Ultra-reliable Network-controlled D2D

Nicolas Malm

School of Electrical Engineering

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo 23.11.2015

Thesis supervisor:

Prof. Olav Tirkkonen

Thesis advisor:

D.Sc. (Tech.) Kalle Ruttik

Aalto University
School of Electrical
Engineering

AALTO UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

ABSTRACT OF THE
MASTER'S THESIS

Author: Nicolas Malm

Title: Ultra-reliable Network-controlled D2D

Date: 23.11.2015 Language: English Number of pages: 7+67

Department of Communications and Networking

Professorship: Communication Theory

Supervisor: Prof. Olav Tirkkonen

Advisor: D.Sc. (Tech.) Kalle Ruttik

Fifth generation cellular networks aim to provide new types of services. Prominent amongst these are industrial automation and vehicle-to-vehicle communications. Such new use cases demand lower latencies and higher reliability along with greater flexibility than current and past generations of cellular technologies allow. Enabling these new service types requires the introduction of device-to-device communications (D2D). This work investigated network-controlled D2D schemes wherein cellular base stations retain control over spectrum usage. D2D nodes assemble into clusters. Each D2D cluster then organises itself as it sees fit within the constraints imposed by the cellular network. A review of proposed D2D control schemes was conducted to identify pertinent interference issues. Measurements were then devised to empirically collect quantitative data on the impact of this interference.

Measurements were conducted using a software-defined radio (SDR) platform. An SDR based system was selected to enable a low cost and highly flexible iterative approach to development while still providing the accuracy of real-world measurement. D2D functionality was added to the chosen SDR system with the essential parts of Long Term Evolution Release 8 implemented. Two series of measurements were performed. The first aimed to determine the adjacent channel interference impact of a cellular user being located near a D2D receiver. The second measurement series collected data on the co-channel interference of spectrum re-use between a D2D link and a moving cellular transmitter. Based on these measurements it was determined that D2D communications within a cellular system is feasible. Furthermore, the required frequency of channel state information reporting as a function of node velocity was determined.

Keywords: D2D, 5G, reliability, SDR, low-latency communications

AALTO-YLIOPISTO
SÄHKÖTEKNIIKAN KORKEAKOULU

DIPLOMITYÖN
TIIVISTELMÄ

Tekijä: Nicolas Malm		
Työn nimi: Korkean luotettavuuden verkkohallinteiset laitteiden väliset yhteydet		
Päivämäärä: 23.11.2015	Kieli: Englanti	Sivumäärä: 7+67
Tietoliikenne- ja tietoverkkotekniikan laitos		
Professuuri: Tietoliikenneteoria		
Työn valvoja: Prof. Olav Tirkkonen		
Työn ohjaaja: TkT Kalle Ruttik		

Viidennen sukupolven solukkoverkoilla pyritään mahdollistamaan uudentyyppisiä palveluja kuten teollisuusautomatiikkaa ja ajoneuvojen välistä viestintää. Tämänkaltaiset uudet käyttötarkoitukset vaativat lyhyempien viiveiden ja korkeamman luotettavuuden ohella myös suurempaa joustavuutta kuin minkä nykyisen sukupolven matkapuhelinverkkoteknologiat sallivat. Edellä mainittujen uusien palvelujen toteuttaminen vaatii suoria laitteiden välisiä yhteyksiä (engl. D2D). Tässä diplomityössä keskityttiin tutkimaan verkkohallinteisia D2D-rakenteita, joissa solukkoverkko hallinnoi spektrin käyttöä. D2D-päätteet liittyvät yhteen muodostaakseen klustereita, jotka hallinnoivat sisäistä tietoliikennettään parhaaksi katsomallaan tavalla solukkoverkon asettamien rajoitusten puitteissa. Kirjallisuuskatsauksen avulla selvitettiin aiemmissa tutkimuksissa esitetyille D2D-ratkaisuille yhteiset interferenssiongelmat. Näiden vaikutusta ja suuruutta tutkittiin mittausten avulla.

Mittaukset toteutettiin ohjelmistoradioalustan (engl. SDR) avulla. SDR-pohjaisen järjestelmän käyttö mahdollisti edullisen ja joustavan tavan kerätä empiirisiä mittaustuloksia. D2D-toiminnallisuus lisättiin Long Term Evolution Release 8:n olennaiset ominaisuudet omaavaan alustaan. Tällä alustalla toteutettiin kaksi mittaussarjaa. Ensimmäisellä kerättiin tuloksia viereisellä kanavalla toimivan matkapuhelimen D2D-vastaanottimelle aiheuttamasta interferenssistä näiden ollessa toistensa läheisyydessä. Toisella mittaussarjalla selvitettiin samalla kanavalla toimivan D2D-yhteyden ja liikkuvan matkapuhelimen välistä interferenssiä. Mittausten perusteella todettiin D2D-toiminnallisuuden lisäämisen solukkoverkkoon olevan mahdollista. Lisäksi laskettiin vaadittava kanavalaadun päivitystiheys päätteiden nopeuden funktiona.

Avainsanat: D2D, 5G, luotettavuus, SDR, lyhytviive yhteydet

Preface

I would like to thank my thesis supervisor Prof. Olav Tirkkonen and advisor D.Sc. (Tech.) Kalle Ruttik for their valuable input as well as for providing me a challenging and engaging work environment. In addition, my colleagues working on the ARF platform have my appreciation for their help in enabling me to perform the measurements contained in this work.

I would further like to thank my parents and friends in helping me get to this point in life. I greatly appreciate the zealous, diligent and insightful feedback offered by Kevin Salvesen. Last but not least, I offer my gratitude to Rekik Haile Magicho for her support, encouragement and in particular for giving me a push to apply for the summer job that would eventually lead to this thesis position.

Otaniemi, 23.11.2015

Nicolas Malm

Contents

Abbreviations

ACI	Adjacent Channel Interference
ARF	Aalto Radio Framework
ARQ	Automatic Repeat Request
BCH	Broadcast Channel
BER	Bit Error Rate
BS	Base Station
C-RAN	Cloud Radio Access Network
C-RNTI	Cell Radio Network Temporary Identifier
CoS	Class of Service
CPU	Central Processing Unit
CQI	Channel Quality Indicator
D2D	Device-to-device
DRX	Discontinuous Reception
eNodeB	Evolved Node B
GPP	General Purpose Processor
GPRS	General Packet Radio System
GPS	Global Positioning System
GTP-U	GPRS Tunnelling Protocol User Plane
HARQ	Hybrid Automatic Repeat Request
IP	Internet Protocol
ISM	industrial, Scientific and Medical
LTE	Long Term Evolution
MAC	Medium Access Control
MCSRV	Modulation, Coding Scheme and Redundancy Value
MIMO	Multiple-input and Multiple-output
MS	Mobile Station
OCI	Operation and Control Interface
OFDM	Orthogonal Frequency-division Multiplexing
OS	Operating System
PHY	Physical Layer
PSS	Primary Synchronization Signal
QoS	Quality of Service
RACH	Random Access Channel
RB	Resource Block
RE	Resource Element
RF	Radio Frequency
RLC	Radio Link Control
RRH	Remote Radio Head
RRM	Radio Resource Management
SDR	Software-defined Radio
SINR	Signal to Interference and Noise Ratio
SPS	Semi-persistent Allocation
SSS	Secondary Synchronization Signal
TTI	Transmission Time Interval
TXP	Transmit Power
UE	User Equipment
V2V	Vehicle-to-vehicle
VHEL	Virtual Hardware Enhancement Layer
WI	Work Item

1 Introduction

Cellular communication systems have become ubiquitous parts of modern life. Since their introduction, networks have become faster, more affordable and coverage has increased greatly. Development of the fifth generation (5G) of cellular technologies aims to improve upon these traditional goals but also expand the scope of possible use cases. Prominent amongst these are vehicle-to-vehicle (V2V) and industrial automation applications [47]. Both require communication networks providing characteristics differing from the traditional offerings of legacy cellular technologies catering to consumers' needs. In particular, two factors have been identified as key to enabling new large classes of applications: latency and reliability. Both safety-of-life orientated V2V and industrial automation require higher degrees of reliability than can be cost effectively offered — if at all — by current networks. Furthermore, since the aim is to coordinate machinery involved in high speed operations, latency must be limited to very short delays.

One part of the 5G approach to enabling the above mentioned goals is device-to-device communication (D2D) [47]. In particular, allowing mobile stations (MS) to exchange messages directly with each other while remaining under the control of the serving base station (BS) without resorting to switching to non-cellular technologies such as wireless local area networks (WLAN). This concept is known as network-controlled D2D and utilises direct connections between end devices located in close proximity to each other without relaying through the serving BS, while retaining the control signalling. In addition to enabling lower latencies and higher reliability, D2D offers the prospect of increased spectral efficiency through the re-use of frequencies within one cell. Resources used by local communications in one part of the cell may be simultaneously in use in another. Current 5G research envisions a drastic increase in the number of connected nodes [47] making spectrum an even scarcer resource than it has been until now, thus placing a premium on spectral efficiency.

This work investigates network-controlled D2D communication. Contributions are made as to the practical impact of implementing D2D schemes as found in the literature. LTE Release 8 is used as a representative current generation standard within which D2D links are hosted. Devices present in the network are assumed to operate in the same frequency band and be mobile inside the same geographical area. Such an arrangement carries the potential to introduce interference between the cellular and D2D nodes. Measurements carried out for this work aim to quantify the magnitude of such interference and its effect on the feasibility of enabling high-reliability communications directly between devices. In order to assess this, a proof-of-concept implementation of D2D is carried out to both obtain the above mentioned interference measurements and to validate the concept of direct links sharing time-frequency resources with cellular ones. Based on this data a model for estimating minimum channel quality reports is developed.

Chapter 3 starts with a discussion of the fundamental challenges faced by high reliability links. This is due to neither LTE nor any of its predecessor cellular technology having incorporated D2D communications, leaving many details and technical solutions as open research questions. Some consequences of the introduction of such new interaction patterns between MSs can however be foreseen [20]. Regardless of the exact technical minutiae of each standard to come, certain general problems will need solving to enable high-reliability wireless communications as detailed in this chapter. In order to identify the problems common to most if not all proposals, this work presents the current generation LTE standard's relevant features emphasising those most pertinent to D2D communication in chapter 2 and an overview of currently proposed approaches and designs found in the literature for implementing network-controlled D2D in cellular networks in chapter 4. Based on this review and classification, two metrics — adjacent channel interference and co-channel interference — are formulated to assess the feasibility of D2D in cellular networks.

Assessing the impact of adding D2D functionality into cellular networks can be investigated using several methods ranging from analytical studies to field measurements. Using actual hardware to obtain data yields more accurate data than simulations due to the near-impossibility of accounting for all factors in the latter. Moreover, when dealing with new concepts it is difficult to know for certain which factors need to be modelled and to what level of accuracy. On the other hand, hardware-only solutions are costly and inflexible. In order to circumvent the high costs of going down this route, for the present work a software-defined radio (SDR) platform developed by Aalto University called ARF was used instead as the main research tool. Being entirely software-based enabled insertion of the D2D functionality created for this study into the existing code base, leveraging the existing LTE code. This solution enabled real-world testing on hardware yet with the flexibility to iterate offered by software-based approaches. Chapter 5 discusses the challenges and adopted solutions in the creation of the ARF SDR platform and its D2D extension.

Chapters 6 and 7 present the methodology used to assess the impact of D2D introduction as well as the results obtained. The objective of this work is to establish whether the addition of D2D nodes into a cellular communication system leads to prohibitive levels of interference or whether this can be managed and if yes, under what circumstances. The aim is not to propose a particular protocol or approach but focuses instead on quantifying the interference impact of additional node types and thus under what conditions co-existence with traditional cellular MSs is feasible if possible at all. Public demonstrations of the created system showed D2D communications to be feasible in cellular networks.

Furthermore, the data obtained will be refined and developed into system parameter limits using a propagation model. From the raw interference level results, distance and speed of the nodes will be considered to establish bounds for maintaining the functionality of the communication channel. These values will then be used to establish the required reporting frequency of channel quality information D2D clusters must observe in order to avoid disruptions to their ability to exchange information reliably in the presence of mobile potentially interfering nodes. In particular, use cases requiring high levels of reliability cannot wait for outages to occur but must instead pro-actively avoid them. Obtaining such operational parameter bounds forms a critical part of system design. Any suggested solution unable to fit into the constraints imposed by the interference environment will suffer from a reduced number of deployment options and use cases or in the worst case, be entirely unsuitable. Moreover, established bounds will help guide future protocol development in the directions most likely to be fruitful. The results obtained show there exists realistic deployment scenarios in terms of node distances and speeds under which direct communication services can be provided.

2 LTE

Long-term Evolution (LTE) is a fourth generation cellular network standard. Standardisation work takes place within the 3GPP - The 3rd Generation Partnership Project. The first version of the standard is Release 8, adopted in 2008 [1]. In this work, LTE serves as the framework within which *device-to-device* (D2D) functionality is integrated. The relevant technical details of the standard [16, 17] are discussed in the remainder of this chapter starting with the high level design and progressing towards the physical layer.

2.1 Architecture and design

The high-level topology of LTE networks resembles its forerunners in being roughly divided into three parts: the *air interface*, the *control plane* of the *core network* and *gateways* (user plane connection to other networks). The air interface is comprised of cells under the control of a *base station* (BS) called an *eNodeB* in LTE. These connect to each other directly instead of through a shared controller as in previous generations. The core network's control plane provides supporting functionality for the network. These include authentication, authorization, billing and admission control. Finally, gateways connect the mobile network to others in order to provide inter-networking.

LTE's high-level design (Figure 1) differs from previous generations in two major ways. Firstly, it abandons circuit switching for voice calls in favour of an entirely IP-packet based architecture. Growth in data traffic lead to this choice. Ericsson reports that mobile data traffic doubled from 2011 to 2012 [15] while Cisco estimates suggest that worldwide data traffic will reach 24.3 exabytes by 2019 [11]. Such growth requires a new approach capable of coping with it. An IP-centric design allows a flat unified backbone architecture capable of connecting to the Internet through a simpler gateway than previous cellular technologies. Since it is likely that the vast majority of the data traffic generated by users on their mobile devices leaves their operator's network for the Internet, this design choice optimises the capacity of cellular backbone networks by reducing the number of intermediate hops.

The second major innovation compared to earlier generations is LTE's forward looking design. It serves as the basis for LTE-Advanced (LTE-A). Allowing updates to deployed cellular networks reduces costs for operators as previously purchased equipment does not become obsolete when updates to the standard are published. Furthermore, since initial LTE releases did not satisfy the technical requirements of fourth generation (4G) technology [31] as set by the International Telecommunication Union (ITU) an upgrade path was deemed desirable to compete with other technologies. LTE-A (Release 10) was submitted by the 3GPP to the ITU as a 4G candidate technology with final acceptance in 2010 [32].

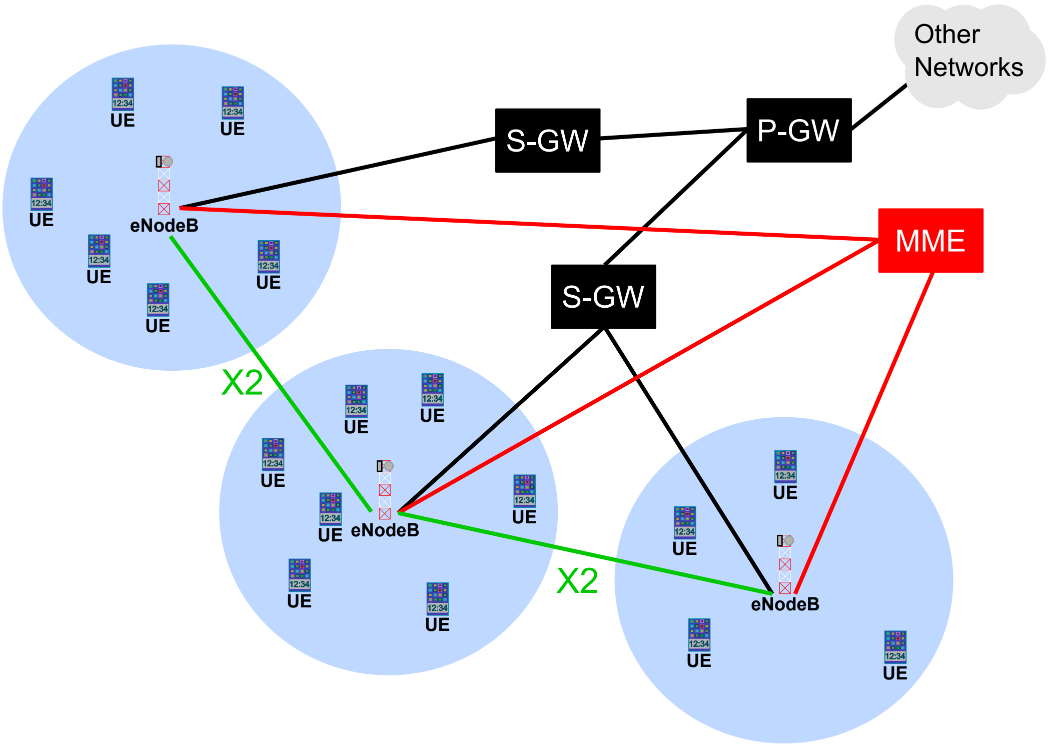

Figure 1: LTE system architecture.

Overall, LTE aims to increase spectral efficiency to support the rapid growth in mobile data while offering lower latency to end users. The motivation behind these improvements is to bring an improved quality of service to mobile communication in order to compete with fixed line alternatives. Operators' own networks obey this trend as well in that their core network evolves at a slower pace than the air interface technologies they use. LTE's flat all-IP core design allows for easier upgrades to the air interface alone as it does not depend on controllers for BSs.

