode model for a PV module [117]. The output panel current for this model can be written as

$$I_{PV} = I_{sc}\left[1 - \nu_3\left[\exp\left(\frac{V_{PV}}{\nu_4 V_{oc}}\right) - 1\right]\right], \tag{7.4}$$

where $\nu_3 = \left(1 - \frac{I_{MPP}}{I_{sc}}\right)\exp\left(\frac{V_{MPP}}{\nu_4 V_{oc}}\right)$ and $\nu_4 = \left(\frac{V_{MPP}}{V_{oc}} - 1\right)/\ln\left(1 - \frac{I_{MPP}}{I_{sc}}\right)$. The current generated by the module depends on several parameters: (i) short circuit current I_{sc}; (ii) open circuit voltage V_{oc}; (iii) maximum power point voltage V_{MPP} and (iv) maximum power point current I_{MPP}. These parameters can be expressed as functions of ambient temperature and global horizontal irradiance as follows:

$$I_{sc} = I_{scs} \times \kappa_c \times \left[1 + \nu_1(T - T_s)\right],$$

$$V_{oc} = V_{ocs} + \nu_2(T - T_s),$$

$$I_{MPP} = I_{MPPS} \times \kappa_c \times \left[1 + \nu_1(T - T_s)\right],$$

$$V_{MPP} = V_{MPPS} + \nu_2(T - T_s),$$

where I_{scs}, V_{ocs}, I_{MPPS}, and V_{MPPS} are defined for standard conditions, $T_s = 25°C$ with ν_1 and ν_2 being the current and the voltage coefficients. These parameters are generally provided in the data sheet of a PV module. From (7.4), the output power of the PV panel can be computed as a function of voltage as $P_{PV} = I_{PV}V_{PV}$. For the ease of tractability, we consider standard conditions ($T = T_s$) for the panel for

Hourly okta	Energy success probability $\left[\mathcal{S}^{\{e\}}\left(P_m\right) = 1 - F_{\kappa_c}\left(\frac{P_m}{\zeta}\right)\right]$
0	$\frac{1}{2}\text{erfc}\left(\frac{P_m - \zeta a_1}{\sqrt{2}\zeta b_1}\right), a_1 = 0.99, b_1 = 0.08.$
≤ 6	$\frac{1}{2}\text{erfc}\left(\frac{P_m - \zeta a_2}{\sqrt{2}\zeta b_2}\right), a_2 = 0.6784, b_2 = 0.2046.$
7	$\exp\left(-\left(\frac{P_m}{\zeta b_3}\right)^{a_3}\right), a_3 = 2.4061, b_3 = 0.5577.$
8	$\frac{1}{\Gamma(a_4)}\Gamma\left(a_4, \frac{P_m}{\zeta b_4}\right), a_4 = 3.5624, b_4 = 0.0867.$

Table 7.2.: Energy success probability for various levels of cloud cover in oktas.

our further analysis. The equation for the harvested power (in Watts) simplifies to

$$P_{PV} = \kappa_c V_{PV}\, I_{scs}\left[1 - \nu_3\left\{\exp\left(\frac{V_{PV}}{\nu_4 V_{ocs}}\right) - 1\right\}\right]. \tag{7.5}$$

Using definition 7.1 and the expression for the output panel power available for transmission, the energy success probability is given as

$$\mathcal{S}^{\{e\}}\left(P_m\right) = \mathbb{P}\left\{\kappa_c \geq \zeta^{-1} P_m\right\}, \tag{7.6}$$

where $\zeta = V_{PV}\, I_{scs}\left[1 - \nu_3\left\{\exp\left(\frac{V_{PV}}{\nu_4 V_{ocs}}\right) - 1\right\}\right]$.

It is evident from (7.6) that $\mathcal{S}^{\{e\}}\left(P_m\right)$ follows the same distribution as κ_c, only scaled by a factor of ζ. The expressions for $\mathcal{S}^{\{e\}}\left(P_m\right)$ for various oktas are presented in Table 7.2, where a_i and $b_i, i \in [1, 2]$ are the mean and standard deviation of the normal distribution of κ_c for $okta \in [0 - 6]$, a_3 and b_3 are the shape and scale parameters of the Weibull distribution of κ_c for okta 7, a_4 and b_4 are the shape and scale parameters of the Gamma distribution of κ_c for okta 8 and $\Gamma(a, x) = \int_x^\infty t^{a-1}\exp(-t)dt$ is the incomplete Gamma function.

7.3.2. Spectrum Access Success Probability in CM2M network

To compute the spectrum success probability, we first consider the outage constraint enforced by the primary network on the transmitting MTDs. The received SIR of a

typical UE in an interference limited scenario is characterized as

$$\begin{aligned} SIR_u(r) &= \frac{h_u l(\|\boldsymbol{x}\|)}{\sum_{i\in\Phi_b\setminus b(o,\|\boldsymbol{x}\|)} h_i l\left(\|\boldsymbol{x}_i\|\right) + \sum_{j\in\Phi_m^{\{TX\}}} \varsigma g_j l\left(\|\boldsymbol{x}_j\|\right)}, \\ &= \frac{h_u l(r)}{I_u(r)+\varsigma I_m} = \frac{h_u l(r)}{I_{tot}}, \end{aligned} \tag{7.7}$$

where $h_u, h_i, g_j \sim \exp(1)$ are random variables capturing the effect of Rayleigh fading; $l(r) = r^{-\alpha}$ is the path-loss function; $\varsigma = \frac{P_m}{P_b}$ is the transmit power ratio of the CM2M and primary networks and $r = \|\boldsymbol{x}\|$ is the distance between the primary BS and the UE. The primary user's QoS constraint can be expressed in terms of its desired SIR threshold θ_u and an outage probability threshold $\gamma_{out}^{\{u\}}$ as

$$\mathbb{P}_{out}^{\{p\}}\left(P_m, p_m\right) = \mathbb{E}_{R_u}\left[\mathbb{P}\left[SIR_u(r) \le \theta_u \mid R_u\right]\right] \le \gamma_{out}^{\{u\}}, \tag{7.8}$$

where R_u is the random distance between the UE and its corresponding BS in a reference Voronoi cell. Notice that the primary user's outage probability is coupled with the aggregate interference generated by the CM2M network. Consequently, secondary access is limited subject to the constraint $\gamma_{out}^{\{u\}}$ in (7.8).

Proposition 7.1. *(Maximum permissible SAP for an MTD.) Given the QoS of a UE in terms of the desired SIR threshold θ_u and its maximum tolerable link outage $\gamma_{out}^{\{u\}}$, the SAP for the MTDs which can operate in a concurrent manner without violating the co-existence constraint can be quantified as*

$$p_m^{max} = \frac{\lambda_b}{\mathcal{S}^{\{e\}}\left(P_m\right)\lambda_m} \frac{\operatorname{sinc}\left(\delta\right)}{\theta_u^{\delta}\varsigma^{\delta}} \left[\frac{\gamma_{out}^{\{u\}}}{\left(1-\gamma_{out}^{\{u\}}\right)}\left(1-\mathcal{F}(\theta_u,\delta)\right)\right], \tag{7.9}$$

where $\delta = 2/\alpha$ and $\mathcal{F}(\theta_u,\delta) = \frac{\delta\theta_u}{1-\delta}\,{}_2F_1(1,1-\delta;2-\delta;-\theta_u)$. Here ${}_2F_1(a,b;c;z)$ is the Gauss Hypergeometric function [118]. For $\alpha = 4$ ($\delta = 1/2$), $\mathcal{F}(\theta_u,\delta) = \sqrt{\theta_u}\arctan\left(\sqrt{\theta_u}\right)$ and the maximum permissible SAP simplifies to

$$p_m^{max} = \frac{2\lambda_b}{\mathcal{S}^{\{e\}}\left(P_m\right)\lambda_m \pi\sqrt{\theta_u\varsigma}} \left(\frac{\gamma_{out}^{\{u\}}}{\left(1-\gamma_{out}^{\{u\}}\right)}\left(1-\sqrt{\theta_u}\arctan\left(\sqrt{\theta_u}\right)\right)\right), \tag{7.10}$$

Proof. Following the steps in [31] and [110], the outage probability of the primary

UE can be written as

$$\mathbb{P}_{out}^{\{p\}}(P_m, p_m) = 1 - \int_{r>0} \mathcal{L}_{I_{tot}}(\theta_u r^\alpha) \exp(-\pi\lambda_b r^2)\, 2\pi\lambda_b r\, dr. \tag{7.11}$$

For the considered scenario, $\mathcal{L}(s) = \mathcal{L}_{I_u(r)}(s) \times \mathcal{L}_{I_m}(\varsigma s)|_{s=\theta_u r^\alpha}$ [7], and $\mathcal{L}_{I_u(r)}(s)$ and $\mathcal{L}_{I_m}(\varsigma s)$ are respectively the Laplace transforms of the interference on the aggregate interference on the UE from the other-cell co-channel interferers and the spectrum sharing MTDs. Using the well-known definition of the generating functional of a HPPP in [36], the following Laplace transforms can be evaluated as

$$\begin{aligned}
\mathcal{L}_{I_u}(\theta_u r^\alpha) &= \exp\left(-2\pi\lambda_b \int_r^\infty (1 - \mathbb{E}_h[\exp(-\theta_u r^\alpha h v^{-\alpha})]) v\, dv\right) \\
&= \exp\left(-2\pi\lambda_b \int_r^\infty \frac{v}{1 + (\theta_u r^\alpha)^{-1} v^\alpha}\, dv\right) \\
&= \exp\left(-\pi\lambda_b r^2 \mathcal{F}(\theta_u, \delta)\right)
\end{aligned} \tag{7.12}$$

and

$$\begin{aligned}
\mathcal{L}_{I_m}(\varsigma\gamma_{th}^{\{p\}} r^\alpha) &= \exp\left(-2\pi\lambda_m^{eff} \int_0^\infty (1 - \mathbb{E}_g[\exp(-\varsigma\theta_u r^\alpha g u^{-\alpha})]) u\, du\right) \\
&= \exp\left(-\frac{\pi\lambda_m p_m \mathcal{S}^{\{e\}}(P_m)\, \theta_u^\delta \varsigma^\delta}{\mathrm{sinc}(\delta)} r^2\right)\},
\end{aligned} \tag{7.13}$$

where $\delta = 2/\alpha$, $\lambda_m^{eff} = \lambda_m p_m \mathcal{S}^{\{e\}}(P_m)$ is the effective density of the active MTDs and $\mathcal{F}(\theta_u, \delta) = \theta_p^\delta \int_{\theta_u^{-\delta}}^\infty (1 + u^{\frac{1}{\delta}})^{-1} du$. The solution for this integral in terms of the hypergeometric function can be obtained by substituting $y = u^{\frac{1}{\delta}}$ and using Eq. 3.194-2.6 from [119].