2.2 eNodeBs

eNodeBs each manage one cell and the *mobile stations* (MS) - termed *user equipments* (UE) in LTE — within it. All communication occurs only with express permission of the eNodeB. These operate as independent entities unlike previous generations which relied on controllers responsible for multiple BSs. This design helps reduce latency as decisions may be made locally instead of being sent further into the operator's network. Admission control as well as resource allocation for data bearers are still made in the core network but all fast scheduling decision are the exclusive responsibility of the eNodeB.

On the data plane, they communicate directly with the serving gateway (S-GW) they are attached to for forwarding data in and out of the mobile backhaul. IP-traffic from and to UEs undergoes tunnelling while transiting through the core network of the operator. The tunnelling protocol used is *GPRS Tunnelling Protocol User Plane* (GTP-U) where GPRS stands for General Packet Radio System. Tunnelling is used to encapsulate the UE generated payload traffic without modification during its transit through the mobile backhaul. GTP-U enables data to be addressed to a particular UE using their *Cell Radio Network Temporary Identifier* (C-RNTI) and furthermore targeting a specified class of service called a *bearer*. Such metadata enables the core network and the radio interface to treat data packets consistently with regards to their service levels. Bearers therefore serve as an abstraction for the treatment of packet at each step of their transmission path. When GTP-U packets reach an eNodeB it will convert the QoS traffic class it is assigned to into a scheduling priority (i.e., bearer) on the air interface for the destination UE.

Handovers in LTE can be performed by the eNodeBs themselves in cooperation with their neighbours in addition to the traditional SGW-assisted handovers of previous generations of cellular technology. Direct inter-eNodeB handovers operate via an exchange of information regarding the UE to be handed over using an interface called *X2*. X2 interfaces link an eNodeB to one or more of its neighbours. These connections carry both signalling between the base stations as well as user plane data traffic being forwarded during the completion of a handover. After completion, user plane data will be sent directly to the new BS by the S-GW.

In addition to handovers X2 interfaces may also be used by eNodeBs to exchange load and interference level reports. These aid in the tuning of parameters related to handover thresholds. By exchanging such information an LTE network gains the ability to balance the load across different cells more effectively. Furthermore, eNodeBs exchanging interference level reports and modifying their behaviour accordingly reduce the amount of interference in the neighbouring cells and thus optimise overall capacity.

2.3 Air interface

The air interface of LTE carries the name of *Uu* [16]. LTE downlink utilises *orthogonal frequency-division multiple access* (OFDMA) while the uplink uses its single carrier variant SC-OFDMA. It supports both *frequency-division duplex* (FDD) and *time-division duplex* (TDD) mode of operation. In the former, the downlink and uplink are separated in frequency whereas in TDD, a single frequency is split according to the *uplink-downlink configuration* (UDConfig). These configuration serve to define the relative proportion of uplink and downlink capacity available. Seven such configurations exist. They differ in the number of downlink (D subframes) and uplink (U subframes) resources they provide. In addition, TDD introduces time gaps between transmission and reception operations when switching from downlink to uplink carried in S subframes. These exist for the purpose of enabling UE hardware to switch from reception to transmission mode. Consequently, they are not used in FDD operation. Selection of the UDConfig parameter constitutes an important tool in network planning. The amount of uplink and downlink resources available can be chosen to match the anticipated demand by picking an appropriate UDConfig. Bandwidths supported range from 1.4 to 20 MHz for both FDD and TDD.

LTE Uu divides time into *frames*, *subframes* and *slots* [16]. Each frame lasts 10ms and subdivides into ten one millisecond subframes. These in turn subdivide into two slots each lasting for seven OFDMA symbols. Frames and subframes are used as the unit of time keeping in LTE. They are used to define the time for one-off events(e.g.: single transmission grants) and recurring events such as system information channel broadcasts. During FDD operation, each subframe on a given carrier frequency is either downlink or uplink whereas for TDD operation the subframe types within a frame vary according to the UDConfig parameter mentioned previously. All time domain scheduling and reporting in LTE occurs at subframe granularity. Since retransmissions occur after some multiple of the subframe period, shorter *transmission time intervals* (TTI) result in lower latencies thus motivating the choice for LTE of 1ms as opposed to the longer values of previous generations. Moreover, *channel quality indicators* (CQI) reporting opportunities increase as there are more TTIs enabling UEs to report more often leading to improved accuracy of channel estimates which in turn leads to higher throughputs.

Frequency resources are divided into subcarriers spaced 15kHz apart [16]. A group of twelve subcarriers in frequency and one subframe in time is termed a *resource block* (RB). The number of RBs available for scheduling in each subframe depends on the bandwidth allocated to the cell in question. Table 1 lists the available resource block counts for each LTE supported cell bandwidth. RBs subdivide into *resource elements* (RE) such that there are twelve REs per OFDMA symbol for a total 168 REs per RB.

Table 1: LTE resource block counts by cell bandwidth.

Bandwidth (MHz)	1.4	3	5	10	15	20
Resource block count	6	15	25	50	75	100

The totality of the resource elements within one subframe constitute the *resource grid*. Different physical channels map to different REs each carrying one modulated symbol. Physical channels thus map to different parts of the grid and need not be contiguous in either time or frequency. Pilot values used for channel estimation are distributed with a regular pattern across the whole resource grid. This provides a more accurate estimate over the totality of resource than if they were grouped. Frequency selective impairments to the channel can thus be more precisely detected and compensated for. Each physical channel can be modulated independently thus producing differing data rates and levels of reliability.

In terms of time-wise scheduling, system channels take precedence over data user data channels. Every time, a system channel is scheduled to be transmitted, the REs it uses will be unavailable for user data. Since system channels are not transmitted during each subframe, the data capacity varies slightly during frames and even from frame to frame. Re-transmission opportunities also vary for the same reason. Re-transmission times are fixed to four subframes ahead in FDD while in TDD, the timings depend on the UDConfig is use. The LTE mechanism for increasing channel reliability is the *hybrid automatic repeat request* (HARQ) [18]. Upon reception of a *negative acknowledgement* (NACK), it performs a re-transmission. The receiver will then attempt to combine the re-transmitted information with the previous — failed — attempt in order to maximise the amount of information available during decoding.

Due to this complex time-variant structure combined with the requirements of OFDMA LTE possesses a robust synchronisation mechanism. UEs begin by performing cell search to detect the presence of *broadcast channels* (BCH). Since the UE does not yet know the cell bandwidth in use at this point, the BCH always occupies the six centremost RBs. After detection of the presence of a cell, the UE proceeds to search for the *primary synchronization signal* (PSS). This will provide frame-level alignment since the PSS is always at the same predetermined position with the resource grid. Next the UE searches for the *secondary synchronization signal* (SSS) to determine which half of the frame the PSS found lies in. Like the PSS, the SSS sequences occur at a fixed position within the frame. Unlike the PSS however, the two SSS sequences differ to enable the determination of the half-frame in question. PSS and SSS signals appear in each frame. UEs continually adjust their sample-level synchronisation to the eNodeB by tracking the drift of the synchronisation signals from frame to frame.

3 High-reliability wireless links

The design of high-reliability wireless links requires mitigation of the adverse characteristics of their operating environment. These inherent properties set the constraints to which the system must adapt. Required levels of reliability are determined by the intended application of the system being designed. Similarly, the operating environment will determine how difficult it will be to achieve this target level of dependability. As such, high-reliability designs must therefore be constructed around combating the undesirable characteristics present in the following sections. This chapter presents the issues faced by wireless links in general and high-reliability ones in particular. Sections 3.1 and 3.2 discuss the problems of fading and interference. Following this, section 3.3 will present different definitions of reliability as well metrics for measuring reliability. The remainder of this chapter is dedicated to techniques used to mitigate the problems described in sections 3.1 and 3.2.

3.1 Fading

The first type of challenge to overcome is *fading* defined as the stochastic time variance of the propagation environment between transmitter and receiver [5]. An identical transmission that would have reached the receiver on the previous transmission opportunity might not do so on the current one or vice-versa. This creates uncertainty that a reliable design must address. It should be noted that fading affects a single link and is therefore independent of the number of terminals present in a network. Increasing the probability of correct reception by the receiver involves increasing the redundancy of the information transmitted. This effectively means trading-off capacity for reliability. Multitudes of solutions for overcoming fading have been proposed such as algorithmic fading prediction [14], waveform design [35], iterative equalizers [51] and adaptive modulation schemes [73] to name but a few.

Fading can also be beneficial in certain situations by reducing the power of transmissions from transmitters other than the desired one. This proves to be extremely helpful in situations where communicating parties are located in close geographical proximity. With high fading, the frequency resources they employ can be re-used with modest physical separation. Cellular systems — and particularly D2D links therein — may use low transmit power in combination with fading to allow re-use of the time-frequency resources allocated to a communicating pair in another part of the same cell without excessive interference levels. Since in cellular-integrated D2D systems the equipment is shared between both node types, it may be possible for the D2D side to take advantage of the existing measurements taken by the cellular part.

Since cellular-integrated D2D systems will employ the same frequencies as the parent system, existing measurements data and methodologies for estimating the effects of fading still apply, thus simplifying design and operation of systems and networks. Additional D2D specific measurement approaches may still prove to be required or beneficial for certain D2D schemes. Such considerations lie outside the scope of this work.

3.2 Interference

The second major problem to combat is *interference*. Unlike fading, interference levels depend on the number of terminals present in a given geographical area. The solutions therefore relate more to network management although other types of solutions have been presented as well (for example, interference cancellation [72]). Managing the detrimental effects of interference includes both prevention as well as mitigation and recovery when it does occur. In systems with a centralized control node, one element carries the responsibility of avoiding overlapping transmission by acting as an arbitrator of time-frequency resource access. The other nodes must request permission to transmit and only do so when authorized. It should be noted however that the first transmission of a node in a network with variable membership cannot, by very nature, follow this procedure as a new node is unknown to the controller. Often, system designs include special resources for such transmission are termed random access resources. These are known to every node through some method, such as from the relevant specification or from a system parameter broadcast resource. The central control node need not aim for zero-interference levels, however. It may be advantageous to allow simultaneous usage of the same time-frequency resource when the communication groups experience high attenuation between them. The drop in signal to noise and interference ratio (SINR) may be small compared to the gain of allowing re-use of the resource in question.

Traditionally [24], cellular systems have mostly been concerned with the amount of interference caused to neighbouring cells with mobile stations (MS). This is because the only receiver in a cell during an uplink phase is the BS. In an uplink D2D scheme as considered in this work, the situation changes due to the introduction of additional receivers during uplink phases. D2D transmission occur in parallel to the uplink transmissions from cellular MS to BS. These D2D receivers may be affected by nearby MS transmitting to the base station. Interferers could be on the same channel as in the traditional case or on an adjacent channel depending on the resource subcarrier allocation selected by the serving BS. This latter type of interference — termed adjacent channel interference (ACI) — depends on the amount of transmit power appearing outside of the allocated resources and is hardware dependent.

3.3 Types of reliability

What is understood as reliability will depend on a particular use case's requirements. Different applications judge system performance by different metrics. These might conflict with each other. It is thus difficult or even impossible to design a generic high-reliability scheme. Applications may still be grouped into broad categories by favoured reliability metric. This section discusses the various ways in which reliability may be measured in a communication system.

Latency critical systems require that messages arrive within a given time frame [46]. Late deliveries may contain outdated information that will be discarded by the node — thus wasting transmission capacity — and may even provoke undesired results. This could occur for example in the case of industrial automation where the timing of operations is critical. In order to meet the requirements of latency critical systems, the communication protocol must limit or even completely avoid the use of retransmissions [40] and other mechanisms that retain data after the initial send request received from the application.

The opposite type of scenario to the one above are systems requiring very low — ideally zero — bit error rates in their payloads. Unlike latency critical systems, these high-integrity systems tolerate delays provided the payload arrives intact at the receiver. In the case of energy unconstrained devices, the protocol used to provide reliability may use complex and powerful coding schemes to protect the payload data since the application tolerates processing delays.

A third type of reliability demands the elimination of false positives. Such application may cope with errors, provided that one type of messages or interference does not get mistaken for another type of message. Furthermore, these systems often rely on the detection of communication link failures to take some corrective actions [65]. For example, an industrial automation system may proceed to shutdown in case the connection to its controller is lost to avoid potentially dangerous malfunctions.

Other applications may present a hybrid profile of the above. Some requirements may not be absolute but rather relative to some threshold metric [41]. For instance, a requirement might be to have correct reception of x% of messages transmitted within a time window with no more than y milliseconds of delay. It should also be noted that production systems often perform additional duties to their core mission. These tasks include performance reporting, logging and software updates. The aforementioned do not require the same level of protection as the mission critical tasks and may therefore be protected by weaker mechanisms.

3.4 Link protection strategies

Link protection may be implemented in different ways depending on the available resources and the level of reliability required for a given application. This section discusses methods and solutions that have been proposed in the literature.

Some use-cases place a premium on reliable delivery of all or part of the transmitted data. In such cases, concessions regarding throughput or latency are often made. This could be partial as, for example, in the case of control channels having a higher level of protection than data channels. Alternatively, the system may provide different *classes of services* (CoS) for different data streams or users. In either case, the mechanisms used involve trade-offs. Re-transmissions increase the odds of correct data delivery at the cost of latency while forward error correction (FEC) codes reduce the number of payload bits per time unit.

In addition to the above mentioned methods, one can also increase the level of link protection by trading-off system capacity for reliability. In such a diversity scheme, the overall number of resources used per user or data stream or a combination of both is increased, thus resulting in a corresponding decrease in the number of entities receiving service simultaneously. This section will discuss methods that fall into this category of reliability mechanisms.

3.4.1 Stand-by allocations

Stand-by allocations are stand-by resources created — but not normally used — to handle failures in a primary (protected) resource by providing diversity in the set of available resources. Alternatively, both resources could be used to transmit the same information simultaneously. This type of diversity scheme helps ensure correct reception of the data in case one of the resources becomes impaired. In a wireless system, such diversity could be provided by a channel on another frequency. This would allow a quick switch-over in case frequency selective fading or interference [46] occurs on the primary resource. When conditions remain within acceptable limits in both channels, the system monitors both links. If the primary deteriorates below a set threshold, operations will switch to the back-up. Concurrent allocation of a new reserve resource will ensure the maintenance of the quality guarantee provided by the system. Note that the new back-up channel may be the original primary channel. Varying the number of reserve resources effectively configures different levels of reliability.

The scheme can be extended to cover multiple classes of services. In this case, traffic belonging to a higher CoS pre-empts lower ones. Effectively, this entails that all available resources of a lower CoS are back-up resources for a higher one. Care must be taken in scheduling to prevent a cascade of reassignments as a data stream of the highest CoS pushes out one from a lower that then evicts an even lower one and so on.

Stand-by allocation must be checked regularly for continued applicability. As either the participating nodes or others may move after the initial allocation, interference levels undergo constant evolution. Testing of backup allocation channel quality between high-reliability nodes constitutes a form of overhead that must be weighted against the gain. The optimal trade-off will depend on the required level of reliability. In general, guaranteeing that such changes do not negatively impact the protected link is impossible due to the possible presence of nodes in a disconnected state. These may wake-up unexpectedly at any time. These might not coincide with the back channel measurement intervals selected for or by the nodes of the high-reliability link and could therefore become "invisible" to the BS. The latter may however introduce constraints to its own allocation policy such that it can minimise the impact of such unanticipated wake-ups.

The first and most important of these constraints is to restrict the time-domain resources it considers for protected channel allocation to those transmission time intervals (TTI) that do not contain any random access channel (RACH) resources. This eliminates any in-band interference from awakening nodes. The second technique is to group all protected links into the same TTI. Since these links are negotiated for longer periods of time and with predictable timings, the BS will not have to contend with unexpected events. This may also be extended to cellular users using semi-persistent allocations (SPS) [16].

Furthermore, the BS must consider the effect of periodic system channels. While some have consistent load (for example, system parameter broadcast channels), others — such as paging channels — will not. These are however easier for the BS to plan around since it possesses full knowledge and control over their timing and content.

The quantity of extra resource to assign constitutes not only a technical issue but also an economic one. Due to the scarcity of spectrum, the level of gained added reliability must be weighed against both other solutions providing similar dependability and the return on this investment. Detailed discussion of such considerations lies beyond the scope of this work.

3.4.2 Neighbouring resource measurements

Another approach to improving link reliability is to not only consider the resources of the protected data stream but also the neighbouring ones as well. If transmitters operating on contiguous frequencies physically approach one another, levels of adjacent channel interference may rise above acceptable levels. Continuous monitoring of adjacent frequency resource helps prevent this from occurring.

The problem is not completely eliminated by this straightforward monitoring solution, however. It may be possible that a nearby transmitter remains mostly silent — most importantly during the measurement intervals — and therefore undetected. The majority of cellular system will put inactive transmitters into a dedicated idle states from which resumption of transmissions can only occur on dedicated separate resources, limiting the risk of this scenario. Each application must weight the acceptability of this low probability event versus the impact it would have to determine whether the situation is acceptable or not and only schedule when it is not. Indeed, while the BS can schedule measurements for those times when it commands MSs to wake up, it cannot anticipate when random access resources will be used by an MS. These emissions appear completely spurious from the point-of-view of a D2D node.

Further compounding this problem, there exists the possibility that a nearby D2D group employs a protocol with very irregular transmission patterns. To mitigate this type of inter-D2D interference the BS can do little by itself but command low power ceilings. Monitoring this type of interference requires some degree of knowledge of the protocol being used inside a D2D group. Information about measurement opportunities could be provided to the BS as part of the D2D permission negotiation with the network.