Notice the difference in the limits of integration in (7.12) and (7.13). This is due to the fact that the other-cell interferer will be separated by a minimum distance of r from the tagged UE while such a constraint is not enforced on the MTDs. Furthermore, the density λ_m of MTDs inside the cell is thinned by the SAP and the energy success probability $\mathcal{S}^{\{e\}}(P_m)$ as only the MTDs with sufficient harvested energy will be able to transmit. Inserting the above two expressions into (7.11) and solving the integral provides us with the outage probability of the UE. Enforcing the outage constraint $\gamma_{out}^{\{u\}}$ as in (7.8) and inverting the equation for p_m concludes the proof. □

Implementation Note: From (7.9), we notice that p_m^{max} is function of both the primary and the M2M network parameters. Thus cloud based coordination

[7] This is because of the independence of the point processes Φ_b and Φ_m.

is beneficial to dynamically reconfigure the SAP based on the prevalent network conditions. More specifically, (7.9) can be implemented in the cloud re-configuration, inference, learning and adaptation engine to proviso self-organization.

Proposition 7.2. *(Spectrum success probability of the MTDs.) The probability that an MTD is able to access the available spectrum and successfully communicate with its aggregator is given as*

$$\mathcal{S}^{\{s\}}\left(P_m\right) = \left(1 + \frac{\lambda_b}{\lambda_c}\varsigma^{-2\delta}\left[\frac{\gamma_{out}^{\{u\}}}{\left(1-\gamma_{out}^{\{u\}}\right)}\left(1-\mathcal{F}(\theta_u,\delta)\right)\right] + \frac{\lambda_b\varsigma^{-\delta}\theta_m^{\delta}}{\lambda_c \mathrm{sinc}\left(\delta\right)}\right)^{-1}. \quad (7.14)$$

Proof. The proof for (7.14) follows similar steps as in the proof of (7.11). The received SIR at the CM2M aggregator placed at the origin is given as

$$\begin{aligned} SIR_m &= \frac{g_c l(r)}{\sum_{i\in\Phi_m^{\{TX\}}\setminus b(o,r)} g_i l\left(\|\boldsymbol{x}_i\|\right) + \sum_{j\in\Phi_u}\varsigma^{-1}h_j l\left(\|\boldsymbol{x}_j\|\right)}, \\ &= \frac{h_c l(r)}{\varsigma^{-1}I_u + I_m(r)} = \frac{h_c r^{-\alpha}}{I_{tot}}. \end{aligned} \quad (7.15)$$

Using the definition 7.3, the spectrum success probability is given as

$$\mathbb{P}\left\{SIR_m \geq \theta_m\right\} = \int_{r>0} \mathcal{L}_{I_{tot}}(\theta_m r^{\alpha}) \exp(-\pi\lambda_c r^2)\, 2\pi\lambda_c r\, dr. \quad (7.16)$$

In this case, $\mathcal{L}_{I_{tot}}(s)$ will be

$$\mathcal{L}_{I_{tot}}(s) = \mathcal{L}_{I_u}(\varsigma^{-1}s) \times \mathcal{L}_{I_m}(s)|_{s=\theta_m r^{\alpha}}.$$

The Laplace transforms are given by the following expressions

$$\begin{aligned} \mathcal{L}_{I_u}(\varsigma^{-1}\theta_m r^{\alpha}) &= \exp\left(-2\pi\lambda_b\int_0^{\infty}(1-\mathbb{E}_h[\exp(-\theta_m r^{\alpha}hv^{-\alpha})])v\, dv\right) \\ &= \exp\left(-\frac{\pi\lambda_b\theta_m^{\delta}\varsigma^{-\delta}}{\mathrm{sinc}\left(\delta\right)}r^2\right) \end{aligned} \quad (7.17)$$

and

$$\begin{aligned} \mathcal{L}_{I_m}(\theta_m r^{\alpha}) &= \exp\left(-2\pi\lambda_m^{eff}\int_0^{\infty}(1-\mathbb{E}_g[\exp(-\theta_m r^{\alpha}gu^{-\alpha})])u\, du\right) \\ &= \exp\left(-\frac{\pi\lambda_m p_m \mathcal{S}^{\{e\}}\left(P_m\right)\theta_m^{\delta}\varsigma^{-\delta}}{\mathrm{sinc}\left(\delta\right)}r^2\right). \end{aligned} \quad (7.18)$$

The lower limits of integration for both the Laplace transforms are zero because of

no interference guard zone. □

An interesting observation from (7.14) is that the spectrum access success for MTDs is independent of the intensity of MTDs λ_m and the energy success probability $\mathcal{S}^{\{e\}}(P_m)$. This is because, as $\mathcal{S}^{\{e\}}(P_m)$ or λ_m increase, the maximum activation probability p_m^{max} decreases accordingly resulting in the activation of fewer MTDs to satisfy the maximum interference constraint.

7.4. Overall Success Link Probability

Employing the existing analytical characterizations, we define the unified performance metric for the energy harvesting empowered CM2M network. Intuitively, an MTD will only be able to successfully communicate with its nearest aggregator if the following conditions are met:

1. The harvested energy is greater than the required transmit power;
2. The MTD is allowed to access spectrum while satisfying the primary UE's QoS constraint; and
3. The ratio of the received signal's power to the interference from all the other MTDs and the primary BSs is greater than the desired SIR threshold θ_m.

The performance metrics $\mathcal{S}^{\{s\}}(P_m)$ and $\mathcal{S}^{\{e\}}(P_m)$ derived in the previous section are both important in characterizing the performance of a CM2M network. Nevertheless, $\mathcal{S}^{\{e\}}(P_m)$ only signifies the energy availability and $\mathcal{S}^{\{s\}}(P_m)$ signifies the spectrum availability. For a more comprehensive analysis, there is a need to combine these two metrics such that all the factors governing the performance of a CM2M network can be captured by a single metric. Hence, we introduce the overall link success probability metric as follows

$$\mathcal{P}_{suc}^{tot}(P_m) = \mathcal{S}^{\{s\}}(P_m) \times \mathcal{S}^{\{e\}}(P_m), \tag{7.19}$$

where $\mathcal{P}_{suc}^{tot}(P_m)$ is the overall success probability, which depends on the communication aspects, MTD transmit power, and the solar panel parameters. Using the expressions in Table 7.2 and with a few manipulations in (7.14), the overall success probability in (7.19) is given as

For okta ≤ 6 :

$$\mathcal{P}_{suc}^{tot}(P_m) = \frac{1/2}{1 + c_1 P_m^{-2\delta} + c_2 P_m^{-\delta}} \text{erfc}\left(\frac{P_m - \zeta a_i}{\sqrt{2}\zeta b_i}\right), i \in [1,2]. \quad (7.20)$$

Here, $i = 1$ represents okta 0, while $i = 2$ represents $1 \leq$ okta ≤ 6.

For okta 7 :

$$\mathcal{P}_{suc}^{tot}(P_m) = \frac{1}{1 + c_1 P_m^{-2\delta} + c_2 P_m^{-\delta}} \exp\left(-\left(\frac{P_m}{\zeta b_3}\right)^{a_3}\right). \quad (7.21)$$

For okta 8 :

$$\mathcal{P}_{suc}^{tot}(P_m) = \frac{1}{1 + c_1 P_m^{-2\delta} + c_2 P_m^{-\delta}} \left[\frac{\Gamma\left(a_4, \frac{P_m}{\zeta b_4}\right)}{\Gamma(a_4)}\right], \quad (7.22)$$

where $c_1 = \frac{\lambda_b}{\lambda_c} P_b^{2\delta} \left[\frac{\gamma_{out}^{\{u\}}}{(1-\gamma_{out}^{\{u\}})}(1 - \mathcal{F}(\theta_u, \delta))\right]$ and $c_2 = \frac{\lambda_b P_b^{\delta} \theta_m^{\delta}}{\lambda_c \text{sinc}(\delta)}$.

It follows from (7.20), (7.21) and (7.22) that for all values of oktas, the effect of modulation dependent decoding thresholds θ_m and θ_u, path loss exponent α, density of the MTDs λ_m and primary BS density λ_b remains the same on the overall success probability. This is because the cloud cover only affects the availability of the energy and the spectrum success probability remains unchanged. However, from (7.19), we observe that both the terms $\mathcal{S}^{\{s\}}(P_m)$ and $\mathcal{S}^{\{e\}}(P_m)$ depend on the device level transmit power P_m, which could be adapted to optimize the CM2M performance. In the following section, we see how P_m could be adapted to achieve better connectivity for the MTDs and also, what impact will the densification of MTDs have on the required transmit power and the coverage characteristics.

7.4.1. Optimal transmit power

As the transmit power P_m of the MTDs increases (and in turn ς increases keeping P_b constant), $\mathcal{S}^{\{s\}}$ also increases as the maximum permissible SAP (p_m^{max}) decreases. This implies that transmission opportunities of MTDs will diminish due to interference protection implemented by the cloud aggregator to guarantee the primary user's QoS requirement. Alternatively, increasing P_m results in a drop in $\mathcal{S}^{\{e\}}$ as more and more harvested power is required to meet the energy demand. However, the low transmit power employed by a CM2M platform may not be able to guarantee the required QoS or QoI at each MTD. Consequently, the transmit power must

Parameter	Description	Value
α	Path loss exponent	4
λ_c/λ_b	Ratio between CM2M aggregator and BS densities	100
$\gamma_{out}^{\{u\}}$	Primary UE SIR outage threshold	0.2
P_b	BS transmit power	32 dBm
θ_m, θ_u	MTD and UE SIR threshold	[23, 23] dBm
I_{scs}, I_{MPPS}	Short circuit, maximum power point current	[15,13.3] mA
V_{ocs}, V_{MPPS}	Open circuit, maximum power point voltage	[1.89, 1.5] V [120]

Table 7.3.: List of parameters used for plotting results.

be optimized through considering energy and spectral success. The inverse relationship between the two terms can be observed in Figure 7.4. Since the overall success probability is the product of these two terms, there must exist an optimal power which maximizes the overall success probability. We see in Figure 7.5 that as the cloud okta increases and consequently the harvested energy decreases, the overall success probability decreases and for a fixed value of cloud okta, there always exists an optimal transmit power which maximizes the overall success probability.

Proposition 7.3. *(Optimal Transmit Power) The optimal transmit power can be numerically evaluated by solving the following equations for* P_m^{opt}.

For okta ≤ 6 ***:***

$$\mathcal{S}^{\{s\}'}\left(P_m^{opt}\right)\left[erf\left(\frac{P_m^{opt}-\zeta a_i}{\sqrt{2}\zeta b_i}\right)+1\right] - \frac{2}{\sqrt{2\pi}\zeta b_i\left(\mathcal{S}^{\{s\}}\left(P_m^{opt}\right)\right)^{-1}} exp\left(-\left(\frac{P_m^{opt}-\zeta a_i}{\sqrt{2}\zeta b_i}\right)^2\right)=0, i\in[1,2]. \tag{7.23}$$

For okta 7 ***:***

$$\mathcal{S}^{\{s\}'}\left(P_m^{opt}\right)\exp\left(-\left(\frac{P_m^{opt}}{\zeta b_3}\right)^{a_3}\right)-\frac{a_3\left(P_m^{opt}\right)^{a_3-1}}{(\zeta b_3)^{a_3}\left(\mathcal{S}^{\{s\}}\left(P_m^{opt}\right)\right)^{-1}}\exp\left(-\left(\frac{P_m^{opt}}{\zeta b_3}\right)^{a_3}\right)=0.$$

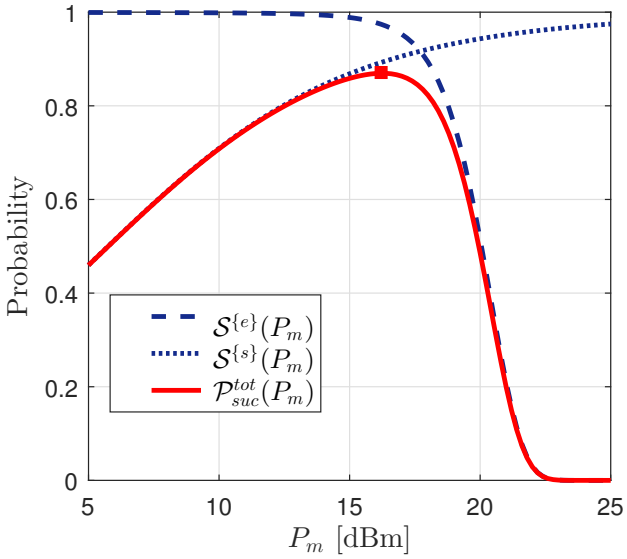


Figure 7.4.: Existence of an optimal power which maximizes the overall success probability, Eq. (7.6), (7.14) and (7.20): $1 \leq okta \leq 6$.

For okta 8 ***:***

$$\mathcal{S}^{\{s\}'}\left(P_m^{opt}\right)\left[\frac{\Gamma\left(a_4, \frac{P_m^{opt}}{\zeta b_4}\right)}{\Gamma(a_4)}\right] - \frac{\left(P_m^{opt}\right)^{a_4-1}\left(\zeta b_4\right)^{-a_4}}{\Gamma(a_4)\left(\mathcal{S}^{\{s\}}\left(P_m^{opt}\right)\right)^{-1}}\exp\left(-\frac{P_m^{opt}}{\zeta b_4}\right) = 0$$

Proof. The optimal transmit power (P_m^{opt}) is the solution of

$$\frac{\partial\, \mathcal{P}_{suc}^{tot}(P_m)}{\partial\, P_m} = 0 \tag{7.24}$$

From (7.24), we obtain the partial derivatives of $\mathcal{P}_{suc}^{tot}$ for various oktas, which can be numerically solved to obtain the optimal transmit power of the MTDs. □

7.4.1.1. Discussion

1. As seen from Figures 7.4 and 7.5, there always exists an optimal transmit power point which maximizes the overall success probability. This is due the fact that as the transmit power increases, the maximum SAP (p_m^{max}) reduces to activate fewer MTDs and the SIR increases and hence the probability of successful transmission increases. However, as the desired transmit power goes high, the available harvested energy becomes insufficient. The optimal transmit power point may lie in either the spectrum limited regime, i.e. when $\mathcal{S}^{\{e\}} > \mathcal{S}^{\{s\}}$ or

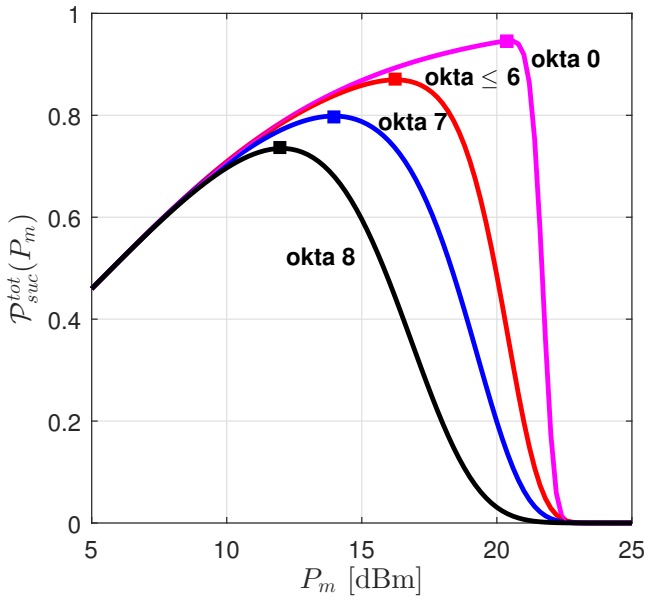


Figure 7.5.: Overall success probability with transmit power adaptation for various okta values, Eq. (7.20)-(7.22).

the energy limited regime, i.e. when $\mathcal{S}^{\{s\}} > \mathcal{S}^{\{e\}}$. The location of the optimal point depends solely on the solar panel parameters and other network related parameters that affect the slope of $\mathcal{S}^{\{e\}}$ and $\mathcal{S}^{\{s\}}$. As illustrated in Figure 7.4, the decrease in $\mathcal{S}^{\{e\}}$ is more rapid than the increase in $\mathcal{S}^{\{s\}}$ for the given set of parameters. The optimal power point lies in the spectrum limited regime as the overall success probability $\mathcal{P}_{suc}^{tot}$ follows the same trend as $\mathcal{S}^{\{s\}}$ as long as $\mathcal{S}^{\{e\}}$ is close to one. When $\mathcal{S}^{\{e\}}$ drops, $\mathcal{P}_{suc}^{tot}$ enters the energy limited regime and decreases accordingly.

2. Cloud cover plays an immensely important role in determining the performance of MTDs in a CM2M network. Comparing the clear sky (okta 0) and heavily cloudy conditions (okta 8) in Figure 7.5, a drop of almost 20% in the maximum overall success probability is observed while for partially cloudy conditions, this drop is fairly reasonable (around 8% and 15% for okta 6 and less and okta 7 respectively). Not only does the cloud cover affect the maximum overall success probability, it also changes the optimal transmit power. This can be intuitively explained by the fact that as the sky becomes clearer, a higher amount of energy can be harvested and also consumed in achieving a high spectral coverage. Thus, the a higher optimal power can be adopted when the

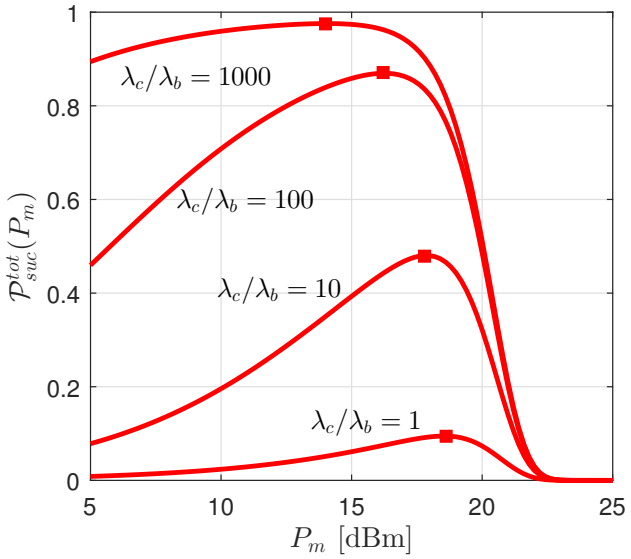


Figure 7.6.: Effect of changing transmit power on energy success, Spectrum success and overall success probabilities for various values of MTD density, Eq. (7.20). $1 \leq okta \leq 6$, $\lambda_b = 10^{-6}$.

harvested energy is high.

3. As the relative density of CM2M aggregators λ_c/λ_b is increased, the spectrum success probability increases and the overall success probability also increases as observed in Figure 7.6. Even though the maximum SAP of MTDs is not related to the density of aggregators (7.9), the non-linear increase in overall success is attributed to the fact that the distance between an arbitrary MTD and its tagged CM2M aggregator decreases with the increase in λ_c. This greatly reduces path loss for the desired signal.

4. We see from Figure 7.7 that for large values of λ_c/λ_b, a smaller optimal power P_m^{opt} can be used by the MTDs to maximize the overall success probability. This is because the effective path loss between the CM2M controller and the generic MTD reduces with increased λ_c and a high overall success probability can be achieved with even smaller transmit power. The useful relationship can be exploited especially for larger okta values, where up to 6 dB less transmit power can achieve the best performance as the number of aggregators per BS goes from 1 to 1000.