3.4.3 Cellular re-scheduling

For D2D systems operating inside a cellular one, there exists a particular class of solutions for the protection of high-reliability D2D links. Instead of affording the protected link better treatment, the other (i.e., cellular) ones are scheduled around the needs of the priority one to ensure its needs are met. Such a system relies on the fact that mobile stations (MS) only transmit when permitted to do so by their serving base station (BS). As such, they can easily be re-allocated since this does not differ from their normal mode of operation. In order for the BS to know when a change becomes necessary, the D2D nodes involved in the high-reliability link send reports detailing the encountered channel conditions in both directions of their link. Should these indicate a worsening of quality approaching unacceptable levels, the BS changes the scheduling of the disturbing MS to clear a protected zone around the D2D receiver [74]. Furthermore, the BS ensures through configuration parameters that no system resource (for example, a random access channel) overlaps with the high-reliability link.

Should the channel between participating nodes in a D2D deteriorate for reasons other than external interference, the BS may assign traditional cellular resources for use by the cluster. These resources will taken from the same pool as the one used by the traditional MSs. Assignment of resources occurs as if the nodes were cellular ones. The quantity of resources requested from the BS will depend on the number of nodes in D2D outage as well as their required throughput. This enables the connection to endure deteriorations in their direct D2D link as well as increased physical separation in the case of mobile D2D nodes. When conditions permit direct communication may be resumed and cellular resources released. Such a solution carries the caveat that the latency of cellular BS-relayed communication proves acceptable to the application in question.

The approach of re-scheduling cellular users may be extended to multiple BSs as well. Utilising the X2-interface, two BSs can agree to perform a handover of a cellular user from one to the other (see 2.2). One scenario in which this solution presents significant benefits arises when the channel between the target BS of the handover and the handed over cellular MS requires less transmission power. It is also possible to extend this scheme further and introduce pro-active BS co-ordination in an attempt to prevent interference in addition to correcting cases where it arises.

The above mentioned case might present itself when a cellular MS resides at the edge of one cell while the D2D link resides over the border in a neighbouring cell. Since nodes in a cellular network are typically mobile, such situation may arise at any time. By pro-actively exchanging reports received by MSs neighbouring BSs may select time-frequency resource in a manner conducive to minimising the risk of conflict. This type of neighbour aware scheduling could potentially reduce the interference suffered by both the cellular and D2D users. It should be noted however that this solution will be most effective when neighbouring cells operate on the same channel. Some benefit may still be realised even when nodes operate on different channels since cellular transmitters may saturate D2D receivers when located nearby at the cell edge where transmission power is greatest.

3.4.4 Admission control

Interference control can also be performed by careful distribution of users amongst the available time-frequency resources. When the network deems the latter insufficient, request for services by new users are denied. Alternatively, some users may be afforded privileged access to resources in which case others will have their allocation reduced or even completely withdrawn. If multiple BSs are located nearby, it may be possible to move a node requesting resources from one to the other in an effort to balance the load and thus aim to reduce the interference level inside any given BS.

Handover measurements could be extended to uplink time slots in order to help BSs determine whether an incoming cellular MS is located physically near to an active D2D cluster. Such information would be helpful in avoiding assigning the new user into time-frequency resources susceptible to cause interference to a protected link.

D2D cluster themselves might perform some form of admission control based on their allocated resources. When a new node requests to join, the existing nodes use some method — such as a master device — to asses whether the join request may be granted. In cases where this is not feasible, the demand may either be rejected outright or more resources may be requested from the controlling BS. If available, the D2D cluster can accept the new node using these newly granted resources.

4 Device-to-device communication

Traditional cellular networks are built around entities called base stations (BS) [3, 24]. These are responsible for a given geographical area known as a cell. All traffic within a cell has classically transited through the BS. Direct links between mobile stations (MS) are known as a device-to-device (D2D) links. They can either be freely created, managed and terminated by the MSs independently or they can be overseen by the BS. The latter case is known as network-controlled D2D and is the focus of this work. More specifically, *in-band underlay* type D2D communications are considered. In-band means that the same spectrum is used for both D2D and cellular communications as opposed to segregating them to different frequency bands. Underlay indicates simultaneous use of spectrum by both D2D and cellular links.

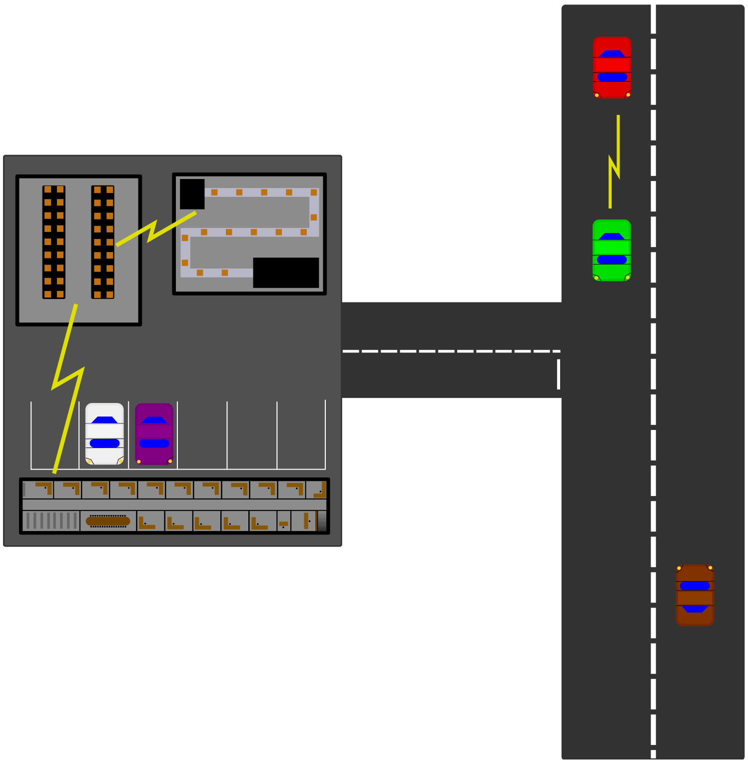

Figure 2: D2D use case illustrating vehicle-to-vehicle communications and wireless industrial automation.

D2D can prove beneficial in applications benefiting from decreased latency or increased reliability as detailed in section 4.1. Figure 2 presents two such use cases. The first one is vehicle-to-vehicle (V2V) communication. Intra-vehicular communication links aim to provide increased levels of safety through shorter response times as well as improved traffic fluidity through increased co-ordination [42]. In Figure 2, the green car will turn onto the side road. As it slows down in preparation for the turn, it will message the red car to notify it of this fact. Such early warning allows the necessary safety gap to be maintained as the deceleration can be anticipated by the following vehicle [37]. The second use case illustrated in the figure is wireless industrial automation. Allowing machinery to co-ordinate with each other can provide greater efficiency. Compared to wired solutions wireless ones enable mobile equipment to be integrated into these communications and in general increases the flexibility in the placement of machinery [70]. Furthermore, contactless communication technologies reduces wear on connectors and cables used to link moving components together [4].

The theoretical review and analysis presented in this chapter was performed in order to identify which issues are common to most D2D control schemes and to identify practical measurements pertinent to them. The measurements presented in chapter 6 aim to obtain quantitative data as to the feasibility of network controlled D2D in general. This chapter begins by discussing the motivation and potential alternatives to D2D communications. Next, the required high-level tasks of discovery, link monitoring and node control will be presented. The last section of this chapter describes the approach used to obtain measurement data for this work.

4.1 Benefits of D2D

D2D links provide several benefits to both the network and individual nodes served by it. Firstly, the end-to-end latency for the D2D devices can be substantially reduced [33]. Eliminating the transit through the BS removes one hop, namely the uplink from the sending node. Additionally, the BS requires a certain amount of time to process the received burst and to prepare the downlink one. This further compounds the latency cost of the extra hop. Further delay may also be introduced by a lack of transmission opportunities for the up- and downlink hops not being available at the optimal time (i.e., right after reception is completed) due to scheduling load.

Secondly, the extra hop through the BS wastes energy compared to a direct transmission [20]. This is especially so when the two D2D nodes are situated close to each other but far from the BS. In such a situation they are forced to use full or close to full transmit power. Even with a low extra energy requirement for transmission, energy must still be expended on the processing, scheduling and re-sending of the information at the BS. Another opportunity for energy saving comes from D2D nodes relaying popular content. The information is first sent from the BS and then relayed through user devices with an opportunity to achieve better energy efficiency [69].

Thirdly, spectral efficiency suffers when the same data undergoes an extra transmission. In cases where the D2D devices' channels to the BS are poor, retransmissions will likely be necessary thus using more capacity. When channel conditions between the D2D devices are favourable, the waste becomes more pronounced. Traditional cellular networks avoid overlap scheduling for MS transmissions since the BS must be able to receive all transmissions. If two MSs were given the same time-frequency resources there would exist the risk of one with a better channel to the BS masking the other. When D2D links are available the BS may in certain circumstances reuse the time-frequency resources allocated to a D2D link elsewhere in its coverage area for a cellular user, thereby increasing the sum-rate capacity of the cell [67].

4.2 Control/data plane functional split

Delegated control D2D as presented in this work extends the split between data and control planes even further than current and past generations of cellular technologies [46]. In case of network-controlled D2D, only signalling traffic will be seen by the base BS while all data — and its processing — will be the responsibility of the participating nodes. This arrangement transforms the BS into a broker of communication resources issuing fixed-term sub-licenses to use part of the spectrum that it had in turn been delegated by the competent regulatory authority. In addition to the arbitration of spectrum usage, the network provides authentication and authorization services in that a MS must first connect to the cellular network in order to be able to make use of the broker function. Synchronization amongst the MSs — and therefore D2D nodes — thus happens implicitly. Figure 3 depicts the control and data connections.

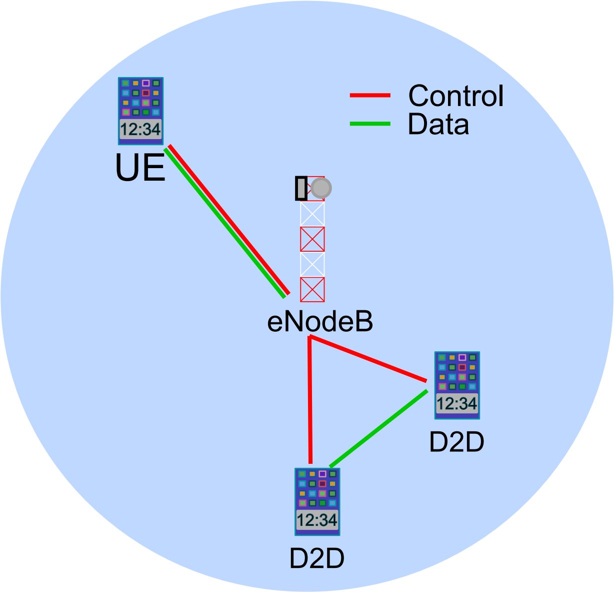

Figure 3: LTE system with D2D nodes.

Due to the above mentioned arbitration role, the nodes participating in the D2D cluster retain the possibility of utilising cellular services as a back-up or in addition to direct communication amongst them. Moreover, the network control prevents outages due to uncontrolled transmissions from other nodes, as would be the case when using *industrial, scientific and medical* (ISM) band technologies. Improving interference conditions typically correlates to reduced numbers of required retransmissions [12]. This combined with the elimination of the two hops to and from the BS reduce the latency of D2D communication vis-à-vis its cellular counterpart.

Previously proposed D2D schemes have focused on methods for avoiding or minimising the interference impact on the cellular technologies. One of the proposed is *offloading*. In offloading, the D2D communication occurs on a separate frequency band after the D2D nodes have discovered each other [49]. Often, the suggested technology for these alternate bands is IEEE 802.11. Offloading suffers from several drawbacks however. Devices must possess multiple radio transceiver chains in order to be able to operate on several frequency bands and possibly using several different technologies simultaneously. Such designs consume more energy and cost more. Secondly, the spectral efficiency achieved suffers due to offloading. This is due to the need for a separate permanent spectrum allocation that may go unused much of the time and in most geographic locations.

Another proposed scheme is the use of additional infrastructure to enable D2D-cellular co-existence. One approach consists in deploying massive *multiple-input and multiple-output* (MIMO) arrays to exploit spatial diversity. MIMO allows for improved directional separation of signals thus helping to spatially filtering out unwanted ones. Other proposals make use of relays or other co-ordination nodes. All such solutions involve additional infrastructure or additional hardware to implement. These once again lead to increased energy consumption and costs.

4.3 Alternative technologies for direct communications

This section presents other current technologies that provide communications capability for devices without the use of control or designated relay nodes. These technologies mainly differ by the use of unlicensed spectrum, limited data rates or higher latencies. They are therefore much more susceptible to interference and delay than cellular based systems. Bluetooth was selected for review due to it being widely available on current UE equipment while Zigbee's intended use for connect smart appliance is quite similar to some of the envisioned D2D scenarios.

4.3.1 Bluetooth

Bluetooth [59] is a short-range wireless communications technology designed for use in personal area networks (PAN) [55, p.249]. It aims to enable several devices in close proximity to exchange data without the need for a cable (originally to replace the wired RS-232 serial communication standard). Most commonly, these devices belong to the PAN of a single user. As such, Bluetooth was designed to operate on an unlicensed band (2.4GHz ISM; shared with 802.11) using point-to-point low speed radios. Its maximum operational range varies depending on the device class from approximately 10m (for class 3) to approximately 100m [62] (for class 1). Furthermore, Bluetooth aims to enable low-energy consumption as it is mainly designed for portable devices with modest battery capacities.

Since Bluetooth is intended to operate without infrastructure present, it discovers other devices through the use of special packets called ID packets. Devices configured to respond to ID packets reply with a message containing a short description of themselves. These replies are then used to populate a list from which a user may select the devices to connect to. Moreover, the list is continuously refreshed since the topology may change often as devices are both mobile and prone to entering sleep mode to conserve energy.

When connected together, the devices form a network that is known as a Piconet [55, p.253] in Bluetooth parlance. One device becomes a master and can control up to seven active slaves and up to 255 parked (i.e., inactive) slaves. Slaves are polled by the master for packets to transmit.

4.3.2 Zigbee

IEEE Standard 802.15.4 [29] — commonly referred to as Zigbee — is technology for building wireless personal area networks [25, pp 93-98]. It aims to provide short-range, long battery life communications to devices requiring only modest amounts of transmission capacity (250kbps raw physical rate). Amongst the most common use cases envisioned for Zigbee are Internet of Things (IoT) and smart infrastructure — such as building automation and sensor networks. Zigbee standards allow operation on unlicensed ISM bands. The main band at 2.4GHz overlaps with the popular 802.11 family of standards. The 802.15.4 standard recommends usage of channels falling in-between the 802.11 ones in order to mitigate interference. Furthermore, Zigbee utilises a CSMA-CA scheme. Random back-off times are waited before transmitting in both the channel free and channel occupied cases to stochastically reduce the number of overlaps. Typical ranges obtainable by Zigbee radios are on the order of 10 to 20 metres.

Zigbee networks are identified by two sequences: the PAN ID and the extended ID. The former is used in most message. Its value is a 16 bit long number randomly selected by the first Zigbee node in a network. This is the first node in a given Zigbee network. The extended ID will be used in case two networks with the same PAN ID come into contact. Such a scenario may occur either when an additional network becomes operational or an existing one expands. In such cases, the extended ID will be used to address nodes of one or the other network in order to assign a new PAN ID. At a finer granularity, individual nodes also have a separate node ID which is also a 16 bit quantity. This value is assigned when they join a network and used to address message to a particular node. Similarly to PAN ID conflicts, if two nodes notice an ID conflict, they use a conflict resolution method based on an extended ID which is assigned by the manufacturer of the chip according to a scheme guaranteeing it to be unique.

Nodes divide into one of three roles. The first node in a network is termed the Zigbee Coordinator (ZC). As mentioned above, it picks the PAN ID and is responsible for conflict detection and resolution relating to it. A second node is the Zigbee Router (ZR). These nodes relay messages from others to create a mesh allowing to extend the range of the network. This effectively trades off capacity for improved coverage. ZRs are not allowed to sleep as they must constantly be ready to relay messages from other nodes. The ZC is also considered a router. The final type of Zigbee node is the Zigbee End Device (ZED). ZEDs communicate exclusively through some ZR designated as their parent node. Since ZEDs do not route messages they are allowed to enter power saving modes in which they are not actively participating in the network for some period of time.

Nodes search for networks by transmitting beacon requests. ZRs within an existing network will reply if within range of the requesting device. The latter will then request permission to join some network it selects. The answer to this request depends on both available capacity to handle additional ZEDs on routers as well as administratively determined settings. New nodes also have the option to join as routers, in which case the previously mentioned limit does not apply. This scheme requires fairly static nodes since network will splinter if nodes lose contact with each other. As such, a persistent network cannot exist independently of current connectivity. If nodes were to come in contact again, they would see a duplicate PAN ID and the networks would proceed to perform conflict resolution — resulting in one changing its ID.

Once a node has joined a network, it performs service discovery. Services are identified by 16 bit numbers assigned by the Zigbee Alliance. Services are grouped into what are termed profiles. These represent sets of capabilities and attributes supported by any node in this group. Such a scheme provides inter-vendor interoperability. Zigbee also support proprietary extensions. Nodes either provide values and actions or make use of them. These are termed servers and clients, respectively.

4.4 D2D discovery

Realizing the benefits of D2D in practice requires measuring and collecting several metrics and elements of state information not traditionally tracked by cellular systems. Nodes wishing to use D2D links to communicate with each other need matching with each other based on common interest. This phase is called discovery. Numerous methods have been proposed; for example: [8, 27, 36, 44]. All proposed discovery schemes must work within the bounds of three main constraints: battery usage, security and scalability.