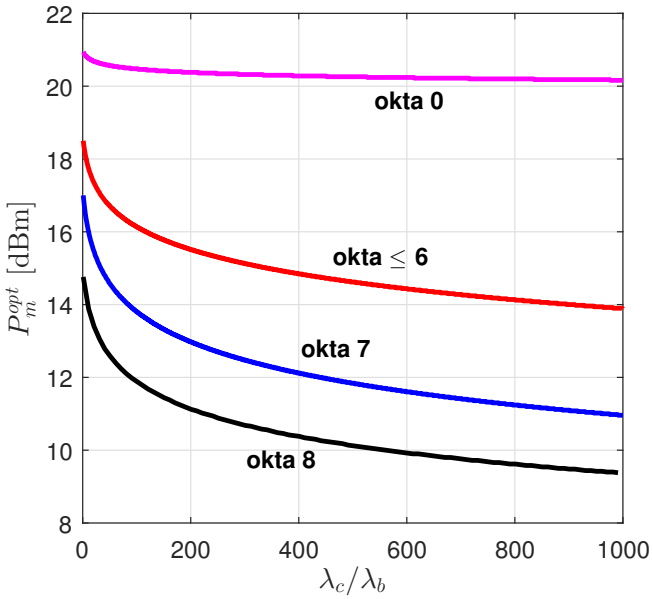


Figure 7.7.: Effect of changing the density of MTDs λ_m on the optimal transmit power P_m^{opt}, $\lambda_b = 10^{-6}$.

7.5. Conclusion

In this chapter, we provided a unified architecture for the cognitive M2M (CM2M) framework. We advocate that the definition of cognition must be extended to incorporate M2M specific design challenges. We solicited a cloud based cognitive underlay spectrum access for the MTD radio platforms. Furthermore, energy harvesting is proposed to attain so called self-sustainable network design. We introduced a novel statistical framework to characterize the energy and spectral success in CM2M networks. The relationship between energy and spectral outages was explored for a reference scenario of solar energy harvesting with stochastic cloud cover. It was shown that both metrics are coupled as 'overall success' as they are governed by same underlying parameter, i.e., transmit power. Finally, there exists a tradeoff between maximizing spectral success and the availability of energy and thus an analytical framework was developed to obtain the optimal transmit power to maximize network level performance.

8 Conclusions and Future Work

In this chapter, we review the main results of this thesis and highlight the important conclusions. We also present the possible extensions and future directions of our work.

8.1. Summary and Conclusions

In this thesis, we have considered various scenarios in which D2D and M2M communication paradigms can be integrated with existing cellular networks. With the help of stochastic geometry, we have developed comprehensive models for the performance analysis of cellular networks in conjunction with M2M and D2D communications. In chapters 3-5, we studied how offloading UEs to D2D mode can result in the enhancement of link spectral efficiency and the overall network throughput. In chapter 3, we considered a simple single cell setup where the BS was equipped with a large antenna array of M antennas. The BS served N UEs inside the cell using zero-forcing beamforming. We investigated the gains that could be achieved in the overall cell capacity if a certain number of UEs were offloaded to operate in D2D mode. We obtained closed-form bounds on the spectral efficiency of an arbitrary UE in cellular and D2D modes. We showed that offloading UEs to D2D mode is particularly beneficial in maximizing the overall capacity when the number of antennas at the BS is small, when the transmit SNRs are high and when the D2D link distance separation is small.

In chapter 4, we studied how D2D communication can be integrated with cellular networks in a multi-cell scenario. We incorporated the realistic path loss model

recommended by the 3GPP for the cellular and D2D links for the analysis of the area spectral efficiency of the network. We compared our results with the reference model where a simple power law path loss model has been used. We demonstrated that with the 3GPP recommended path loss model, the results not only change quantitatively from the reference scenario but also qualitatively. The area spectral efficiency for reference scenario saturates with the increase in the density of the BSs. This, however, is not the case when the realistic path loss model is employed and there exists a unique density of BSs that maximizes the area spectral efficiency of the network.

Chapter 5 focused on the content-centric approach to establish D2D connections. If a BS intends to offload a UE to be served directly by one of its neighboring UEs, it is necessary that neighboring UEs possess the content that has been requested. Our analysis considered that the UE requests follow the well-known Zipf law and the helper UEs willing to serve the requesting UE are equipped with caches. Similar to the EPC-level discovery of D2D helpers proposed by the 3GPP, we assumed that the helper selection, authentication and authorization was carried out by the network. We proposed that the UE will be served by one of its k closest D2D helpers inside the cell subject to content availability. Our main analytical contribution was to develop a tight approximation for the distribution of distance between an arbitrary UE and the ith closest helper within the cell. Using this distribution, we characterized the average rates experienced by the UE in cellular and D2D modes. Our results revealed an interesting trade off between the number of candidate D2D helpers k and the average rate experienced by the UE. We demonstrated that D2D communication is most useful when popular contents are requested, as they are more likely to be present in the caches of helpers closer to the requesting UE.

In chapter 6, we explored a new dimension of D2D communication, which was to use D2D links to connect UEs with MTDs in close proximity for the aggregation of M2M data. The advantage of this approach is to reduce the burden of massive access on the BS while ensuring that the M2M data is relayed efficiently. We assumed that the UEs and MTDs are both deployed independently according to a homogeneous Poisson point processes. We accounted for the proximity factor in establishing a D2D connection by employing the Poisson hard sphere (PHS) model for the UE coverage regions. According to PHS model, a UE coverage region is a disk centered at it, whose radius depends on the distance between its nearest UE. The UE only served the MTDs located within its coverage region. The PHS model is

unlike the commonly used Voronoi tessellation approach, where the coverage regions collectively span the entire space, i.e. all MTDs are covered. The PHS model is also flexible as the coverage regions are of varying sizes. Furthermore, the PHS model ensures that the coverage regions are non-overlapping making the MTD association process simpler. The performance determining metrics for the D2D communication in this case was not to maximize throughput, but to ensure that the rate requirements of both the UEs and MTDs are satisfied while maximizing the number of MTDs served by a given UE. We showed that the QoS of a UE is not affected if it aggregates data from a few MTDs and trunks this data along with its own data.

In chapter 7, we considered a cellular network underlaid with a separate M2M network, where the MTDs transmitted to their closest generic aggregator in cellular downlink. The MTDs were assumed to be energy constrained as they harvested energy from the sun and only the MTDs with sufficient energy were able to transmit to the aggregator. We adopted a stochastic cloud cover model to account for the variation in the harvested energy at a particular time. To avoid disruption to the cellular network, we proposed that the MTDs employ cognition by keeping the aggregate interference to cellular UEs at a certain level. We introduced a cellular outage probability threshold, which dictated the probability of activation of the MTDs. We showed that the harvested energy available at the MTD to transmit and the spectrum access success of the MTDs exhibit an inverse relationship in terms of the transmit power of MTDs. As we increase the power level at which we wish to transmit, a better coverage is guaranteed, but the probability of being able to do this decreases due to the dependence on only the instantaneous harvested energy. We defined a key performance metric called the overall success probability, which captured the trade off in the energy success and the spectrum access success probabilities. We concluded that there exists an optimum transmit power, which maximizes the overall success probability. We further showed that by increasing the density of aggregators inside the cognitive M2M network, we can achieve better performance with a lower MTD transmit power.

8.2. Future Directions

We have identified several future directions for our work listed below.

Steering Nulls at D2D Receivers

In chapter 3, we studied the impact of D2D communication in a massive MIMO systems. Our analysis assumed perfect channel state information (CSI) at the BS and did not take into account the interference generated from the other cells. As a next step, it would be interesting to study the interplay between D2D and massive MIMO with imperfect CSI and pilot contamination. It is also important to study how the additional degrees of freedom at the BS could be used to steer nulls at the D2D receivers to minimize the impact of interference and whether this will affect the performance of cellular UEs.

Modeling and Analysis with Realistic Path loss Model

As stressed in chapter 4, the results and the related inferences may significantly change when LoS and NLoS links are taken into account in the path loss model. Our analysis in all the chapters except chapter 4 consider a simple power law path loss model due to increased tractability. The modification of the analysis based on the 3GPP recommended path loss model still remains an open issue. Additionally, the UE association with the BS is based on highest long term received signal strength and this boils down to association with the BS with the minimum path loss. For power law path loss model, minimum path loss association is synonymous with the association with the nearest BS. But, this is not the case when the realistic path loss model is employed and therefore the association process requires rigorous analysis and a straight forward extension of our analytical results for realistic path loss model is not possible.

Analysis of Spectrum Partition Factor

Our analysis in chapters 4 and 5 considers in-band overlay D2D communication, where a portion of spectrum is reserved for D2D communication. It would be a nice extension to see if this partitioning factor is adapted with respect to the variation in cellular and D2D traffic to efficiently utilize the available spectrum. For example in chapter 5, we have seen that when unpopular content is requested or when the UEs do not demand very high data rates, it is better for the BS to serve the UEs and not offload them to D2D mode. But due to static bandwidth assignment, the available spectrum for D2D communication remains underutilized. By controlling the spectrum partitioning this problem can be mitigated.

Design of Collaborative Caching schemes for D2D helpers

As already shown in chapter 5, the probability for a UE to served by its neighboring D2D helpers critically depends on the helper's cache hit rates and the content popularity. Our work assumes a simple caching mechanism, where each memory unit is independently filled. An interesting future direction would be to see how the contents should be cached in the k candidate D2D helpers' caches to store as much different content as possible to maximize the data rate experienced by the UE. A comparison of various caching strategies would be extremely useful.

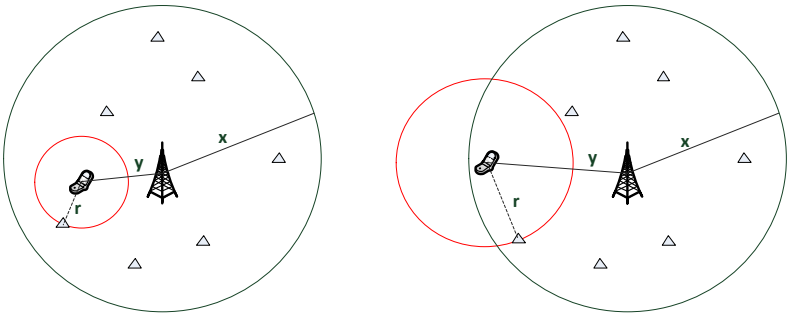
Furthermore, we currently assume that the contents are homogeneously popular both in space and time. It will be interesting to introduce the aspect of spatial and temporal locality in the content popularity as some content may be requested in certain locations and at certain times. Various studies on the behavior of social interaction between the users could play a significant role in analyzing the patterns of content popularities. Devising optimal caching strategies based on the social interaction along with the proximity information of the users will help exploit the full potential of D2D communication.

Design of an Incentive Mechanism for M2M Aggregation

Our work in chapter 6 is based on an inherent assumption that the users are willing to aggregate M2M data from the MTDs. This argument is supported by ensuring that the QoS requirements of a UE are not affected. However, in reality there are many factors that would make the users decide otherwise. These factors include, but are not limited to, battery drainage, security and privacy concerns, etc. For encouraging the users to participate in the aggregation of M2M data, the network operators have to give some incentives to the users, such as free minutes or downloads. Game theoretical techniques can be employed to help achieve rational strategies of the users.

A Proofs of Chapter 5

A.1. Proof of Theorem 5.1



(a) **(b)**

Figure A.1.: Distance to the nearest D2D neighbor within the circular Voronoi cell: (a) when $b(o,r)$ is inside B_{max}, (b) when $b(o,r)$ partly overlaps B_{max}.

The probability that the distance R_i between the requesting UE and the ith nearest D2D helper within the cell is at least r is the probability that there are at least $i-1$ helpers inside the region $\mathcal{A}(r)$. We therefore have

$$F_{R_i|X>Y,N_d\geq i}(r) = 1 - \frac{\Gamma(i,\mathcal{A}(r))}{\Gamma(i)} \tag{A.1}$$

where $\mathcal{A}(r)$ is the area of intersection between B_{max} of radius x and and a disk $b(u,r)$ of radius r centered at the UE at location u. As shown in Fig. A.1, this area of intersection can be categorized into two regimes given as follows.

- Regime 1 - When $b(u,r)$ partly overlaps B_{max}, i.e. $x-y<r<x+y$. The overlapping region $\mathcal{A}(r)$ in this case can be written as [121]

$$\mathcal{A}(r) = \nabla(r, y, x) \quad = \quad r^2 \arccos\left(\frac{\omega_1}{2y\,r}\right) + x^2 \arccos\left(\frac{\omega_2}{2y\,x}\right) - \frac{1}{2}\sqrt{4y^2x^2 - \omega_2^2}, \tag{A.2}$$

where $\omega_1 = r^2 + y^2 - x^2$ and $\omega_2 = x^2 + y - r^2$.

- Regime 2 - When $b(u, r)$ lies inside B_{max} i.e. $0 < r < x - y$. The overlapping region in this case is straightforward and is given as $\mathcal{A}(r) = \pi r^2$.

Differentiating (A.1) with respect to r gives $f_{i,1}(r, y, x)$ in (5.22) and $f_{i,2}(r)$ in (5.23) for regimes 1 and 2 respectively. Then, given that the UE and at least i D2D helpers lie inside B_{max}, the unconditional distance distribution $f_{R_i}(r)$ is obtained by averaging (5.22) and (5.23) over X and Y as follows

$$f_{R_i}(r) = \frac{\int_0^\infty f_Y(y)\left[\int_{a_1}^{a_2} f_{i,1}(r, y, x)\, f_X(x)\, dx + \int_{a_2}^{\infty} f_{i,2}(r)\, f_X(x)\, dx\right] dy}{\mathbb{P}\left[X > Y\right].\mathbb{P}\left[N_d \geq i\right]}. \tag{A.3}$$

The limits of integration a_1 and a_2 are obtained by rearranging the limits of r. For Regime 2, $x > a_2 = r + y$ and for Regime 1, $a_1 = r - y < x < a_2$. But, since $x > y$, we have $a_1 = \max(y, r - y)$. Solving the integrals in (A.3) and substituting the values for $\mathbb{P}\left[X > Y\right] = p_{in}$ from (5.18) and $\mathbb{P}\left[N_d \geq i\right] = p_{N_d}^{(i)}$ from (5.19), we obtain (5.21).

A.2. Proof of Proposition 5.3

Because of the exponentially distributed channel power, we can write (5.34) as

$$\mathcal{R}_{d,i,\Pi}^{(k)} = \mathbb{E}_{R_i}\left[\exp\left(-s_d\sigma^2/P_d\right)\mathcal{L}_{I_d}\left(s_d\right)\right], \tag{A.4}$$

where $s_d = \tau_d r^\alpha$ and $\mathcal{L}_{I_d}\left(s_d\right) = \mathbb{E}_{I_d}\left[\exp\left(-s_d I_d\right)\right]$ is the Laplace transform of D2D interference. Since, at most one D2D connection is active in a cell in a given time instant, we employ a key assumption that Φ_d^{int} is also HPPP[1] with intensity $\lambda_d^{int} = p_d^{int} \times \lambda_m$. Here, p_d^{int} from (5.37) is the probability that there is at least one UE

[1]The equi-dense HPPP assumptions ignores the correlations due to the position of helpers inside a cell, but is more tractable. [22]

being served in D2D mode inside an arbitrary cell. We can then write $\mathcal{L}_{I_d}(s_d)$ as

$$\mathcal{L}_{I_d}(s_d) = \mathbb{E}_Q\left[\exp\left(-2\pi\lambda_d^{int}\int_q^\infty \frac{\nu}{1+s_d^{-1}\nu^\alpha}\,d\nu\right)\right], \tag{A.5}$$

The lower limit of the integral in (A.5) represents the minimum distance separation between the typical UE and the interfering D2D helpers in other cells. Notice that the lower limit $q = \|z_i\|$ in this case is different from the distance r, which is the distance to the ith nearest neighbor within the cell. This is because the interfering D2D helper may be closer to the typical UE than the ith nearest D2D helper within the cell. Therefore, a distance separation of r doesn't ensure a guard zone. The value of q is governed by the nearest neighbor distribution of HPPPs because of the equi-dense HPPP approximation. Therefore, $f_Q(q) = 2\pi\lambda_d^{int} q\exp(-\lambda_d^{int}\pi q^2)$. Owing to the convexity of functions of the exponential function, we apply Jensen's inequality to achieve a tight bound for (A.5)

$$\mathcal{L}_{I_d}(s_d) \approx \exp\left(-2\pi\lambda_d^{int}\mathbb{E}_Q\left[\int_q^\infty \frac{\nu}{1+s_d^{-1}\nu^\alpha}\,d\nu\right]\right). \tag{A.6}$$

Substituting (A.6) into (A.4) gives (5.39).

The overall D2D coverage probability for NS and US schemes and a particular content request in (5.38) is obtained by taking expectation over i and by conditioning over the probability of D2D mode from Prop. 5.1.

B Proofs of chapter 6

B.1. Proof of Theorem 6.1

The coverage probability for MX case can be characterized as

$$\begin{aligned}\mathcal{S}_a^{MX}(\theta) &= \mathbb{E}_{R_{arb}}\left[\mathbb{P}\left[SIR_a^{MX} \geq \theta | R_{arb} = r\right]\right] = \mathbb{E}_{R_{arb}}\left[\frac{h\, r^{-\alpha}}{I_m} \geq \theta\right] \\ &\overset{(a)}{=} \mathbb{E}_{R_{arb}}\left[\mathcal{L}_{I_m}(s_a)\right]. \end{aligned} \tag{B.1}$$

Here, (a) follows from the exponential distribution of the channel gain h, $s_a = \theta r^{\alpha}$ and $\mathcal{L}_{I_m}(.)$ is the Laplace transform of I_m which is the interference experienced by the UE from the MTDs outside its S-cell. For analytical tractability, we assume that the set of active MTDs Φ_m^{act} in MX case constitute a HPPP with density λ_m^{act}. At a given time, as only one MTD inside a S-cell (if there is any) will be transmitting, the effective density of Φ_m^{act} will be $\lambda_m^{act} = \lambda_u \times \mathbb{P}\left[N_m \geq 1\right] = \lambda_u(1+\mu)^{-1}$. Hence, we have

$$\begin{aligned}\mathcal{L}_{I_m}(s_a) &= \mathbb{E}\left[\prod_{\omega_j \in \Phi_m^{act}\setminus o} \exp\left(-s_a h_j ||\omega_j||^{-\alpha}\right)\right] \\ &\overset{(b)}{=} \mathbb{E}_Q\left[\exp\left(-2\pi\lambda_m^{act}\int_q^{\infty} \mathbb{E}_h\left[1-\exp\left(-s_a h\, \nu^{-\alpha}\right)\right]\nu d\nu\right) \mid Q\right], \end{aligned} \tag{B.2}$$

where (b) follows from the probability generating functional (PGFL) of PPPs. Since the channel power gain is independent of Φ_m^{act}, the expectation with respect to h is moved inside the integral. The lower limit of integral q represents the minimum distance separation between the UE and the interfering MTD. The value of q is governed by $f_Q(q) = 2\pi\lambda_m^{act} q \exp\left(-\lambda_m^{act}\pi q^2\right)$ because of the HPPP assumption, where Q is the distance from the typical UE to the nearest UE having at least one MTD in its S-cell Taking expectation with respect to h, we get $\mathcal{L}_{I_m}(s_a) =$

$\mathbb{E}_Q\left[\exp\left(-2\pi\lambda_m^{act}\int_q^\infty \frac{\nu}{1+s_a^{-1}\nu^\alpha}d\nu\right) \mid Q\right]$. For further simplification, we exploit the convexity of the exponential function and apply Jensen's inequality to shift the expectation operator inside the exponential function to obtain (6.14).