The first of these — battery life — results from discovery being an extra function that mobile devices will need to perform in the future compared to the current situation. While a cellular MS need only track its location area and current serving BS, D2D nodes will need to also search other suitable D2D nodes to communicate with. Since a node can stay in a disconnected state for extended periods of time while moving, the services discovered the previous time the node was active may no longer be reachable. Similarly, the nodes that previously had advertised a service might themselves have moved since they were discovered. A mobile station (MS) must therefore periodically proceed to update its known catalogue of available services and devices. As the user might have configured his device to react to certain types of announcements, it is not sufficient to activate the discovery mechanism only when interacted with by the user. Discovery mechanisms must balance between the need to be responsive enough to changes in their environment with the low energy consumption benefits of sleep mode. As the number of cellular connected devices is expect to grow to tens of billions [47], broadcast-type systems where all nodes must constantly listen for discovery traffic seem infeasible. Moreover, this type of solution would quickly lead to high interference levels leading devices to boost their own transmission powers to compensate and thus increase their energy consumption.

A second challenge faced by discovery schemes is security. Traditionally, all users present in a cellular network have been authenticated and authorized by the operator of the network. D2D relaxes this security assumption in that although users must authenticate to utilise the network control functionality provided by base stations (BS), operator infrastructure plays no direct role in control of the communication between nodes. Instead of a trusted network provider with whom the user establishes a contractual relationship, the remote node belongs to a "stranger" entity.

Discovery schemes can be categorised based on their degree of decentralisation. The centralised approach places all discovery functions under the spectrum owning operator's direct control. All end user nodes simply indicate their interest in certain types of services or nodes and then wait for the BS to notify them of a suitable match. In such schemes, D2D differs little from traditional cellular communication. The opposite type of solution is represented by a fully externalised approach where nodes discover and negotiate their communication parameters entirely outside of the cellular network. Once this step is completed one or more nodes contact a BS to request communication resources. Externalised service discovery imposes the lightest burden on the operator along with the greatest flexibility since a wide range of discovery methods may be used without requiring any changes on the operator's behalf. Lack of network involvement presents the drawback of providing the least guarantees on the identity of nodes. The network does not filter or validate messages in any way effectively outsourcing authentication concerns to the community of D2D nodes. [20]

An additional security concern is introduced in D2D schemes utilizing hierarchical schemes. Since all nodes participating in a D2D cluster might not be connected to a BS, traffic can be relayed by those that are. Compromised master nodes may therefore be used to gain access to networks without the network being able to distinguish between communications emanating from the master itself and those from other nodes being relayed.

Lastly, D2D discovery schemes face scalability concerns. With the above mentioned growth in the number of nodes present in networks and with each device potentially interested in and participating in numerous clusters, device and service discovery reporting overhead can potentially grow very large. Mitigating this requires some mechanism that enables devices to limit their discovery effort to only targets of interest. However, since determining what is of interest requires at least partial knowledge of what is available, this introduces a circular dependency.

One potential approach consists in the network providing one or more centralised registries for nodes located inside a given area. Each node will report what services and nodes they are providing as well as which ones they are seeking. Any device seeking to make use of a particular service would then query the registry's database to obtain information on its availability, and in the case of a positive response its technical parameters. This co-ordination approach simplifies the operation of D2D-enabled User Equipments (UE) considerably since they are only required to perform a single query to a pre-determined entity (the serving BS).

Such a centralised scheme suffers from several drawbacks however. The complexity of support device discovery shifts from the D2D nodes to the BS. The latter must be able to categorise every service nodes register with it as being available. Creating a set method for classification of any arbitrary piece of data constitutes a daunting task as evidenced by the popularity of search engines and the difficulties encountered in the development of named data networking (NDN) [50, 66] to create a scalable, efficient and routable global naming scheme. Due to the mobility of nodes the repository of devices and services they offer would likely be constantly evolving, adding overhead to network signalling.

A second drawback of centralised schemes is that they risk creating a single point-of-failure, a bottleneck for system performance or both. Since the BS is still required to perform its cellular network management duties, adding the further task of administering a D2D node repository requires additional processing. This will come at the cost of new hardware and increased energy consumption as well as potentially increased latency [64]. This latter issue is of particular concern in light of current efforts to reduce the delays introduced by the air interface [24].

On the opposite end of the continuum, the decentralised approach delegates all D2D discovery functionality to the nodes themselves. One such method accomplishes this by allocating a broadcast time-frequency resource for D2D discovery. Nodes perform the discovery and negotiation of services amongst themselves and only contact the BS to request the required resources. The decentralised approach also suffers from several drawbacks. Firstly, being uncontrolled by the network, interference conditions are not guaranteed to be within acceptable bounds. Similarly to ISM band technologies, the number of nodes attempting to make use of the shared resources simultaneously may grow so high that none may obtain sufficient levels of service quality. These nodes may then attempt retransmissions with higher power level if they are still below any mandated maximum imposed by the BS. Secondly, due to the lack of network validation of announcements, nodes cannot trust the claimed identities of other nodes to the same degree as in the centralised scheme. In the latter, all nodes - D2D and cellular - authenticate with the BS and maintain an encrypted connection to it. If the network does not involve itself in the D2D discovery phase, this no longer holds.

4.5 Link monitoring

Once interested devices are known to the network, it must determine the feasibility of each requested pairing. Indeed, in lieu of the benefits explained above, poorly chosen D2D links can deteriorate system performance instead. Long distance links generate more interference since they employ more transmit power in order to reach the remote end and this power does not only radiate towards the intended recipient but also surrounding MSs. In terms of interference control, it is also important to consider the impact of and on neighbouring cells. Should these not be synchronized and frame-aligned, there exists a risk of transmitting simultaneously with a downlink burst thereby increasing the interference on cell edge MSs attempting to receive it. Furthermore, if D2D nodes are poorly chosen they may also suffer from poor link conditions between each other. This will lead to the need for retransmissions which in turn leads to increased latency and decreased spectral efficiency.

Any measurement system is only useful if the collected results can be communicated to where they are needed in a timely fashion. In the case of D2D systems operating inside a cellular one, there exists a need to maintain a cellular connection from the BS to at least one node of the cluster. New reporting mechanisms also need to be developed to convey channel information from the D2D cluster to the controlling BS [52]. While this type of reporting mechanism appears relatively straightforward to implement by virtue of a similar mechanism already existing within LTE, the more challenging aspect lies in determining what information to report and when.

The number of cluster members maintaining a cellular connection will depend on the intra-cluster control protocol in use. For some cases, all nodes establish both types of connections. For other types — such as the master-slave configuration — only some nodes will communicate with the BS directly. Yet, the measurements from all nodes are relevant. The manner of tackling this issue will depend on the internals of the D2D protocol while being required to fit into the control protocol used by the network. As such, a new interface must be created to connect the cellular function and D2D function of a mobile node.

The fundamental trade-off in D2D connection maintenance will be between the quality of reports and signalling overhead. Base station scheduling benefits from more accurate, complete and frequent reporting while the available capacity for the nodes participating in the D2D cluster decreases due to it. Finding the optimal balance between these two competing objectives requires information about the application domain for which the cluster was deployed. Since D2D can be implemented as a generic framework for allowing direct device-to-device communication inside a cellular system, this balance can be struck individually for each D2D cluster as they are free to select D2D control protocols independently. Clusters must however take into consideration cellular users. These will both suffer and cause co-channel as well as *adjacent channel interference* (ACI). Periodic reporting of interference level information allows the BS to predict which MS are moving closer to each other. This information becomes particularly important in attempting to avoid strong cellular MS transmissions saturating the radio front-ends of D2D receivers located nearby. Saturation occurs when interfering signals on other channels cause loss of receiver sensitivity due to intermodulation artefacts due to amplifier non-linearities [43].

Radio environment changes are difficult to predict in practice. Consequently, situations may arise where despite the best effort of the BS, D2D nodes detect channel conditions approaching inoperable levels. While for some applications being in outage until the next reporting time may be acceptable, others require even higher levels of reliability and lower latency. Use cases demanding utmost dependability require a notification mechanism to send a warning to the BS concerning the channel quality.

Node movements may also result in situations where not all participants in a D2D cluster reside under the control of the same BS. Handovers for the cellular connection may be required in order to maintain the cellular connection used for network control. This effectively imposes the further requirement to extend the X2 interface to enable BSs to exchange measurement data received from D2D nodes in order to successfully co-ordinate their operation across cell boundaries.

4.6 Control protocol

Although no D2D protocol has been deployed into cellular networks yet, some tasks that will be required to be handled by any such protocol in the future can be inferred from a survey of the literature [3]. Operation within the framework of a cellular network requires abiding by the limits set by the BS. However, much of the benefits D2D aims to provide stem from the flexibility each D2D cluster enjoys in organising itself. Achieving a suitable compromise between D2D flexibility and BS operational control requires taking into consideration a multitude of issues concerning the interoperability of the cellular standard and D2D use case in question. While such a detailed design effort lies beyond the scope of this work the remainder of this chapter will discuss general aspects of interaction between D2D and cellular components that arose and were considered during the creation of this work along with its associated software platform.

The first consideration in D2D control protocol design concerns the number of cellular-connected nodes (illustrated in 4). Having all nodes connected to the BS (see part A in Figure 4) provides the greatest reliability. Out-of-coverage events for one node will only affect the node itself instead of the node and all of its dependent nodes. A drawback for large cluster sizes is that the resource and address space consumption of D2D nodes may become large and unduly tax BS resources. Several in-between alternatives exist. Firstly, a D2D cluster can select primary and secondary master nodes, with the latter serving as a back-up in case of primary node failure. Both primary and secondary register with the BS initially to ensure only authenticated and authorized nodes may claim control of the D2D cluster in case the primary master becomes unreachable. A second solution divides nodes into hierarchical tiers (see part C in Figure 4). Each first tier node being responsible for a group of second tier ones which may in turn be responsible for a group third tier ones and so on. First tiers nodes maintain a cellular connection and act independently towards the BS while sharing the allocated D2D resources.

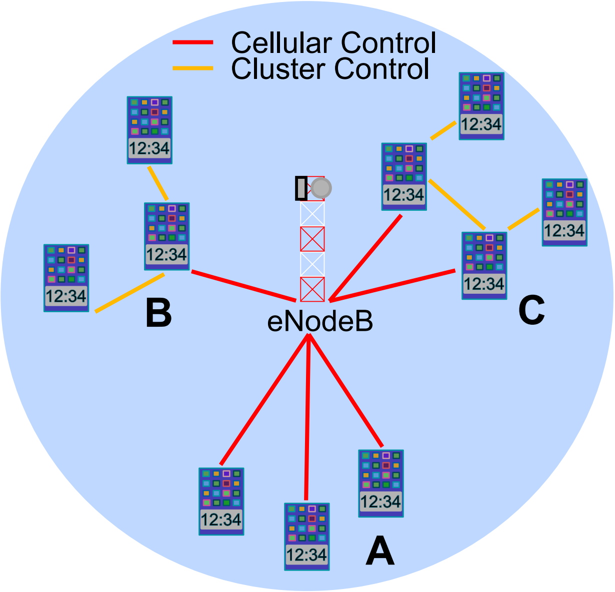

Figure 4: D2D clustering structure types. A = All connected, B = Single master, C = Hierarchical.

Network control organisation impacts the scaling of the D2D network as well. Single master systems (see part B in Figure 4) must ensure all dependent nodes are able to reach the master. This places constraints on the geographical area over which the system may operate. Systems with a single master node reporting to the BS require every other node to report to it and to receive instructions from it. Dependent nodes located farther away will require higher transmit powers to ensure proper reception. Multi-master systems are able to alleviate this issue through a selection of master nodes such that the average distance — and thus transmit power required — to nodes decreases. However, predicting interference conditions precisely is in general impossible, leading to greater uncertainty and variability in performance for single master systems due to the single point of failure. A further such vulnerability stems from hardware and software failures in the sole central master node.

Very large node counts risk saturating the computational and memory resources available on the master nodes and thus form bottlenecks for scalability of the system [75]. Every report from a dependent node needs to be processed and potentially forwarded to the BS. This process requires computations to prepare the data for transmission as well as buffering. Memory requirements increase further in the presence of retransmission mechanisms. System design involving multiple cellular connected nodes may split these memory and computational requirements amongst multiple devices.

A second important consideration in the design of D2D systems is node cost and complexity. Every application presents a different set of requirements and constraints. Adapting a single control protocol to the various demands presents a great challenge. Some applications require very simple and low cost nodes limiting the amount of complexity that may be imposed upon their design. Splitting the D2D protocol into two separate parts allows both the cellular and intra-cluster communications to be handled efficiently. The latter can be tailored to the limited capabilities of the simple nodes. All node types will communicate with each other using the intra-cluster protocol to perform their assigned tasks. Master nodes must handle the task of reporting pertinent information to the BS. Having a limited number of such cellular-enabled nodes limits the number of nodes needing to be replaced in order to support a new cellular technology while allow the retention of the simpler dependent nodes. Conversely, dependent nodes may be updated without needing any modifications to the cellular network facing configuration.

In light of the above factors, a hierarchical D2D control scheme presents the additional advantage of allowing dependent nodes of various capabilities to be handled within a single cluster. Master nodes handle communication both towards the rest of the cluster and to the BS. In the event of a loss of communication with one master node another one steps in to replace it. If no new suitable master node can be found, the remaining master nodes will at least be able to notify the BS of this fact. Additionally, re-establishment of communication may be attempted using cellular network relaying if the disconnected D2D nodes support it.

Thirdly, control protocols overhead requires careful consideration. On the one hand, frequent reporting of *channel quality indicator* (CQI) information determines the accuracy of channel estimation and therefore system performance. On the other hand, frequent reporting reduces the available payload capacity. Amongst the factors determining the optimal balance, one of the most important ones is speed of the nodes. Higher velocity results in more frequent changes to channel conditions and thereby a heavier reporting burden. D2D control protocols must therefore be able to respond rapidly to changing channel conditions in order to maintain reliable communications [46]. Single master system risk saturating the capacity of this single node while approaches with all nodes network-connected will utilise a large amount of cellular resources. Hierarchical systems mitigate these issues by having multiple master nodes. These split the work amongst themselves while also providing the opportunity to filter out or aggregate CQI reports towards the BS. Only reports relevant to the network are forwarded out of the dependent group.

Finally, in order to support high-reliability links, a control protocol must support multiple disjoint resource allocations obtained from the network. Since interference conditions are in general impossible to fully predict ahead of time, no amount of due diligence on the behalf of the BS can ensure that a given frequency block remains usable without interruption. In these cases a back-up resource is desirable and in the case of latency-critical systems, unavoidable. Disjoint frequency blocks add to the CQI measurement burden due to the need to monitor different parts of the spectrum. The control protocol must thus possess the ability to instruct controlled nodes to perform channel quality estimation and reporting in a manner that does not unduly interfere with ongoing communications.

4.7 D2D scheme implemented for measurements

The D2D scheme selected for implementation in the measurement code is a network-controlled scheme with all nodes possessing both a D2D and cellular connection. Resource allocations are requested from the BS. As such the BS scheduler must be modified to accommodate the new request and reservation type. Scheduling inside the allocated resource is organised by the D2D nodes themselves. This means determining when each device performs transmission and reception does not the involve the network. D2D nodes will make such a request after determining they are able to establish direct communication at the desired level of reliability.

Communication inside the D2D cluster utilises the same waveform, modulation and coding scheme as for LTE uplink. Consequently, the D2D receiver component strongly resembles the BS's receiver. In order to distinguish uplink transmissions from UEs from those made by D2D transmitters, a different pilot sequence is used. Subframes not used for D2D communication are processed as they would be when operating in cellular mode.

5 ARF platform

Measurements (see section 6) were performed on the Aalto Radio Framework (ARF) platform which is a C++ *software-defined radio* (SDR) partial implementation of the TDD LTE Release 8 specification. The code comprises the *physical* (PHY) layer (TS-36.211), the *medium access control* (MAC) layer (TS-36.212) and the *radio link control* (RLC) layer (TS-36.213) as well as supporting interfaces to input and output data to and from the LTE stack. For IP data this is done through a virtual network device called TUN interfaces in the Linux kernel. All of the baseband processing occurs on a general purpose PC running a stock Linux kernel based operating system. The various components of the system have been distributed across a number of threads in order to improve system performance.

The aim of the ARF platform is to provide a flexible and easy-to-use research tool for the study of radio technologies. Due to its all-software nature, modifications and extensions to the code are possible at all radio protocol stack levels. This combined with the modular nature of the architecture enables insertion of new functionality in one area while retaining the previous functionality elsewhere in contrast to traditional hardware based implementations. These use *application specific integrated circuits* that are difficult and costly to modify. Furthermore, since hardware-based solution are commercial products, access to their designs is usually not possible outside of the company producing them. Time-division duplexing needs only one frequency band to be monitored as opposed to two for frequency-division duplexing approaches. The ARF platform was chosen for the above mentioned reasons. Achieving the above mentioned benefits requires compromises in other areas. Firstly, the lack of purpose-built hardware typically results in lower achievable data transfer speeds. Secondly the ability of the front-end to support a wide range of frequencies comes at the cost of smaller usable bandwidths. Finally, the ARF platform does not target operational deployments and is as such not designed for commercial levels of reliability. It does however aim to provide sufficient reliability to obtain valid scientific measurement results.

D2D features needed for this work were thus added into the existing cellular LTE framework without needing to write a new tool altogether. Such an approach furthermore provides the benefit of enabling the testing of co-existence as well as network-control of D2D nodes. An additional advantage of the flexibility comes from being able to iterate the design quickly without additional hardware related costs or time spent on the verification of real-time system timing requirements compliance.

The ARF platform provides in addition to the above mentioned flexibility, sufficient performance to enable two-way time-division LTE communication. In turn, this enables research into low-latency operations which constitutes one of the main goals of D2D. The following sections will begin by describing the ARF platform's internal structure relevant to this work followed by a details concerning the main components. Lastly, sections 5.5 and 5.6 discuss the methodology used for adding D2D functionality will be discussed along with the issues impacting the SDR platform's design.