B.2. Proof of Theorem 6.2

The SIR coverage probability to decode k MTDs can be written as

$$\mathcal{S}_a^{SIC}(\theta) = \prod_{i=1}^{k} \mathcal{S}_{a,i}^{SIC}(\theta), \tag{B.3}$$

where $\mathcal{S}_{a,i}^{SIC}(\theta)$ is the probability to decode the ith nearest MTD ($i = \{1, .., k\}$) given that all $i-1$ MTDs have already been decoded. It is given as

$$\begin{aligned} \mathcal{S}_{a,i}^{SIC}(\theta) &= \mathbb{E}_{R_i}\left[\mathbb{P}\left\{SIR_{a,i}^{SIC} \geq \theta | R_i\right\}\right], \\ &= \mathbb{E}_{R_i}\left[\mathcal{L}_{I_{in}}(s_a)\,\mathcal{L}_{I_{out}}(s_a)\right]. \end{aligned} \tag{B.4}$$

Here, $\mathcal{L}_{I_{in}}(.)$ and $\mathcal{L}_{I_{out}}(.)$ are respectively the Laplace transforms of the aggregate interference experienced by the typical UE from MTDs inside its S-cell and other S-cells. The derivation of $\mathcal{L}_{I_{out}}(.)$ is rather straight forward as we combine the effect of MTDs in each interfering S-cell owing to the fact that the MTDs are uniformly distributed inside the S-cell. We consider that the collection of active MTDs constitutes a HPPP with density $\lambda_m^{out} = \lambda_u$ with effective power $N_m P_m$, where N_m is the random number of MTDs inside each S-cell. Hence, we have $\mathcal{L}_{I_{out}}(s_a) = \mathbb{E}_{\Phi_m^{out}}\left[\prod_{v_j \in \Phi_m^{out}} \mathbb{E}_{h,N_m}\left[\exp\left(-s_a N_m h ||v_j||^{-\alpha}\right)\right]\right]$. Making use of Jensen's inequality and taking the expectation with respect to N_m inside the exponent and applying the PGFL of HPPPs, we get $\mathcal{L}_{I_{out}}(s_a) \approx \mathbb{E}_{R_{uncon}}\left[\exp\left(-2\pi\lambda_u\int_w^\infty \frac{\nu}{1+\left(N_m^{avg} s_a\right)^{-1}\nu^\alpha}d\nu\right)\right]$. The lower limit w ensures that the interference power is coming from outside the S-cell under consideration. It indicates the distance between the typical UE and the nearest interfering MTD and follows the same distribution as R_{uncon} in (6.10) as we consider the location of UEs as the central source of interference.

Using similar techniques, the Laplace transform of interference $\mathcal{L}_{I_{in}}(s_a)$ experi-

enced from MTDs inside the typical S-cell can be written as

$$\mathcal{L}_{I_{in}}(s_a) = \mathbb{E}_{X|X\geq r}\left[\exp\left(-2\pi\lambda_m \int_r^x \frac{\nu}{1+s_a^{-1}\nu^{\alpha}} d\nu\right)\right], \tag{B.5}$$

where the lower limit of the integral ensures that the MTDs $\{1, ..i-1\}$ do not contribute to the interference as their data has been previously decoded and the closest interferer is at least at a distance r from the typical UE. On the other hand, the upper limit x denotes the boundary of the typical S-cell. Notice here that for the validity of the analysis, the expectation is with respect to the radius of S-cell with the condition that the radius of S-cell is grater than r. This conditional distribution can be expressed as

$$\mathbb{P}\left[X = x | X \geq r\right] = \frac{f_X(x)}{\mathbb{P}\left[X \geq r\right]} = \frac{f_X(x)}{\exp\left(-\lambda_u \eta^{-2}\pi r^2\right)}, \qquad x > r. \tag{B.6}$$

Substituting $\mathcal{L}_{I_{in}}(s_a)$ and $\mathcal{L}_{I_{out}}(s_a)$ in (B.4) and (B.3) gives (6.20).

B.3. Proof of Theorem 6.3

The coverage probability of the UE can be expressed as

$$\mathcal{S}_t(\theta)|Y = \begin{cases} \mathbb{P}\left[\frac{\rho_0 h}{I_u} \geq \theta\right] & 0 \leq y \leq R_{max}, \\ \mathbb{P}\left[\frac{P_u^{max} h\, y^{-\alpha}}{I_u} \geq \theta\right] & y > R_{max}. \end{cases} \tag{B.7}$$

Alternatively, we can write $\mathcal{S}_t(\theta)|Y = \mathbb{P}\left[h \geq s_t I_u\right]$, where

$$s_t = \begin{cases} s_1 = \theta/\rho_0 & 0 \leq y \leq R_{max}, \\ s_2 = \theta y^{\alpha}/P_u^{max} & y > R_{max}. \end{cases}$$

Because of the exponentially distributed channel power h, we have $\mathcal{S}_t(\theta)|Y = \mathcal{L}_{I_u}(s_t)$, where $\mathcal{L}_{I_u}(s_t)$ is the Laplace transform of the aggregate interference experienced by the typical BS in uplink from active UEs in other macrocells. It can be written as

$$\mathcal{L}_{I_u}(s_t) = \mathbb{E}_{\Phi_u^{act}}\left[\prod_{z_j \in \Phi_u^{act}} \mathbb{E}_{h,P_u}\left[\exp\left(-s_t P_u h\, ||z_j||^{-\alpha}\right)\right]\right].$$

The interfering UEs $z_j \in \Phi_u^{act}$ do not comprise a HPPP, but a Poisson-Voronoi perturbed lattice due to the correlations between the sizes of adjacent macrocells. Hence, the characterization of the Laplace transform of interference in intractable. We employ the approximation proposed in [92] to approximate Φ_u^{act} as an inhomogeneous PPP with intensity measure function $\Lambda_u(d\nu) = 2\pi\lambda_u \nu d\nu$, where $\nu = ||z_j||$ is the distance of a UE to the BS at the origin. We therefore have

$$
\begin{aligned}
\mathcal{L}_{I_u}(s_t) &= \exp\left(-\int_{\nu>0} \mathbb{E}_{P_u}\left[\frac{1}{1+\left(s_t P_u\right)^{-1}\nu^{\alpha}}\right]\Lambda_u(d\nu)\right), \\
&\overset{(a)}{=} \exp\left(-\left(1-p_{max}\right)\mathbb{E}_U\left[\int_u^{\infty}\frac{\Lambda_u(d\nu)}{1+\left(s_t\rho_0 u^{\alpha}\right)^{-1}\nu^{\alpha}} \mid U < R_{max}\right]\right) \\
&\quad \cdot \exp\left(-p_{max}\mathbb{E}_U\left[\int_u^{\infty}\frac{\Lambda_u(d\nu)}{1+\left(s_t P_u^{max}\right)^{-1}\nu^{\alpha}} \mid U > R_{max}\right]\right) \\
&\overset{(b)}{=} \exp\left(-\pi\lambda_u\left(1-p_{max}\right)\mathbb{E}_U\left[\int_1^{\infty}\frac{u^2}{1+\left(s_t\rho_0\right)^{-1}t^{\alpha/2}}dt \mid U < R_{max}\right]\right) \\
&\quad \cdot \exp\left(-\pi\lambda_u\left(1-p_{max}\right)\mathbb{E}_U\left[\int_1^{\infty}\frac{u^2}{1+\left(s_t P_u^{max}\right)^{-1}u^{\alpha}t^{\alpha/2}}dt \mid U < R_{max}\right]\right) \\
&\quad \exp\Bigg(-\left(1-p_{max}\right)\mathbb{E}_U\left[u^2 \mid U < R_{max}\right]\mathcal{C}(\alpha, s_t\rho_0, 1)\Bigg). \\
&\quad \cdot \exp\Bigg(-p_{max}\mathbb{E}_U\left[u^2\mathcal{C}(\alpha, s_t P_u^{max}u^{-\alpha}, 1) \mid U > R_{max}\right]\Bigg), \qquad \text{(B.8)}
\end{aligned}
$$

where the distance U between the interfering UE z_j and its associated BS is also a random variable and follows the same distribution as Y. We substitute the value of P_u in step (a) and exchange the order of the integral and the expectation. Notice that the lower limit u of the integrals in (a) is due to the nearest BS association. This implies that z_j is at least at a separation of u from the typical BS. Step (b) follows by employing the substitution $t = (v/u)^2$. Solving the integral for t, we obtain (6.28).

Bibliography

[1] "Cisco visual networking index global mobile data traffic forecast update, 2016-2021 white paper." http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html. [Online; accessed 20/06/2017].

[2] "Cisco services dynamics." http://www.ciscoservicesdynamics.com/edition/5/With-no-signs-of-slowing-data-just-keeps-growing. [Online; accessed 04/07/2017].

[3] M. Chen12, J. Wan, and F. Li, "Machine-to-machine communications: architectures, standards and applications," *KSII transaction on internet and information systems*, vol. 6, no. 2, pp. 480–497, 2012.

[4] "5g: A technology vision." www.huawei.com/5gwhitepaper. [Online; accessed 20/06/2017].

[5] 3GPP-TR22.803, "Feasibility study for proximity services (ProSe)." http://www.3gpp.org/ftp/specs/archive/23_series/23.703/. [Online; accessed 20/06/2017].

[6] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, 2014.

[7] G. Chandrasekaran, N. Wang, and R. Tafazolli, "Caching on the move: Towards D2D-based information centric networking for mobile content distribution," in *40th IEEE Conference on Local Computer Networks (LCN)*, pp. 312–320, 2015.

[8] A. Asadi, Q. Wang, and V. Mancuso, "A Survey on Device-to-Device Communication In Cellular Networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 1801–1819, 2014.

[9] R. Hekmat, *Ad-hoc networks: fundamental properties and network topologies*. Springer Science & Business Media, 2006.

[10] H.-Y. Hsieh and R. Sivakumar, "On using peer-to-peer communication in cellular wireless data networks," *IEEE Transactions on Mobile Computing*, vol. 3, no. 1, pp. 57–72, 2004.

[11] Y.-C. Liang, K.-C. Chen, G. Y. Li, and P. Mahonen, "Cognitive radio networking and communications: An overview," *IEEE transactions on vehicular technology*, vol. 60, no. 7, pp. 3386–3407, 2011.

[12] "Qualcomm research on LTE Direct, a device-to-device discovery technology." `https://www.qualcomm.com/invention/research/projects/lte-direct`. [Online; accessed 12/07/2017].

[13] 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, no. 4, pp. 40–50, 2014.

[14] W. Zhao and S. Wang, "Resource sharing scheme for device-to-device communication underlaying cellular networks," *IEEE Transactions on Communications*, vol. 63, no. 12, pp. 4838–4848, 2015.

[15] H. Tang and Z. Ding, "Mixed mode transmission and resource allocation for D2D communication," *IEEE Transactions on Wireless Communications*, vol. 15, no. 1, pp. 162–175, 2016.

[16] K. Doppler, M. Rinne, C. Wijting, C. B. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to lte-advanced networks," vol. 47, no. 12, pp. 42–49, 2009.

[17] C.-H. Yu, K. Doppler, C. B. 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.

[18] 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, no. 8, pp. 3541–3551, 2013.

[19] G. Fodor, E. Dahlman, G. Mildh, S. Parkvall, N. Reider, G. Miklós, and Z. Turányi, "Design Aspects of Network Assisted Device-to-Device Communications," *IEEE Communications Magazine*, vol. 50, no. 3, pp. 170–177, 2012.

[20] C. Xu, L. Song, Z. Han, Q. Zhao, X. Wang, and B. Jiao, "Interference-aware resource allocation for device-to-device communications as an underlay using sequential second price auction," in *IEEE International Conference on Communications (ICC)*, pp. 445–449, 2012.

[21] W. Xu, L. Liang, H. Zhang, S. Jin, J. C. Li, and M. Lei, "Performance enhanced transmission in device-to-device communications: Beamforming or interference cancellation?," in *Global Communications Conference (GLOBECOM), 2012 IEEE*, pp. 4296–4301, IEEE, 2012.

[22] H. ElSawy, E. Hossain, and M.-S. Alouini, "Analytical modeling of mode selection and power control for underlay D2D communication in cellular networks," *IEEE Transactions on Communications*, vol. 62, no. 11, pp. 4147–4161, 2014.

[23] 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, no. 7, pp. 3665–3676, 2014.

[24] S. E. Hajri and M. Assaad, "An exclusion zone for massive MIMO with underlay D2D communication," in *2015 International Symposium on Wireless Communication Systems (ISWCS)*, pp. 471–475, IEEE, 2015.

[25] 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, no. 3, pp. 118–127, 2014.

[26] J. F. Schmidt, M. K. Atiq, U. Schilcher, and C. Bettstetter, "Underlay device-to-device communications in LTE-A: uplink or downlink?," in *26th Annual IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1542–1546, 2015.

[27] X. Lin, J. G. Andrews, and A. Ghosh, "Spectrum Sharing for Device-to-Device Communication in Cellular Networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 12, pp. 6727–6740, 2014.

[28] A. Asadi and V. Mancuso, "WiFi Direct and LTE D2D in action," in *IEEE IFIP Wireless Days (WD)*, pp. 1–8, 2013.

[29] B. Raghothaman, E. Deng, R. Pragada, G. Sternberg, T. Deng, and K. Vanganuru, "Architecture and protocols for LTE-based device to device communication," in *IEEE International Conference on Computing, Networking and Communications (ICNC)*, pp. 895–899, 2013.

[30] M. Hicham, N. Abghour, and M. Ouzzif, "Device-to-device (D2D) communication under LTE-Advanced networks," *International Journal of Wireless & Mobile Networks (IJWMN)*, vol. 8, no. 1, 2016.

[31] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3122–3134, 2011.

[32] M. Haenggi, *Stochastic Geometry for Wireless Networks*. Cambridge University Press, 2012.

[33] Z. Han, *Game theory in wireless and communication networks: theory, models, and applications*. Cambridge University Press, 2012.

[34] N. Deo, *Graph theory with applications to engineering and computer science*. Courier Dover Publications, 2017.

[35] D. Stoyan, W. S. Kendall, J. Mecke, and L. Ruschendorf, *Stochastic geometry and its applications*, vol. 2. Wiley New York, 1987.

[36] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic geometry and its applications*. John Wiley & Sons, 2013.

[37] D. Gesbert, M. Kountouris, R. W. Heath Jr, C.-B. Chae, and T. Salzer, "Shifting the MIMO paradigm," vol. 24, no. 5, pp. 36–46, 2007.

[38] T. L. Marzetta, "Massive MIMO: an introduction," *Bell Labs Technical Journal*, vol. 20, pp. 11–22, 2015.

[39] Y. Ni, S. Jin, W. Xu, Y. Wang, M. Matthaiou, and H. Zhu, "Beamforming and interference cancellation for D2D communication underlaying cellular networks," vol. 64, no. 2, pp. 832–846, 2016.

[40] X. Lin, R. W. Heath, and J. G. Andrews, "The interplay between massive MIMO and underlaid D2D networking," vol. 14, no. 6, pp. 3337–3351, 2015.

[41] S. Shalmashi, E. Björnson, M. Kountouris, K. W. Sung, and M. Debbah, "Energy efficiency and sum rate tradeoffs for massive mimo systems with underlaid device-to-device communications," *arXiv preprint arXiv:1506.00598*, 2015.

[42] S. Shalmashi, E. Björnson, S. B. Slimane, and M. Debbah, "Closed-form optimality characterization of network-assisted Device-to-Device communications," in *2014 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 508–513, IEEE, 2014.

[43] J. Hoydis, K. Hosseini, S. t. Brink, and M. Debbah, "Making smart use of excess antennas: Massive MIMO, small cells, and TDD," *Bell Labs Technical Journal*, vol. 18, no. 2, pp. 5–21, 2013.

[44] J. C. Li, M. Lei, and F. Gao, "Device-to-device (D2D) communication in mu-mimo cellular networks," in *Global Communications Conference (GLOBECOM), 2012 IEEE*, pp. 3583–3587, IEEE, 2012.

[45] A. Afzal, S. A. R. Zaidi, D. McLernon, and M. Ghogho, "On the analysis of cellular networks with caching and coordinated Device-to-Device communication," in *IEEE International Conference on Communications (ICC)*, pp. 1–7, 2016.

[46] D. Moltchanov, "Distance distributions in random networks," *Ad Hoc Networks*, vol. 10, no. 6, pp. 1146–1166, 2012.

[47] T. Bai and R. W. Heath, "Coverage and rate analysis for millimeter-wave cellular networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 2, pp. 1100–1114, 2015.

[48] X. Zhang and J. G. Andrews, "Downlink cellular network analysis with multi-slope path loss models," *CoRR*, vol. abs/1408.0549, 2014.

[49] M. Ding, D. Lopez-Perez, G. Mao, P. Wang, and Z. Lin, "Will the area spectral efficiency monotonically grow as small cells go dense?," *arXiv preprint arXiv:1505.01920*, 2015.

[50] M. Di Renzo, "Stochastic geometry modeling and analysis of multi-tier millimeter wave cellular networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 9, pp. 5038–5057, 2015.

[51] "3GPP TR 36.828: Further enhancements to LTE time division duplex (TDD) for (DL-UL) interference management and traffic adaptation," 2012.

[52] "3GPP TR 36.843: Study on LTE device to device proximity services; radio aspects," 2014.

[53] 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, no. 8, pp. 82–89, 2014.

[54] S. Woo, E. Jeong, S. Park, J. Lee, S. Ihm, and K. Park, "Comparison of Caching Strategies in Modern Cellular Backhaul Networks," *ACM Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, pp. 319–332, 2013.

[55] X. Lin, J. Andrews, A. Ghosh, and R. Ratasuk, "An Overview of 3GPP Device-to-Device Proximity Services," *IEEE Communications Magazine*, vol. 52, no. 4, pp. 40–48, 2014.

[56] F. Malandrino, C. Casetti, and C.-F. Chiasserini, "Toward D2D-Enhanced Heterogeneous Networks," *IEEE Communications Magazine*, vol. 52, no. 11, pp. 94–100, 2014.

[57] M. Ji, G. Caire, and A. Molisch, "Fundamental limits of caching in wireless D2D networks," *IEEE Transactions on Information Theory*, vol. 62, no. 2, pp. 849 – 869, 2014.

[58] 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, no. 1, pp. 176–189, 2016.

[59] B. Perabathini, E. Bastug, M. Kountouris, M. Debbah, and A. Conte, "Caching at the Edge: a Green Perspective for 5G Networks," *CoRR*, vol. abs/1503.05365, 2015.

[60] S. A. R. Zaidi, M. Ghogho, and D. C. McLernon, "Information Centric Modeling for Two-tier Cache Enabled Cellular Networks," *IEEE International Conference on Communications (ICC)*, 2015.

[61] S. Tamoor-ul Hassan, M. Bennis, P. H. Nardelli, and M. Latva-Aho, "Modeling and Analysis of Content Caching in Wireless Small Cell Networks," *arXiv preprint arXiv:1507.00182*, 2015.

[62] E. Bastug, M. Bennis, and M. Debbah, "Cache-Enabled Small Cell Networks: Modeling and Tradeoffs," *11th IEEE International Symposium on Wireless Communications Systems (ISWCS)*, pp. 649–653, 2014.

[63] M. Afshang, H. S. Dhillon, and P. H. J. Chong, "Fundamentals of cluster-centric content placement in cache-enabled Device-to-Device networks," *arXiv preprint arXiv:1509.04747*, 2015.

[64] Z. Chen, J. Lee, T. Q. Quek, and M. Kountouris, "Cooperative caching and transmission design in cluster-centric small cell networks," *arXiv preprint arXiv:1601.00321*, 2016.

[65] M. Afshang, H. S. Dhillon, and P. H. J. Chong, "Modeling and Performance Analysis of Clustered Device-to-Device Networks," *arXiv preprint arXiv:1508.02668*, 2015.

[66] A. Altieri, P. Piantanida, L. R. Vega, and C. G. Galarza, "On Fundamental Trade-offs of Device-to-Device Communications in Large Wireless Networks," *CoRR*, vol. abs/1405.2295, 2014.

[67] U. Niesen and M. A. Maddah-Ali, "Coded Caching for Delay-Sensitive Content," *arXiv preprint arXiv:1407.4489*, 2014.

[68] A. Dabirmoghaddam, M. M. Barijough, and J. Garcia-Luna-Aceves, "Understanding Optimal Caching and Opportunistic Caching at the Edge of Information-Centric Networks," *ACM Proceedings of the 1st international conference on Information-centric networking*, pp. 47–56, 2014.