5.1 SDR framework

Software-defined radios are communication systems implemented using software. Unlike traditional radio systems, they utilise *general purpose processors* (GPP) to perform the necessary computations and operations to achieve the transfer of information. Typically, the only other piece of hardware involved in addition to the computer is known as a *radio front-end* (RF–front-end) performs the analogue-to-digital (ADC) and digital-to-analogue (DAC) conversions, modulation to and from the carrier frequency and the transmission and reception of radio frequency (RF) signals. SDR platform may further be classified into *hard real-time* and *soft real-time* systems. The former have bounded execution time on all functions while the later do not. Verification of timing requirement compliance during the design phase enables hard real-time systems to offer greater reliability. This however comes at the cost of rigidity as all changes must be carefully considered so as not to disturb the timing of the execution flow. As such, any feature that may potentially exceed its allotted computation time may not be implemented at all even when such lengthy execution times are rare. Soft real-time systems on the other hand only need to be fast enough on average to be able to perform their duties. The cost for this flexibility comes from the increased variance in performance resulting from execution time being stochastic. Subframes may be late simply due to functions being called in a suboptimal order due to one function taking a longer-than-usual time to complete.

The ARF platform implements an extended version of the soft real-time SDR concept — an SDR framework. SDR frameworks provide a supporting infrastructure for the development of various radio technologies without the need to re-write every part of the necessary code. Functionality typically found in every piece of software and every radio protocol exist within the framework and are shared by the various RF technologies implemented in a manner akin to the use of libraries in software development. The fundamental task of the platform is to provide a radio research tool providing rapid testing of concepts while minimising the amount of code needing to be written for each experiment. Additionally, the platform solves performance, timing and other computer engineering related problems hat are not germane to radio technology research. Having been solved once, they need not be re-implemented for each new experiment.

Architecturally (see Figure 5) [34], the platform is comprised of the upper layers and a *Virtual Hardware Enhancement Layer* (VHEL) (see section 5.3). This design enables the protocol-specific parts to be written focusing only on the implementation of the radio protocol as described above without need to consider the non-idealities related to the use of a general purpose, *non–real-time* OS. Such a subdivision provides isolation of concerns. Each subsystem need not know about the complexities of others. Communication between the different parts of the processing path occurs through

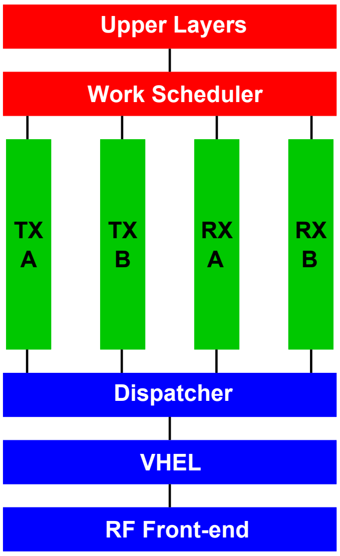

Figure 5: ARF platform architecture.

the use of timestamped buffers. The unit of time used is the subframe number in the same way as in many cellular technologies, thus providing a natural model for the implementation of said technologies. A multiplexer will use this timing metadata to assemble subframes correctly even if the generating components vary in their execution speed. Moreover, the multiplexer keeps transmitted and received subframes in order within their own queues as well as relative to the configured TX/RX pattern. Individual functions need therefore not concern themselves with the order of execution of other function and instead simply fulfil their own task for the subframe indicated in the metadata. This allows pre-generation of transmission bursts as the VHEL will ensure that samples are sent to the radio front-end at the correct time regardless of when and in what order they were generated by the upper layers. The VHEL is responsible for taking the baseband signals generated as above and sending them to the RF–front-end as well as receives data from it and passing it to the baseband logic. In order to do so, the VHEL must convert between the logical subframe level timing used in the radio protocol logic to the sample level timing used by the front-end.

Higher layer radio protocol implementations operate on two continuous streams of subframes provided by the VHEL — one for TX and one for RX. Each of these is operated on by a separate processing chain to allow for full-duplex operation. Both the transmission and reception chains are subdivided into separate modules executing the specified functionalities of the PHY, MAC and RLC layers. Each of these can be easily modified allowing for experimentation on any part of the protocol stack. Multiple different protocols or instances of the same protocol may be executing concurrently within one instance of the platform. This enables the use of features such as sectorised base stations, relays and multi-protocol access points.

In parallel to the radio protocol specific higher layers there are platform higher-level functions. These enable the control of the platform in various ways. At the core of the platform resides the *Operations and Control Interface* (OCI). This module receives configuration and control information from multiple sources including configuration files, user input and a *centralized cloud radio access network* (C-RAN) controller. All such information merges within the OCI which in turn further distributes it to the relevant parts of the code executing radio protocol functionality. In addition, the OCI possesses the ability to pause and resume execution of the code. Such functionality aides during development since a node may be suspended automatically when certain events occur.

Finally, the framework also provides generic protocol-agnostic functionalities intended to support development and operations. This consistency of measurement and analysis functionality across various radio protocols improves the platform's ability to provide comparable quantitative data useful for comparing one protocol or implementation of it to another. The support facilities also provide statistics that are useful in implementation of software-defined radio systems, such as processing time per subframe for example.

5.2 Pipes

Pipe is the name given to the basic worker unit in the ARF platform architecture. They belong to the radio protocol specific portion of the architecture as described in the previous subsection. Pipes are responsible for generating or decoding one subframe according to the instruction generated by the upper layers. Consequently, they operate on the smallest scheduled time unit in the radio protocol being implemented. Such a scheme simplifies their design by removing the need to track state information and furthermore helps in saving power as only the upper layer module needs to wake up every subframe in order to determine what, if any, further processing is required. This enables the implementation of various *discontinuous reception* (DRX) schemes. Furthermore, if a node mostly receives, there is no need to keep the transmitter pipes active — contributing to further power consumption reductions.

In additional to being the atomic unit of scheduling for the radio protocol being implemented, pipes also function as the atomic unit of operating system (OS) scheduling since they are implemented as threads. Threads were chosen for several reasons. Firstly, they provide improved performance by exploiting the processing core level parallelism present in modern CPUs. Moreover, they allowed computationally intensive functions to be distributed more evenly. Secondly, threads — unlike processes — share memory, thus enabling fast exchange of information. Lastly, by grouping the processing functions for a network node into a single process, a two tier prioritisation may be used. Some processes (nodes) may be given more or less resources through the operating system while still retaining the ability of each process to distribute its allotted resources as it sees fit amongst its own threads. Creation of threads occurs at initialisation time in order to avoid the costly [71] operation of spawning new ones during execution. Consequently, threads may be assigned to specific *central processing unit* (CPU) cores to distribute load manually or this may be left to the OS's built-in scheduler. The performance of the latter will vary depending on the load and activity pattern of each thread due to its dual nature as both OS and radio protocol work scheduling unit. The latter stems from the need to assign work to the stateless pipes (and therefore the thread they execute on) in the correct order to ensure completion of subframes in the desired sequence. Typical Linux schedulers allocate execution time using timer with resolution on the order of tens of milliseconds [23] whereas LTE for example requires millisecond precision. On a fully loaded system this can result in milliseconds of skew in the regularity of process scheduling. Any delays beyond one millisecond will accumulate over time leading to subframes being dropped as late. This can happen when processing exceeds its allotted time or the thread starts work late due to inefficient scheduling. Scheduling precision on the order of the subframe duration or less is therefore desirable.

The passing of higher layer control information to pipes occurs through timestamped control information buffers termed *Work Items* (WI). Each WI gets assigned to one pipe that will execute the tasks specified therein. When processing completes, the result will be passed on further: to the VHEL in the TX case and to the upper layers in the RX case. In the latter, the WI's timestamp also serves to associate it with the correct sample data buffer sent up from the VHEL. This is important since one pipe can have multiple WIs queued up for processing. Each one will be fully processed before any system clock or other parameter values are checked for changes. This design cleanly separates the responsibility of creating a subframe from the determination of the feasibility of it reaching the next step of its processing chain in time.

Once processing starts on a subframe its environment (protocol state machine and configuration parameters) remains constant. WIs do not change during the processing. Such a design is necessary to avoid the need to lock global resources from within a pipe. Were this do not done, two or pipes could attempt to modify the same variable simultaneously leading to corruption of the data. Frequent locking and synchronisation would impose an unbearable source of latency [48] on operations as well as risking deadlock between pipes. The stateless and isolated nature of pipe execution environments enables their sharing amongst multiple instances of the same protocol. Doing so reduces the number of threads and CPU cores needed to run node instances on a single computer and amortises the overhead of the platform support thread compared to running multiple processes. Pipes all inherit from a common base class to facilitate their scheduling by the *Work Scheduler* which matches radio protocol WIs to available threads implementing the correct type of pipe.

Since pipes are protocol specific, tracking the execution time of each provides some indication of the per-protocol load situation. Such statistics aid in load-balancing and resource allocation between different protocols. However, it should be noted that such load-balancing applies only to the radio side as inactive threads do not take up CPU resources resulting in no penalty from having large numbers of different protocol pipes ready. As pipes form the main user of CPU time, optimising their usage and behaviour are expected to provide the greatest gains in terms of computational resource consumption.

5.3 VHEL

Since the upper layers work on whole subframes but the radio front-end operates on sample streams, the VHEL must convert between the two. In order to accomplish this, a pre-determined amount of samples are collected and transformed into an RX subframe or vice-versa for transmission. The VHEL's main loop operates on a RX1-TX-RX2 pattern wherein RX1 and RX2 are for the first and second slots of the subframe respectively (see section 2.3 for LTE air interface details). Subframe transmission and reception must follow the protocol's timing constraints precisely in order to maintain the frame structure for all nodes. Interleaving the TX sample sending provides half-a-subframe's worth of extra time to ready the outgoing burst while not risking the subframe being late at the front-end as would be the case by transmitting at the end of the timing window. By maintaining separate TX and RX queues the illusion of steady parallel flows of complete subframes is presented to upper layers.

The VHEL is additionally responsible for maintaining sample level synchronization based on input from the correlator used to detect and follow the protocol's frame structure [34]. In the case of LTE, the correlator tracks the primary synchronization signal (PSS) and the secondary synchronization signal (SSS) (see section 2.3) and send the computed amount of drift with regards to the nominal position to the VHEL. Drift values indicate how many ADC samples earlier or later than expected the correlation peak was found. Constant monitoring and correction of the sample drift enables errors to be kept in check over long periods of time by ensuring the right number of samples are read for each subframe. The exact number depends on the protocol in use and its configuration. Once the correct sample count has been extracted, the VHEL ticks the logical clocks used in higher layer processing forward by one subframe. The whole stack's timing therefore derives from the constant inbound sample stream from the radio front-end.

As the RX sample flow provides the clocking for the system, the TX part of the VHEL main loop (RX1-TX-RX2) cannot wait indefinitely for baseband processing to complete for the current subframe since doing so would skew this timing. Sample data must be present at the *remote radio head* (RRH) in time for its transmission time vis-à-vis the protocol's frame structure. The VHEL should however not pass through the transmission step too fast either as nothing is gained by blocking on the receive call for RX2 longer (i.e., waiting for all samples to arrive). Since a non–real-time OS exhibits varying delays on operations determination of the optimal waiting time cannot be done statically. An algorithm is instead used to estimate the appropriate delay utilizing feedback from the RRH. Reports of sent samples arriving late mean the waiting time was too long. A requirement for multiple consecutive measurement periods to be over an adjustable late subframe count threshold before action is taken accounts for the bursty nature of the losses. Conversely, if no or few subframes are considered to be ready in time, the waiting time is likely too short. When subframes arrive late from the upper layers the VHEL sends zeros instead and adjusts timing related metadata as if the burst were on time in order to maintaining timing alignment. Upper layers handle these lateness induced losses automatically as similar losses occur in radio channels due to fading. LTE therefore has an *automatic repeat request* mechanism for handling such cases which will also recover from subframes being late due to baseband processing exceeding its allotted processing time. Whether the loss was due to unfavourable channel conditions or the transmitter's inability to produce the required data in a timely manner is indistinguishable to the LTE ARQ functions. Re-transmissions will therefore operate without modification or additional input, freeing the VHEL from the need to maintain protocol integrity.

Separation of radio protocol and radio front-end management in VHELs enables scaling of the SDR platform to multiple computers. It is possible to separate the generation of subframes from their conversion into samples. One computer handles the radio protocol while another manages radio front-ends. Such separations enable multiple different use-cases. Firstly, it becomes possible to allocate additional computers to radio protocol (especially physical layer) processing since this task constitutes the majority of the computational load of the system. A single VHEL server may thus aggregate the protocol data of several compute machines. Secondly, computing and RF–front-end equipment may be shared by multiple entities with the VHEL server acting as arbiter. A third use case is to increase the reliability of the system by reducing the number of tasks performed by each machine.

5.4 Upper layers

Upper layers make decisions with regards to what is to be done for a particular subframe. These decisions are then executed by the pipes. Often, such MAC and higher layer protocols rely on state machines describing their operation or some part thereof. The inputs to this decision-making comes from received messages and in some protocols, a dedicated control element (for example, the MME in LTE/EPC) residing elsewhere in the network. Upper layers additionally carry the responsibility to determine the parameters one can use for transmitting such as frequencies to use, power levels, timings in TDD and so on. Removing the burden of tracking such parameters from the pipes provides two benefits. Firstly, it allows pipes to be streamlined into reusable physical layer processing blocks. Secondly, implementation of control logic in the upper layers becomes simplified since all the information is centralized and detangled from its implementing logic.

A further benefit of this split design is the ability to implement both the physical layers processing and MAC layers according to different programming approaches. Both can utilise the optimal approach for its implementation without imposing any restrictions on the other. The sole defined part of the architecture is the manner in which data passes from one to the other and how synchronization as well as timing may be maintained. In the case of the ARF platform, the co-ordination function takes the form of the Work Scheduler. It receives WIs from the upper layers and assigns them to a suitable pipe for processing and ultimately transmission. In the receive direction, the processes is reversed. Once a pipe finishes decoding a received burst, it notifies the Work Scheduler which will in turn pass the decoded information onwards to the upper layers. The latter is implemented as a hybrid of a distributed state machine and agent oriented programming paradigm [53]. This approach was chosen as it corresponds well with the LTE standard. Each functionality (MAC, RLC, etc.) corresponds to one C++ class implementing the state transitions according to signalling messages as defined in the standard. Each of these blocks also acts as an agents to enable inter-layer communication.

Radio resource management (RRM) functionality is also implemented in the manner described above. However, due to the centralised nature of some control elements in LTE — and cellular networks in general — the ARF platform also supports exchanging messages to entities outside of the local base station. These messages are exchanged with the application specific logic of the program. After processing by the latter, messages are sent forward as IP packets through a TUN interface. X2 and data plane traffic follow the same processing path to enable the application layer to perform classification of traffic. There are several reasons for this design choice. Firstly, separating the classification of incoming IP packets from the core radio related code enables both to be developed separately. Experimentation with different core network designs and protocols become possible without requiring any changes to the single base station related parts. One notable use case for such work would be quantification of *software-defined networking* (SDN) approaches. A second benefit of the split design stems from the isolation of timing sensitive functions — such as MAC layer scheduling — from the less time critical processing of incoming and outgoing IP packets. Reducing the fraction of the code base required to meet strict timing constraints simplifies programming while reducing the potential number of delay sources. Thirdly, by not tying the data and control input of the platform to IP, it becomes possible to implement a consolidated C-RAN architecture without the need to traverse a network protocol stack. In these, the data may come from another base station code instance running on the same physical server. Use of any network protocols for data transfer constitutes unnecessary overhead in these cases since the information already exists in memory.

5.5 D2D extension

Extension of the ARF platform to support the selected D2D communication scheme (see 4.7) required changes in two areas. Firstly, the MAC level scheduler needed to be expanded to support scheduling reception of subframes during uplink TTI for the D2D nodes since these were based on the code for UEs. The latter do not listen for incoming bursts during uplink. Additionally, the D2D version of the scheduler follows its own internal logic for determining which U subframes (see section 2.3 for LTE air interface details) to use for D2D transmission and which ones for reception.

This decision is made independently of the eNodeB in so far as it does not violate any constraints (e.g., allocated resource blocks) given by the eNodeB and with the proviso that synchronisation to the cellular network must exist for any transmission to take place. This last condition helps to ensure that no interference will be caused due to connectivity loss induced timing offsets.

Secondly, the UE pipes were augmented with D2D transmission and reception code paths. Transmitter functionality stems from the standard UE components while the receiver's implementation originates in the eNodeB receiver chain. Both were modified to take their allocated resource parameters from the D2D configuration instead of the normal grant data structures determined by the base station. Since the D2D code path was added to the UE one, as opposed to replacing it, switching from one behaviour to the other requires only a change in a single configuration parameter. This, in turn enables testing of the behaviour of the system with UE or D2D nodes in identical conditions without the need to swap equipment.

5.6 Challenges

Software-define radio (SDR) systems face a number of additional challenges compared to their traditional purpose-built counterparts [34]. These challenges grow even more pronounced when the underlying operating system is non–real-time. SDRs must manage the variability introduced by the use of non-dedicated hardware potentially executing other tasks in parallel, thus creating a much more variable execution environment than traditional systems. Even if the typical time to process the required workload is fast enough, the tail end of the latency distribution may have have orders of magnitude greater [39]. Furthermore, the generic nature of the computation infrastructure limits the opportunities for hardware assisted data movement and synchronization. This section details the issues presented by non–real-time GPP environments as well as the techniques used to tackle them in the ARF platform.

Non–real-time general purpose OSs aim to support a wide variety of applications and in doing so offer a wide variety of services and features. All of these require system resources to execute. The precise timing of the activation of these features and services is generally not fixed but depends on external events such as user input or messages received over some network. Such unpredictability clashes with the millisecond level timing and periodicity of LTE air interface frame structure. Whenever another process's thread wakes up, it may be allocated by the OS to a CPU core when an ARF thread needs to perform some time critical task. An inability to impart to the thread scheduler any ordering of threads compounds the problem further. Resultant timing slips do not necessarily cause error on the air interface depending on what functionality was delayed and the length of the interruption. Since activity in other processes impacts the timing of threads in the ARF platform, reliability enhancements can only aim to minimise, but not completely eliminate, the risk of sequencing failures. This issue is exacerbated at higher system loads.

The ARF platform supports the assignment of higher priority levels to its various threads in an effort to minimise the number of interrupts they suffer from other processes as well as to ensure that the relative order of the threads be conducive to error free operation. The highest priority is afforded to the VHEL thread since the entire platform derives its system clock from the inbound sample flow originating from a radio front-end designated as being the master timing source. Timing information obtained from this system clock allows other modules (i.e., threads) to detect whether they have been suspended for too long by comparing this time with the timestamps in buffers and WIs. One significant cause of missed timings are operations across the boundary between the millisecond level radio part of the code and external sources. The exchange mechanism providing the bridge between the asynchronous side communicating in an event-based fashion through a TUN interface and the synchronous side of the air interface must not block the air interface. Consequently, the asynchronous part executes with the least priority.

CPU frequency scaling affects the time required to perform a given set of operations. Since a general purpose Linux operating system typically schedules processes for tens of milliseconds [23], performance and power level tracking and adjustments apply a similar granularity for software based powers state management. Averaging over tens of milliseconds results in several LTE frames being amalgamated. This in return means an unrepresentative average appears since there will be subframes with no or little processing due to scheduling and uplink-downlink periodicity mixed with potentially very processing heavy subframes. The latter might need full CPU power to be available yet the OS sees only the coarse average and determines that CPU frequency may be scaled back. Furthermore, this scaling impacts the time required to complete processing steps. Delays added in such fashion may cause deadlines to be missed. Compensating for such variance is a difficult task as different hardware and software combinations display different behaviour in terms of the above mentioned scheduling and power management behaviour.

Network subsystem performance and configuration forms another critical factor for the performance of the SDR platform. Longer buffering times add latency without providing much benefit since any data held in these buffer becomes irrelevant if not transmitted promptly. One possible cause of extra delays is the interrupt moderation mechanism present in many OSs and network card drivers. It aims to reduce the number of CPU interrupts caused by the arrival of network packets [54]. When the limit is reached, any new packet that arrives will not be handed over immediately by the network card but delayed to grouped together into a single interrupt instead. This can be problematic in software radio systems as a single transmission cannot be properly processed without all of its samples being present. If the packets containing the end of a burst are held up due to interrupt moderation, the system experiences extra latency.

Development and operation of the platform faces difficulties from the above mentioned issues as well. Executing the code under the control of debuggers, profilers or other runtime analysers slows processing down. Several undesirable phenomena may appear as a result. In some cases, the tool will detect errors in performance critical section and begin more thorough investigation and logging of them causing an unacceptable delay resulting in loss of functionality. Another possible outcome of the use of runtime tools is the apparent disappearance of errors — typically race conditions. As the tool slows down code unevenly, it is possible for threads to execute in the desired order despite there being a programming error present. The same situation applies to any internal logging feature of the platform. These need to be designed in such a way that one thread does not block another. For example, logging functionality must allow several threads to use it simultaneously while still preserving the integrity of the data. Care must be taken when logging timestamps along with data. Since the platform operates on the level of subframes but processing steps takes some small fraction of a millisecond, cumulating delays may not be apparent when using only subframe numbers. The log might show that subframes N, N + 1, N + 2 and so on were on time when in fact each of time took some additional time over one millisecond. Cumulatively, these delays cause missed deadlines some time in the future.

6 Measurements

Enabling device-to-device (D2D) communication requires not only a careful design to enable its internal operation but also consideration of its impact on the technology it will be integrated in. As discussed in section 4.6, a D2D control protocol must ensure compliance with interference limits imposed by the cellular technology within which it operates while simultaneously attempting to optimise the available communication capacity, latency or other target performance metric.

This chapter presents measurement carried out to determine suitable values to achieve the above mentioned goals as well as briefly describes a proof-of-concept demonstration carried out to show the feasibility of D2D communication with the ARF platform. The first measurement carried out aimed to assess the amount of adjacent channel interference experienced by the D2D link in the presence of cellular traffic in order to determine feasibility of integration. The second test aims to assess the required reporting frequency of D2D nodes in order to maintain communications at various speeds.

6.1 Proof of concept

A demonstrator of the D2D capabilities of the platform was created first. It consisted of two links. The first one is a traditional LTE Release 8 like eNodeB-to-UE link over which a video stream is transmitted on downlink using constant bit-rate encoding of around 100kB/s MPEG2. The second link is a D2D link transmitting *bit error rate* (BER) measurement data. This second link can operate in either a synchronized or unsynchronized mode. In the latter it interferes with the cellular link, result in a visible degradation of the video quality. This demonstrator was showcased at the EuCNC'2014 conference.

The equipment used was four HP EliteBook 8570w laptops. These possess an Intel Core i7-3820QM CPU and 16GB of RAM. National Instrument NI2932 (see section 6.2.1) devices served as the radio front-ends with a carrier frequency of 2.48GHz. The test-bed platform code ran on an unmodified Ubuntu 12.04 operating system with a stock series 3 Linux kernel for the cellular UE and D2D nodes and Ubuntu 14.04 for the eNodeB. C++ code was compiled with the GCC g++ compiler version series 4.6 on Ubuntu 12.04 computers and with version series 4.8 on 14.04.

This proof-of-concept demonstrated the need for some form of co-ordination mechanism in order to enable D2D and cellular nodes to co-exist. When switched on but without synchronization, the D2D link experiences significantly lower throughput (on the order of 10kB/s) than when synchronized (approximately 100kB/s). This degradation was also visible subjectively through the failure of the video player to maintain video quality due to the reduced number of correctly received video data carrying IP-packets.

6.2 Equipment used

All measurements were realised using commercial off-the-shelf equipment without any custom modifications. These were selected due to their availability rather than any particular requirement. Any similar piece of equipment using the same standards and software interfaces could have been used. The same computers and radio front-ends were used for both measurements.

6.2.1 USRPs

Conversion of digital samples to and from analogue radio-frequency signals was performed by National Instruments USRP-2932 [30] front-ends using UHD driver version 003.008.000. The latter handles the communication between the computer hosting the baseband processing and the USRP unit. As such, user code never interacts directly with the hardware. Certain parameters values are requested from the UHD and it then in turn direct the front-end to take appropriate action. For the measurements presented in this chapter, the most important parameter is transmit power. The driver accepts a value between 0 and 30 for it which is then mapped to an actual power level up to the maximum of 20dBm. Indicated power levels in the remainder of this work refer to this power setting as opposed to the actual power level.

Another important factor to consider for the performance of the USRP-2932s [30] is the accuracy of their timing sources. These units possess a temperature-compensated crystal oscillator which requires some time to warm up and achieve best timing stability. Each front-end was run for at least half-an-hour before measurements started. Furthermore, each unit was connected to a GPS antenna to enable the in-built GPS receiver to synchronise to common reference and thus further increase timing accuracy.

6.2.2 Cable and connector loss measurement

The measurements in 6.4 require knowledge of the attenuation between transmitter and receiver in order to enable the computation of the equivalent distance. It is therefore important to obtain as accurate as possible values for the losses in the system in use. Cables and attenuators each contribute some amount of attenuation. These values are typically provided by the manufacturer and can be used in determining the losses of a system. For the purposes of the measurements conducted in this work, only the total attenuation is relevant. Furthermore, since the aim of the experiments were to obtain data for realistic distance, constant added losses in the results do not present a problem. Were they not present, more attenuators would have been used instead. Moreover, measuring — as opposed to calculating — takes into account wear of the components as well as any other effect that cannot be determined solely based on manufacturer data sheets.

In order to determining the value of the cable and connector losses, the USRP front-ends were detached from the measurement setup. In their stead, for each both

D2D TX — D2D RX and D2D TX — UE, a Rohde & Schwarz SMBV100A [58] signal generator and a Tektronix RSA 6114A [63] spectrum analyser were connected. The former was set to transmit a pure sine wave at the carrier frequency of the test LTE cell (2.48GHz) with a transmit power of 0dBm. Received power was then measured by the spectrum analyser and used to determine the total losses. For the both the D2D TX and UE to attenuation towards the D2D RX was approximately 16.5dB. In addition to attenuation measurements, the channel's frequency response was probed by setting the signal generator to transmit QPSK modulated data and the spectrum analyser to display constellation points for the received symbols. Using visual inspection, it was determined the channel does not exhibit any significant distortion.

6.2.3 Computers

The computers used to run the ARF platform during the adjacent channel interference and co-channel interference measurements were Fujitsu Celcius desktops. Tables 2 and 3 present the main components from a performance viewpoint. Additionally, all computers were linked to their respective USRP front-end with a direct Gigabit Ethernet connection.

Table 2: CPUs and RAM used in the measurement computers for adjacent channel interference measurements.

	eNodeB	UE	D2D RX
CPU	Intel E3-1230	Intel E3-1235	Intel E3-1230v3
RAM	16GB	8GB	16GB
OS	Ubuntu 14.04	Ubuntu 14.04	Ubuntu 14.04
Kernel	3.13.0	3.13.0	3.13.0

Table 3: CPUs and RAM used in the measurement computers for co-channel interference measurements.

	eNodeB	UE	D2D RX	D2D TX
CPU	Intel E3-1230	Intel E3-1230v3	Intel E3-1235	Intel i7-3820QM
RAM	16GB	16GB	8GB	16GB
OS	Ubuntu 14.04	Ubuntu 14.04	Ubuntu 14.04	Ubuntu 14.04
Kernel	3.13.0	3.13.0	3.13.0	3.13.0

6.3 Adjacent channel interference

Achieving ultra-reliable D2D communication inside a cellular network presents new challenges not faced by previous generations of cellular technology. The proof-of-concept demonstrator described in section 6.1 clearly demonstrates the need for network control. As discussed in sections 3.4.2 and 3.4.3 the impact of nearby cellular mobile stations (MS) on D2D links has the potential to impair communications for both node types. For network control to be effective, sufficient information must be available to the base station (BS) to make appropriate resource allocation decisions. Adjacent channel interference measurements provide useful data for minimising conflicts between cellular and D2D users.

6.3.1 Measurement setup and procedure

Measurements were performed with three nodes configured as depicted in Figure 6. All radio front-ends were connected to each other using cables. Doing so minimised the impact of outside interference sources such as the campus wireless network operating on the same frequency band. The eNodeB provided the synchronisation source for the two other nodes. Results were recorded at the D2D receiver (D2D RX in the figure) based on the detected interference power emanating from the UE node. The computers executing the ARF platform code had no other user processes running on them while system ones were left in their default configuration for desktop use.

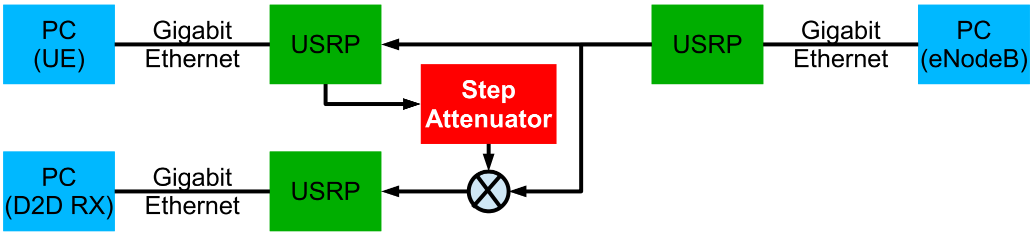

Figure 6: Measurement setup.

SC-OFDMA was used as the waveform for the UE. No D2D transmitter was present. The D2D RX node solely measured the impact of adjacent channel interference from the cellular user. The D2D cell was configured to operate on a carrier frequency of 2.48GHz and occupy a 5MHz (25 resource blocks) bandwidth. Of these the four lowest in frequency were assigned to the D2D link followed by four RBs assigned to the UE on the next highest ones.

The step attenuator used was a RSC model from Rohde & Schwarz [57]. Its purpose was to emulate the effect of distance between the UE and D2D RX. In addition to this variable attenuation 30dB fixed passive attenuators were used to prevent equipment damage when the step attenuator was set to 0dB attenuation.

Table 4 lists the attenuation settings used in the measurements proper for the step attenuator. Used transmission power settings can be found in Table 5. Additionally, a setting of 100dB was used to record the noise floor. The selected attenuation values cover distances from very near to moderately far using Equation (2) in section 7.2. These attenuation values are relevant to the expected short range nature of D2D communications as found in the literature and discussed in Chapter 4.

Table 4: Step attenuator settings used for the measurements.

Attenuation (dB)	0	5	10	20	30	40

Table 5: Cellular UE transmit power settings.

Transmission power (dBm)	5	10	15	20	25	30

Measurements were performed by recording each combination of step attenuator and cellular UE transmit power. The D2D receiver recorded both the in-band transmission power of the cellular UE as well as the power levels in its own allocation. Calibration of the results was performed by recording the interference levels seen in the D2D allocations with the step attenuator set to 100dB. Each measurement was manually timed to be approximately 65 seconds long.

The method described above presents the advantage of recording the noise rise due to the proximity of cellular users to D2D ones. Such results are more easily transferable than absolute values. The use of uplink resource was selected to minimise any issues due to degradation of synchronisation signals causing sample timing errors. In order to further improve the stability of transmitter and receiver, *global positioning satellite* (GPS) calibration of the internal oscillators was performed to reduce the difference between the timing sources of each front-end.

6.3.2 Traffic characterisation and channel coding

The cellular UE transmitted random payload byte values that were channel coded according to the Long-term Evolution (LTE) standard using Turbo Coding with a *modulation, coding scheme and redundancy value* (MCSRV) of 9. Subframes were transmitted in every uplink subframe of a time-division LTE (TD-LTE) frame structure with uplink-downlink configuration scheme 1. The only other transmissions were those of the evolved NodeB (eNodeB) during downlink subframes.

Figure 7 depicts the frame structure used. The UE sent data on uplink subframes marked "U" in the table. Each subframe was filled to capacity with random bytes as

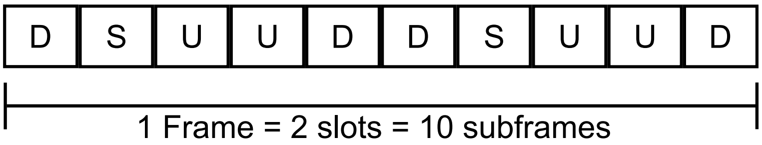

Figure 7: TD-LTE frame structure 1. D = downlink, S = special, U = uplink.

mentioned above. Transmissions and recording of data points was started only after synchronisation with the base station had been achieved.

6.4 Co-channel interference

Co-channel interference measurements aim to establish the impact on the D2D links performance when a single time-frequency sharing cellular UE move closer relative to it. Such a scenario occurs when D2D nodes enable an increase in spectral efficiency as described in section 4.1. While such an arrangement provides the possibility of efficiency gains, it also presents risks of the opposite occurring should interference not be properly controlled. These measurements provide an assessment of the magnitude of the impact of such co-existence in the case of a single static D2D pair being interfered by a single cellular UE.

Measurements were performed in 0.1dB steps around an attenuation value of interest as determined by testing prior to recording data. Determination of this value range was done by observing the success rate of decoding and determining between which values a change occurs from virtually no successful decodings to practically all successful. Attenuation levels were used as a proxy for distance in order to enable the use cabling instead of antennas to prevent non-controlled source of interference from influencing the results. Measurement over cables was therefore selected to provide a more consistent controlled environment. Several transmit power levels of the cellular UE were used during performance measurement. This data was later converted during analysis to represent distance. Higher TXP corresponds to being place farther away from the BS. The distance (i.e., the attenuation) between D2D transmitter and receiver was held constant.

6.4.1 Measurement setup and procedure

Three computers were used to run the ARF platform for the measurements with the eNodeB and UE sharing one machine. This arrangement provided the most even balance of loads since the base station only sends pre-determined pilots and synchronisation without performing any significant computations and as such has a low enough computational load to enable it to share computing resource with the UE.

Transmission from the D2D transmitter and cellular UE shared the five lowest frequency RBs of a 5MHz (25 RB) time-division LTE frames structure 1 cell (see Figure 7). Figure 8 depicts the attenuations used in the measurement setup. Both transmitters were connected to the D2D receiver through cables and attenuators

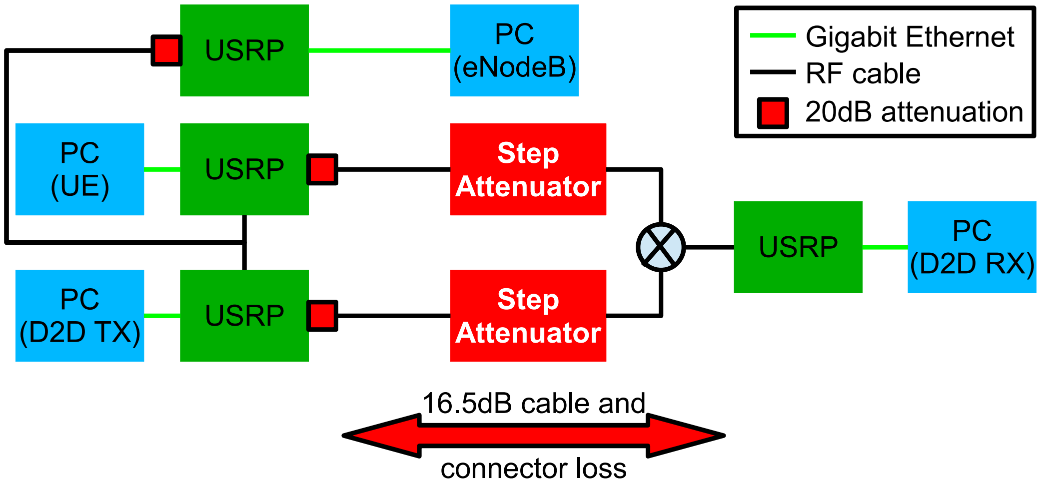

Figure 8: Co-channel measurement setup attenuations.

such that there was a fixed attenuation of 20dB, approximately 16.5dB of losses (see section 6.2.2) and a Rohde & Schwarz RSC step attenuator. Each measurement run was manually timed to be around 65 seconds long (~6500 subframes total). The recorded metrics were the SINR level and decoding success rate of the D2D receiver.