[69] E. Baştuğ, M. Bennis, and M. Debbah, "A Transfer Learning Approach for Cache-Enabled Wireless Networks," *arXiv preprint arXiv:1503.05448*, 2015.

[70] C. Fricker, P. Robert, J. Roberts, and N. Sbihi, "Impact of Traffic Mix on Caching Performance in a Content-Centric Network," *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 310–315, 2012.

[71] B. Blaszczyszyn and A. Giovanidis, "Optimal geographic caching in cellular networks," *IEEE International Conference on Communications (ICC)*, pp. 3358–3363, 2015.

[72] K. Avrachenkov, X. Bai, and J. Goseling, "Optimization of caching devices with geometric constraints," *arXiv preprint arXiv:1602.03635*, 2016.

[73] D. Liu and C. Yang, "Energy Efficiency of Downlink Networks with Caching at Base Stations," *CoRR*, vol. abs/1505.06615, 2015.

[74] M. Wildemeersch, T. Q. Quek, M. Kountouris, A. Rabbachin, and C. H. Slump, "Successive interference cancellation in heterogeneous networks," *IEEE transactions on communications*, vol. 62, no. 12, pp. 4440–4453, 2014.

[75] S. M. Yu and S.-L. Kim, "Downlink capacity and base station density in cellular networks," *11th IEEE International Symposium on Modeling & Optimization in Mobile, Ad Hoc & Wireless Networks (WiOpt)*, pp. 119–124, 2013.

[76] S. Foss and S. Zuyev, "On a Voronoi Aggregative Process Related to a Bivariate Poisson Process," *Advances in Applied Probability*, pp. 965–981, 1996.

[77] S. Singh, H. S. Dhillon, and J. G. Andrews, "Offloading in heterogeneous networks: Modeling, analysis, and design insights," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 2484–2497, 2013.

[78] M. Andersson, "Short range low power wireless devices and Internet of Things (IoT): White paper." `https://www.u-blox.com/sites/default/files/ShortRange-InternetOfThings_WhitePaper_(UBX-14054570).pdf`. [Online; accessed 09/08/2016].

[79] E. Dahlman, G. Mildh, S. Parkvall, J. Peisa, J. Sachs, Y. Selén, and J. Sköld, "5G wireless access: requirements and realization," *IEEE Communications Magazine*, vol. 52, no. 12, pp. 42–47, 2014.

[80] "Gsma public policy position: Spectrum for the internet of things." https://www.gsma.com/spectrum/wp-content/uploads/2017/05/Spectrum-IOT-Position-Paper.pdf. [Online; accessed 15/11/2017].

[81] S.-Y. Lien, K.-C. Chen, and Y. Lin, "Toward ubiquitous massive accesses in 3GPP machine-to-machine communications," *IEEE Communications Magazine*, vol. 49, no. 4, pp. 66–74, 2011.

[82] A. Biral, M. Centenaro, A. Zanella, L. Vangelista, and M. Zorzi, "The challenges of M2M massive access in wireless cellular networks," *Digital Communications and Networks*, vol. 1, no. 1, pp. 1–19, 2015.

[83] H. S. Dhillon, H. C. Huang, H. Viswanathan, and R. A. Valenzuela, "Power-efficient system design for cellular-based machine-to-machine communications," *IEEE Transactions on Wireless Communications*, vol. 12, no. 11, pp. 5740–5753, 2013.

[84] H. S. Dhillon, H. Huang, H. Viswanathan, and R. A. Valenzuela, "Fundamentals of throughput maximization with random arrivals for M2M communications," *IEEE Transactions on Communications*, vol. 62, no. 11, pp. 4094–4109, 2014.

[85] M. Hasan, E. Hossain, and D. Niyato, "Random access for machine-to-machine communication in lte-advanced networks: issues and approaches," *IEEE Communications Magazine*, vol. 51, no. 6, pp. 86–93, 2013.

[86] A. Aijaz, M. Tshangini, M. R. Nakhai, X. Chu, and A.-H. Aghvami, "Energy-efficient uplink resource allocation in LTE networks with M2M/H2H co-existence under statistical qos guarantees," *IEEE Transactions on Communications*, vol. 62, no. 7, pp. 2353–2365, 2014.

[87] A. Afzal, S. A. R. Zaidi, M. Z. Shakir, M. A. Imran, M. Ghogho, A. V. Vasilakos, D. C. McLernon, and K. Qaraqe, "The cognitive internet of things: a unified perspective," *Mobile Networks and Applications*, vol. 20, no. 1, pp. 72–85, 2015.

[88] F. Ghavimi and H.-H. Chen, "M2M communications in 3GPP LTE/LTE-a networks: architectures, service requirements, challenges, and applications," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 525–549, 2015.

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

[90] Z. Dawy, W. Saad, A. Ghosh, J. G. Andrews, and E. Yaacoub, "Towards massive machine type cellular communications," *arXiv preprint arXiv:1512.03452*, 2015.

[91] K. Zhou and N. Nikaein, "Packet aggregation for machine type communications in LTE with random access channel," *IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 262–267, 2013.

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

[93] G. Rigazzi, N. K. Pratas, P. Popovski, and R. Fantacci, "Aggregation and trunking of M2M traffic via D2D connections," *IEEE International Conference on Communications (ICC)*, pp. 2973–2978, 2015.

[94] O. Häggström and R. Meester, "Nearest neighbor and hard sphere models in continuum percolation," *Random Structures & Algorithms*, vol. 9, no. 3, pp. 295–315, 1996.

[95] C. P. Royall, W. C. Poon, and E. R. Weeks, "In search of colloidal hard spheres," *Soft Matter*, vol. 9, no. 1, pp. 17–27, 2013.

[96] J. Stienen, *Die Vergroeberung von Karbiden in reinen Eisen-Kohlenstoff-Staehlen*. na, 1982.

[97] A. Giovanidis, "How to group wireless nodes together?," *arXiv preprint arXiv:1602.03906*, 2016.

[98] R. K. Ganti and M. Haenggi, "Interference and outage in clustered wireless ad hoc networks," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 4067–4086, 2009.

[99] R. H. Aquino, S. A. R. Zaidi, D. McLernon, and M. Ghogho, "Modelling and performance evaluation of non-uniform two-tier cellular networks through Stienen model," *IEEE International Conference on Communications (ICC)*, 2016.

[100] M. O. Hasna, M.-S. Alouini, A. Bastami, and E. S. Ebbini, "Performance analysis of cellular mobile systems with successive co-channel interference cancellation," *IEEE transactions on wireless communications*, vol. 2, no. 1, pp. 29–40, 2003.

[101] X. Zhang and M. Haenggi, "The performance of successive interference cancellation in random wireless networks," *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 6368–6388, 2014.

[102] M. Wildemeersch, T. Q. Quek, M. Kountouris, and C. H. Slump, "Successive interference cancellation in uplink cellular networks," *14th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 310–314, 2013.

[103] D. Moltchanov, "Survey paper: Distance distributions in random networks," *Ad Hoc Networks*, vol. 10, no. 6, pp. 1146–1166, 2012.

[104] E. I. Unit, "The internet of things business index," *ARM*, pp. 1–22, 2014.

[105] M. A. McHenry, P. A. Tenhula, D. McCloskey, D. A. Roberson, and C. S. Hood, "Chicago spectrum occupancy measurements & analysis and a long-term studies proposal," in *Proceedings of the first international workshop on Technology and policy for accessing spectrum*, p. 1, ACM, 2006.

[106] M. A. McHenry, "Nsf spectrum occupancy measurements project summary," *Shared spectrum company report*, 2005.

[107] E. FCC, "Docket 98-153," *First Report and Order: Revision of Part*, vol. 15, 2002.

[108] E. I. Unit, "Ericsson energy and carbon report: on the impact of networked society," *Ericsson*, pp. 1–11, 2013.

[109] J. G. Andrews, S. Singh, Q. Ye, X. Lin, and H. S. Dhillon, "An overview of load balancing in hetnets: Old myths and open problems," *Wireless Communications, IEEE*, vol. 21, no. 2, pp. 18–25, 2014.

[110] S. A. R. Zaidi, M. Ghogho, and D. C. McLernon, "Breaking the area spectral efficiency wall in cognitive underlay networks," *IEEE Journal on Selected Areas in Communications*, 2014.

[111] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu, *Spatial tessellations: concepts and applications of Voronoi diagrams*, vol. 501. John Wiley & Sons, 2009.

[112] F. Yildiz, "Potential ambient energy-harvesting sources and techniques.," *Journal of Technology Studies*, vol. 35, no. 1, pp. 40–48, 2009.

[113] J. Bright, C. Smith, P. Taylor, and R. Crook, "Stochastic generation of minutely irradiance time series derived from mean hourly weather observation data," *Solar Energy*, 2014.

[114] K. Stamnes, S.-C. Tsay, W. Wiscombe, and I. Laszlo, "Disort, a general-purpose fortran program for discrete-ordinate-method radiative transfer in scattering and emitting layered media: documentation of methodology," 2000.

[115] B. Mayer, A. Kylling, C. Emde, U. Hamann, and R. Buras, "libradtran users guide," 2011.

[116] S. A. R. Zaidi, M. Z. Shakir, M. A. Imran, M. Ghogho, A. Vasilakos, K. Qaraqe, and D. McLernon, "Cognitive Internet of Things: A unified perspective," *International Conference on Software-Defined and Virtualized Future Wireless Networks*, 2014.

[117] E. Rodrigues, R. Melício, V. Mendes, and J. Catalão, "Simulation of a solar cell considering single-diode equivalent circuit model," in *International conference on renewable energies and power quality, Spain*, 2011.

[118] G. Gasper and M. Rahman, *Basic hypergeometric series*, vol. 96. Cambridge university press, 2004.

[119] A. Jeffrey and D. Zwillinger, *Table of integrals, series, and products*. Academic Press, 2007.

[120] "Ixolar solar bit solar cell." `http://ixapps.ixys.com/DataSheet/KXOB22-04X3F_Nov16.pdf`. [Online; accessed 15/08/2017].

[121] E. W. Weisstein, "Circle-circle intersection." `http://mathworld.wolfram.com/Circle-CircleIntersection.html`. [Online; accessed 15/09/2017].