6.4.2 Traffic characterisation and channel coding

The traffic characteristics of the test traffic were similar to those used in the adjacent channel interference measurements with some additions. Firstly, the D2D receiver now attempts decoding of each subframe by performing Turbo Code decoding using five iterations. This count was chosen in an effort to balance the increased probability of successful decoding provided by a larger number of iterations and the computational demands placed on the computers used for testing. Too large a number would have resulted in subframe processing exceed its available time resulting in the next subframe being dropped in order to catch up to the current frame time as determined by the eNodeB.

The second difference stems from the fact that both transmission need to share the same resources and the receiver must therefore know which one is intended for it and which one is interference. Otherwise, when the interference grows large enough, the D2D receiver would simply switch to attempting decoding of the UE transmission. In order to provide such a separation a different physical uplink shared channel (PUSCH) index was used.

7 Results and analysis

Measurements carried out as described in chapter 6 aimed at acquiring empirical data to aid determination of numerical bounds for the main design consideration of device-to-device (D2D) communication as identified in the theoretical portion of this work presented in chapter 4. The first consideration is the impact on interference conditions of integrating new node types into cellular systems with activity patterns differing from those traditionally encountered in them. In order to assess this, measurements of the interference of cellular User Equipment (UE) in frequency-wise adjacent channels to D2D receivers was measured. The results are presented in subsection 7.1. The second consideration relates to the design of D2D control protocols and signalling. Channel quality reports form an essential part of the functionality of cellular systems and this also applies to D2D nodes. High-reliability communications in particular dependent heavily on accurate and fast detection of worsening channel conditions to take corrective action. Co-channel interference measurements were carried out to estimate the required reporting frequency. The results are presented in subsection 7.2.

7.1 Adjacent channel interference

The possibility of adjacent channel interference arises when network-controlled D2D functionality is added into cellular networks regardless of the exact protocol implementation details as discussed in chapter 4. Such interference is caused by the use of imperfect radio frequency equipment radiating power outside of the desired frequency range. The measurement of the impact of adjacent channel interference was therefore performed and here expressed with regards to the interference power level rather than successful decoding rates as these depend on the exact implementation details (e.g., D2D modulation, channel coding, etc.). Moreover, the results are presented relative to the noise floor of the equipment used. Obtaining such a baseline is important in increasing the generality of the obtained results since another type of radio receiver will likely present a different noise floor profile. Results in this section are presented as values above this baseline and should therefore be transferable to other types and models of hardware. Figure 9 presents the measured noise floor of the used USRP radio front-end presented in section 6.2.1. An empirically recorded noise floor baseline might not be even due to hardware factors such as the design of receive-side filters. Measuring the level seen on each subcarrier allows the result to be presented as the amount of added interference power resulting from the test scenario itself. Both the UE and the D2D receiver were synchronized to the same base station (BS). GPS receivers were used to help further align the clocks of all radio front-ends used. They were also allowed to warm up prior to use in a further effort to reach stable operating conditions. These steps were taken in an effort to minimise the variance in results due to difference in individual hardware units.

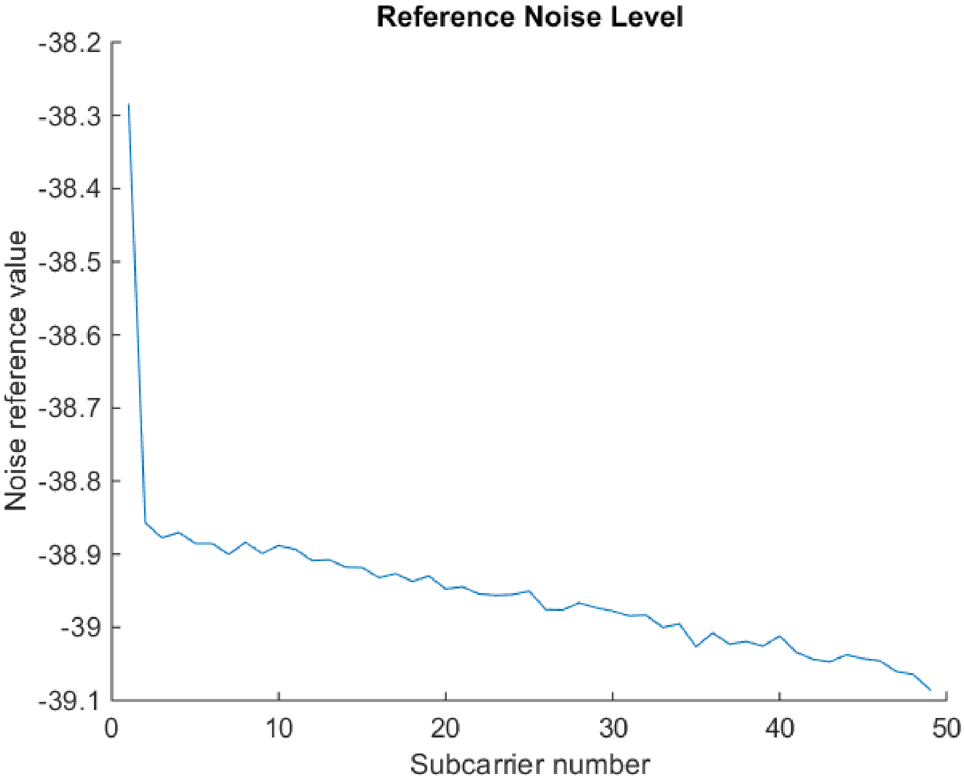

Figure 9: Measured noise floor baseline per subcarrier.

Data was recorded for all of the SC-OFDM subcarriers (see section 2.3) allocated to the D2D link. In Figure 9 as well as all subsequent figures of this section, the subcarriers are numbered 1-48 with 1 being the subcarrier closest in frequency to the adjacent channel where the interfering UE transmits. These forty-eight subcarriers correspond to four resource blocks — namely, the four lowest ones in terms of frequency. Subcarrier 1 (closest to the UE's frequency resources) presents a noticeably higher value than the other ones. This is likely due to the DC component engendered by the carrier wave. Values in Figure 9 are unitless and derive from the magnitude of the samples received by the baseband processing software. These values were recorded after the USRP's amplifier and have been scaled by the UHD driver. As such they do not correspond to any commonly used unit but in principle represent power per-subcarrier over all the OFDM symbols in one subframe. Figures in this section only depict the frequency resources allocated to the D2D link. Forty-eight subcarriers results in a signal bandwidth of four resource blocks of 180kHz each, yielding 720kHz aggregate, which is much less than the coherence bandwidth for a 2.4GHz indoor type channel [26, 68]. Some of the projected applications for D2D — such as vehicle-to-vehicle and massive machine type communication — require low throughput and can therefore be accommodated in narrowband signals.

In the labelling of the following three graphs (Figures 10-12), the transmit power levels (TXP) indicate the setting of the UHD driver of USRP-2932s used for testing.

The Y-axis values present the difference of the recorded interference power level for that particular set of parameters and the reference noise level as computed by Equation (1).

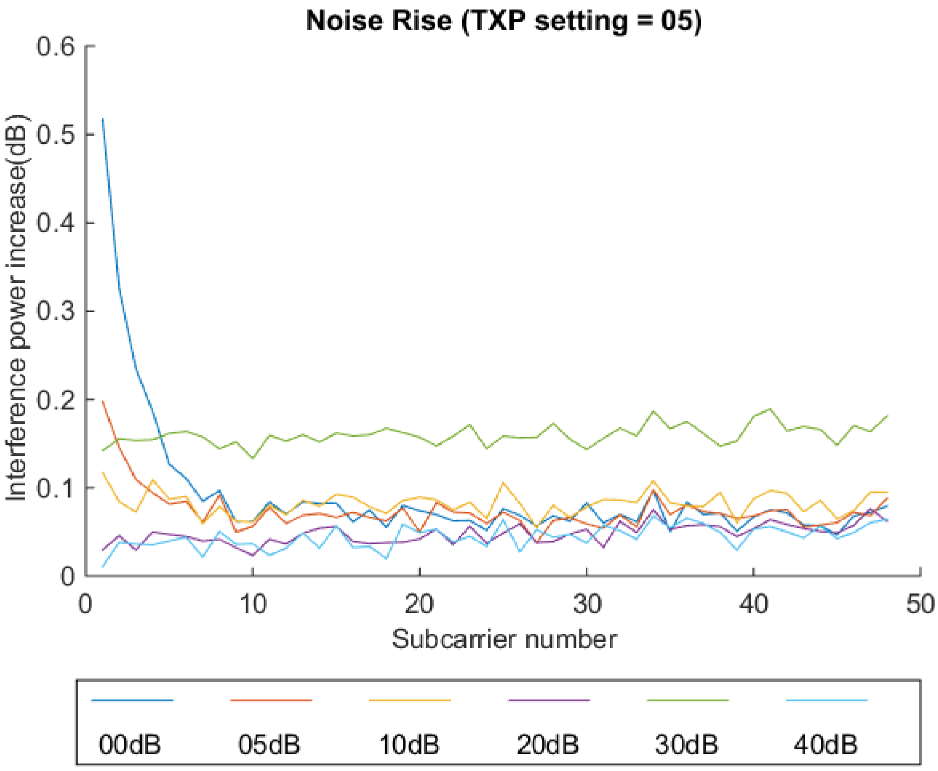

Figure 10: Measured interference power for transmit power setting 05 for UE attenuation levels 00-40dB.

$$P_R = P_M - P_N \quad (1)$$

where, P_R is the net increase in interference power, P_M is the measured power, P_N is the measured noise floor baseline

Figures 10-12 all present a similar rapidly decreasing pattern across all subcarriers. Interference power is highest on the subcarriers immediately adjacent to the ones used by the cellular transmitter. All three figures feature a sharp peak on subcarrier 1 for attenuation levels 0dB, 5dB and 10dB but absent for the higher levels of 20dB, 30dB and 40dB. This is similar to the noise floor baseline in Figure 9. In addition to being TXP-insensitive, the pattern does not vary with attenuation levels between UE and D2D receiver.

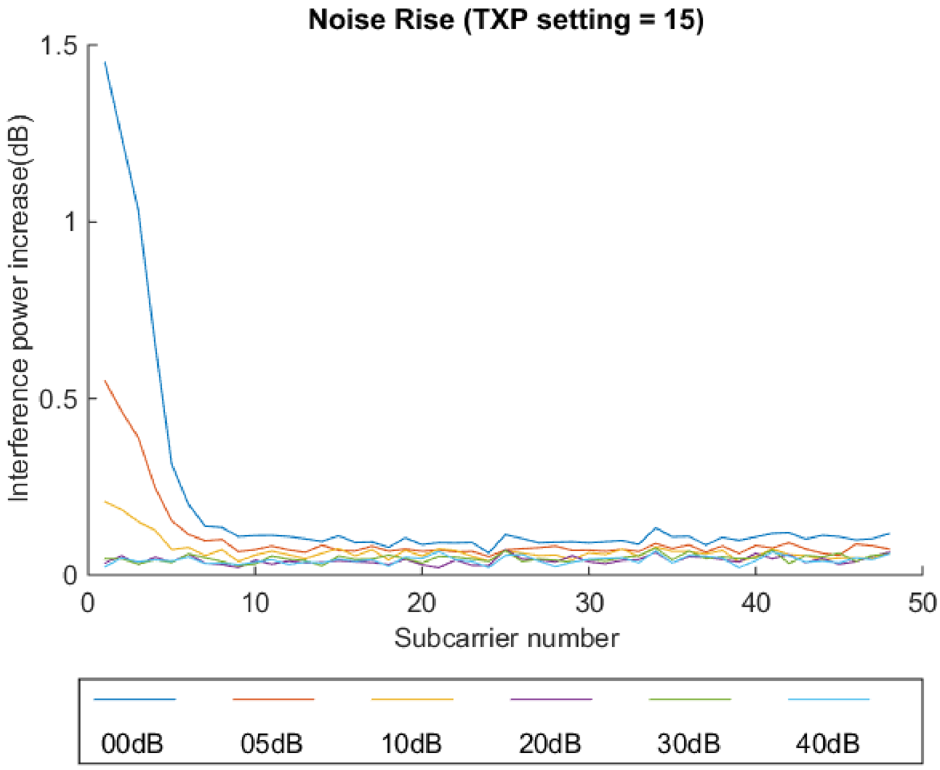

Figure 11: Measured interference power for transmit power setting 15 for UE attenuation levels 00-40dB.

In addition to being affected by the characteristics of the hardware in use, the observed patterns of high-to-low values also depend on the waveform in use. Different waveforms may cause differing amounts and distributions of interference of adjacent channels. As such, D2D interference protection margin requirements are likely to vary somewhat depending on the cellular technology providing the network control. Such margins could be added by the BS scheduler to help protect against interference level fluctuations between channel quality indicator (CQI) reports. In the case of LTE the waveform used is OFDM but this may change for 5G. Since the former did not have an explicit design requirement concerning D2D compatibility yet still exhibits relatively low interference levels in at least the particular set of circumstances tested here, it appears likely that the overall design of the next cellular technology generation will provide an improved environment for direct communications. The extent to which the interference will impact performance depends on the channel coding in use. It has been found [61] that Turbo Code decoding bit-error rate (BER) suffers in the presence of channel estimation errors. As discussed in section 3.4.2, transmission burst timings from UEs cannot always be fully anticipated leading to difficulties in channel estimation due to unpredictable interference level variations. A compounding factor comes from the uneven nature of the interference power in these figures. Channel estimators for D2D receivers must therefore be developed taking this new adjacent channel interference constraint into account. Alternatively, a protection margin could be introduced in an effort to keep SINR high enough to tolerate the higher interference levels on subcarriers close to other users' allocations. If the margin is set to the average value instead the subcarriers on the frequency-wise edges of the resource blocks in use may become unusable. Turbo coding exhibits a sharp drop in the probability of successful decoding around a narrow SINR range called the waterfall region [7]. Above it is the error floor region where BER performance varies less markedly as a result of SINR fluctuations. A protection margin would contribute to keep operations in the SINR range above this waterfall region of the error rate curve and thus helps ensure reliable data reception.

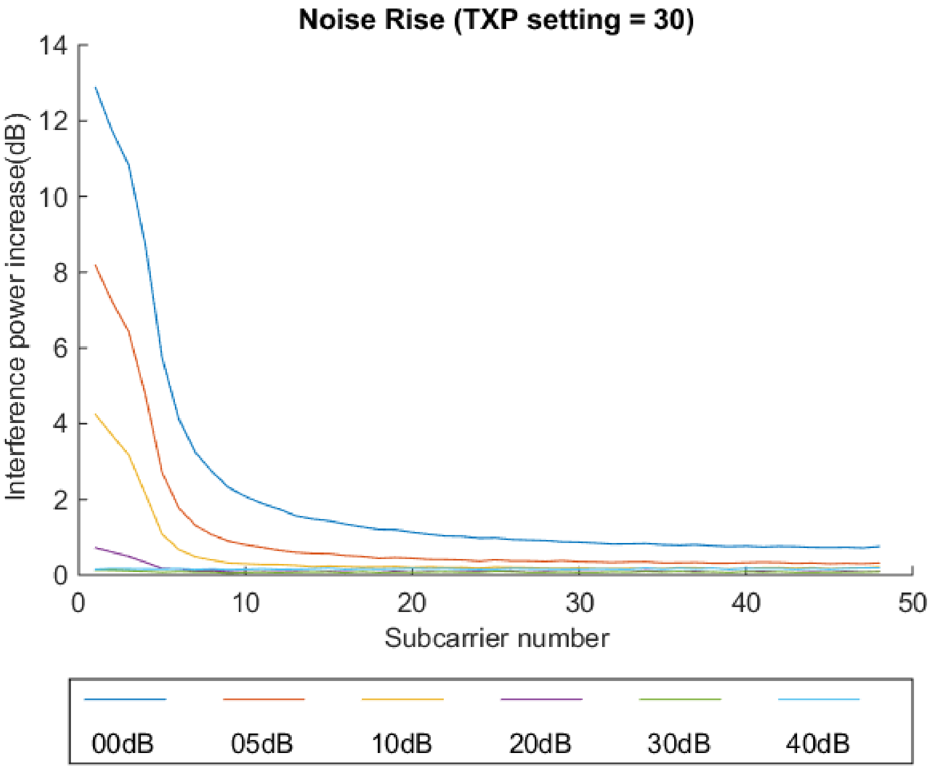

Figure 12: Measured interference power for transmit power setting 30 for UE attenuation levels 00-40dB.

From the above presented results, it can be concluded that the addition of D2D transmitters into a cellular communication system does not introduce intolerable interference levels in lightly loaded cells. UE-generated adjacent channel interference (ACI) should nonetheless be taken into account when designing high-reliability links. These measurements quantise the ACI level suffered by D2D receivers operating in the vicinity of cellular UE transmitters. Figures 10-12 show that the interference power is not even across all subcarriers. For example, in the case of 0dB attenuation in Figure 12 values range from 13dB on subcarrier 1 to 1dB on subcarrier 48. D2D nodes use such measurements to monitor the link for unacceptable degradation. They will notify their controlling BS when interference exceed some level plus a protection margin. Determining an appropriate level for alerting the BS of worsening conditions depends on the channel coding in use. Such measurements are left for future work. Since the tested power setting approaches the maximum power output level for UEs of 23dBm [19] required by the LTE specification, it provides a reasonable estimate of worst-case scenario behaviour.

The above measurements and analysis only considered the presence of a single UE located in the vicinity of the D2D receiver. As the number of users of a cell increases, so will the density of active transmitters. Remote cellular users' equipments will also contribute some interference even if at a very low level due to distance. The exact amount of this unwanted power will depend on the propagation conditions between the interferers and the D2D receiver. Work being conducted by research projects and industry groups on 5G technologies aims to deploy ultra high density cells [2, 45, 47]. This will require further analysis of the number of nearby transmitters (both UE and D2D) a D2D system might encounter in crowded scenarios. Nonetheless, the measurement results presented in Figures 10-12 indicated there exists at least some deployment possibilities.

7.2 Co-channel interference

The introduction of D2D nodes into a cellular network introduces new challenges to the interference environment seen by all nodes. In addition to the adjacent channel interference issues discussed in the previous section, the possibility of re-use of time-frequency resource for communication not involving the BS requires analysis of the impact of re-use on system reliability and performance. Re-use would in this case mean use of part or all of the resource blocks allocated to the D2D cluster for UE communication elsewhere in the cell. The results presented in this section relate to the possibility of co-existence between UE and D2D nodes on the same resource blocks as described in section 4.1 and measured as described in section 6.4.

Figures 13-15 depict graphically measured SINR values and Turbo Code decoding success rates for the D2D receiver as a function of the attenuation applied to an interfering UE. All three figures possess a similar sigmoidal shape. They are shifted versions of each other in terms of absolute value of the attenuation depending on the D2D transmitter power which was the sole parameter modified between the three runs. Each curve breaks down into three rough regions: an outage region, a transition region and a connected region in order from lower to higher SINR levels.

The figures each display some drops in the success rate when the attenuation of the interfering UE is increased. Based on this sole factor, such a drop should not occur. The reason for this behaviour arises from the soft–real-time nature of the ARF platform used in this experiment as detailed in chapter 5. When suboptimal scheduling occurs, performance of timing and latency sensitive applications suffers [13, 22, 38]. In the ARF platform, when subframes are produced late this results in them not being successfully received even though they could have been without such timing problems. Consequently, the small drops in the success rate curves present in Figures 13-15 are treated as measurement errors. The overall rising trend visible in the success rates as attenuation increases follows that expected from theory.

Since the measurements used attenuation as a proxy for distance (see section 6.4) between nodes, a model must be used in order to convert between the attenuation and distance. Literature was reviewed [9, 10, 56] to determine an appropriate mathematical relation. Based on these works, a first order approximation was selected to be:

$$L = 10\alpha log(d) \quad (2)$$
$$\leftrightarrow d = 10^{\frac{L}{10\alpha}}$$

where, α is the path loss exponent, d the inter-node distance in meters, L is the path loss in dB

The model for relating L to d in Equation (2) is very simple. Making such an approximation in this instance is justified as no specific D2D standard, implementation or environment is being considered. Instead the measurements of this section aim to establish the general feasibility of augmenting cellular systems with D2D nodes and to obtain quantitative data at order-of-magnitude precision on the interference implications of doing so. Labels in Figures 13-15 present attenuation instead of converted distance to allow for re-analysis with a more precise model if required. Table 6 gives the distance corresponding to the total attenuation for select values of the UE - D2D receiver step attenuator setting. In addition to the latter, the fixed

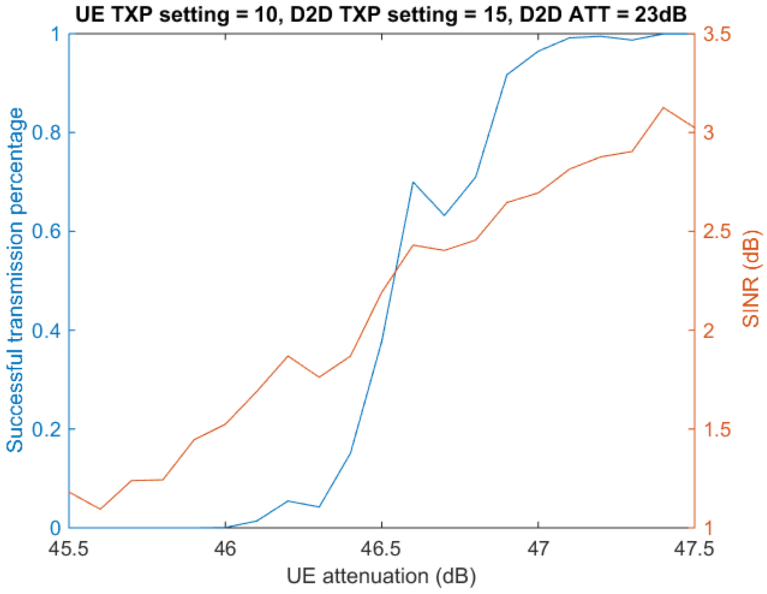

Figure 13: Successfully received subframe percentage. UE TXP = 10, D2D TXP = 15, D2D ATT = 23dB.

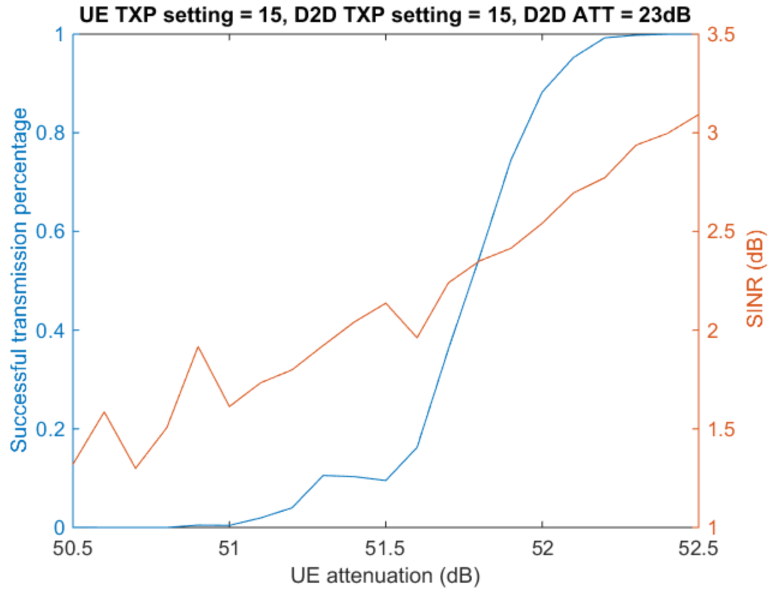

Figure 14: Successfully received subframe percentage. UE TXP = 15, D2D TXP = 15, D2D ATT = 23dB.

attenuation and losses discussed in 6.4.1 of 20dB and 16.5dB, respectively.

Table 6: Approximate attenuation setting to distance conversions. Attenuation values indicated UE - D2D RX and do not include fixed attenuation or cable and connector losses. Values are in dB for attenuation and meters for distances.

Attenuation	9	10	11	12	13	14	15	16	17	18	19
$\alpha = 3.0$	33	35	38	41	45	48	52	56	61	66	71
$\alpha = 3.5$	20	21	23	24	26	28	30	32	34	36	39
$\alpha = 4.0$	14	15	15	16	17	18	19	21	22	23	24

Turbo Code decoders exhibit a sharp drop in the successful decoding rate over a small change in SINR in the waterfall region as discussed in 7.1. This behaviour can be seen in Figures 13-15. Above and beyond the extremal attenuation values present here, the performance curve remains relatively flat and was not included for this reason. The SINR curve rises much more gradually whereas the successful decoding rate climbs faster while starting later at a higher UE attenuation value. It is therefore critical to anticipate changes in channel quality before SINR enters the waterfall region to prevent very rapid link throughput degradation. The SINR/FER curves exhibit similar but more numerous dips to the decoding success rate curves. This is due to the SINR being affected by both the D2D transmitter and the UE

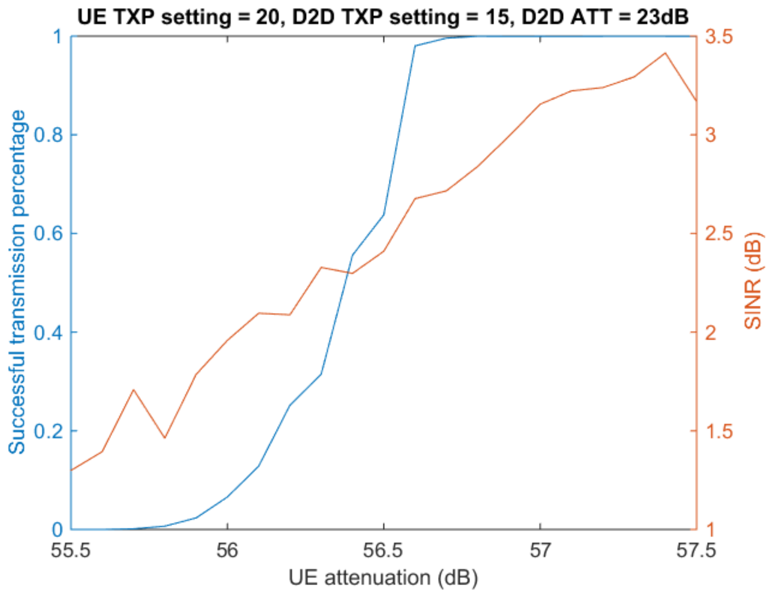

Figure 15: Successfully received subframe percentage. UE TXP = 20, D2D TXP = 15, D2D ATT = 23dB.

and can be expressed as follows :

$$SINR = \frac{P_{D2D}}{P_{UE}+P_N} \quad (3)$$

where, P_{D2D} is the D2D signal's receiver power, P_{UE} is the UE interference power, P_N is the noise power

From Equation (3), we can see that SINR suffers from fluctuations when processing in either the transmitter or receiver is late. The difference in the number of dips between the SINR and decoding success rate curves comes from the fact that for the latter the signal quality either suffices to decode or does not. In effect, the decoding rates perform a binary mapping of the SINR to decoding success. As such, so long as the average conditions improve, the decoding rate average curve will also tend to rise.

Larger UE transmit powers often correspond to nodes located further away from their serving base stations. Such cell-edge location present a greater interference challenge to the D2D link as can be seen from the figures above. When UE TXP increases, the distance from it to the D2D receiver must also increase. This leads to a larger exclusion zone for resource re-use. An exclusion zone is here understood to be the area in which no UE should be scheduled on the same time-frequency resources as a protected high-reliability D2D cluster operates. Nodes located closer to the centre of the cell enables denser packing of time-frequency sharing devices. The extent of this phenomenon depends on the size of cells in use, which in turn depends on the network's cell layout. This is in turn indirectly partly dependent on the cellular technology since each presents different technical characteristics and offers different possibilities. Ultra-dense networks is one of the concepts being studied for use in 5G networks [24, 47]. Greater cell densities to cover a given area than currently will result in reduced transmit power needs and will therefore help to alleviate the issue of D2D re-use exclusion zones growing large at cell edges due to greater TXP. On the other hand, node densities will likely increase. This may result in the advantages afforded by reduced average transmit powers being partially or entirely offset.

7.3 Interference impact on CQI reporting

The results presented in the preceding sections can be used as the basis for determining the required update periodicity of a system incorporating a protected D2D link. In each of Figures 13-15, a rise in interference of 2dB proves enough to bring the link from practically error free to being virtually in outage. As such, the maximum tolerable SINR change of 2dB obtained from the measurements imposes an upper limit on the time between periodic channel quality indicators (CQI). The required periodicity of updates may be computed from the maximum allowable change, the node speed and Equation (2) :

$$\delta t = \frac{10^{\frac{ST}{10}}}{10^{\frac{10\alpha log10(d) - 10\alpha log10(ref)}{10}}} \quad (4)$$

where, α is the path loss exponent, d the distance after one second in meters (ref + node velocity), ref is the starting point used as reference, ST the maximum allowable SINR drop between CQI reports, δt the maximum period between CQI updates

Equation (4) yields a maximum update period in seconds. Values obtained using this equation are presented in tables 7 and 8 with values expressed in milliseconds for clarity's sake. It should be noted that the results only consider the effect of interferers and not for example fading effects. The thresholds used are all below 2dB for the reasons discussed above. Values were calculated for nodes operating at two different distances from their serving BSs: 25m and 100m. The selected distances are at the lower end of cell sizes typically considered [6, 21, 28] and therefore represent more challenging conditions than average in terms of path loss change. The reference distance is used as the starting point with the node moving away from the BS. For a given speed and threshold, the values in tables 7 and 8 indicate how often CQI reports should be made to avoid exceed channel quality changing by said threshold value or more.

The results in Table 7 vary by an order of magnitude. Even the smaller values are on the order of hundreds of milliseconds appear entirely feasible and furthermore low

Table 7: Required inter-CQI update periods for $\alpha = 3.5$ around a distance to serving base station of 100m for various device speeds and SINR change thresholds. Values in milliseconds.

Speed (m/s)	1.67	5	10	20	30
ST = 0.5dB	981.8	757.2	520.5	258.2	135.4
ST = 1.0dB	1101.6	849.6	584.0	289.7	152.0
ST = 1.5dB	1236.0	953.3	655.3	325.0	170.5

overhead since LTE itself has a system broadcast channel changing every forty milliseconds. It should be again emphasised that these values only consider interference level changes caused by a moving interferer's path loss varying over time. Any real world system would have to account for many more factors affecting radio wave propagation.

Table 8: Required inter-CQI update periods for $\alpha = 3.5$ around a distance to serving base station of 25m for various device speeds and SINR change thresholds. Values in milliseconds.

Speed (m/s)	1.67	5	10	20	30
ST = 0.5dB	666.3	258.2	74.5	9.8	2.0
ST = 1.0dB	747.6	289.7	83.6	11.0	2.2
ST = 1.5dB	838.8	325.0	93.8	12.4	2.5

Table 8 presents the values obtained for the case of a 25m distance to the serving base station reference point (centre of the cell). Results for this case varying by two orders-of-magnitude whereas for the 100m case the variation was only one order-of-magnitude. This illustrates the greater changes in attenuation for the same distance travelled for smaller absolute distances due to path loss. At the longer end of the scale, periods on the order of hundreds of milliseconds again appear entirely feasible. At the lower end of the scale however, periods on the order of milliseconds prove problematic since the length of a frame is 10ms in LTE. Consequently, if a D2D cluster has only one transmission grant per frame from its serving BS, nodes will be unable to report sufficiently often. Periods on the order of ten milliseconds may pose problems in LTE as well since this means reporting will occur each frame. Such a constraint considerably reduces the scheduling freedom of the eNodeB. The D2D cluster must have a scheduling grant for reporting every frame in order to be able to meet its reporting requirements. Allocation can be done either per-subframe or in a more permanent fashion. Alternatively, the eNodeB's scheduling freedom becomes constrained by which time-frequency resources it can schedule high-velocity UEs in.

Frequent reporting also causes some or all of the D2D nodes to perform link quality measurements continuously if the D2D cluster has been granted only one time-frequency resource per-frame. Such a case leads to the nodes spending most of their time performing measurements instead of their primary task. While it may seem overly pessimistic to only consider the case where the D2D link transmitting data operates at the lower bound of its operational SINR region, it must be kept in mind that cellular UEs do not transmit constantly and such idle periods make scheduling much harder for the serving BS from the viewpoint of interference avoidance with D2D links. The number presented in tables 7 and 8 therefore indicate a worst case scenario where the D2D cluster is already experiencing challenging conditions due to, for example, channel conditions deteriorating. Such worst-case considerations are especially important for high-reliability links. Single master systems (see section 4.6) appear less suitable for applications requiring high reliability links and high mobility. With frequent reporting, the single master will experience high load. A further factor to consider is the increased power consumption from frequent CQI reporting and high signalling load in general for mobile nodes requiring links to be kept permanently available for low latency messaging.

The previous discussion and calculations only considered path loss. A real system would also be affected by fading. Mobile nodes will encounter different channel conditions when they change locations. This change does not apply linearly between the starting and destination location due to the Doppler effect. Relative motion between the transmitter and receiver will cause velocity dependent time-variance in the channel. The period of time during which a channel remains relatively constant by some definition is known as the coherence time. The equations presented in [60] are used to obtain an approximation for the channel coherence time :

$$f_d = \frac{V}{\lambda} \quad (5)$$

and

$$T_0 \approx \frac{1}{f_d} \quad (6)$$

where, f_d is the Doppler spread, T_0 is the coherence time, V is the node velocity and λ is the wavelength

Table 9: Approximate channel coherence times due to Doppler spread for a carrier frequency of 2.48GHz. Values in milliseconds.

Speed (m/s)	1.67	5	10	20	30
Coherence time	72.5	24.2	12.1	6.1	4.0

The values in Table 9 are small compared to the majority of values in tables 7 and 8. Consequently, maintaining a suitably good channel estimate between the receiver and transmitter will require more frequent updates than the results based on path loss alone would suggest. It should be note however that at the lower end of the value range in 8, inter-report times are shorter than the shortest coherence times in 9. Since the measurements performed for this measurements were performed with static nodes over cables, the assessment of Doppler and multipath induced channel time-variance effects cannot be quantified further than the above presented estimation.

7.4 Applicability of results

The results obtained are broadly applicable to other D2D implementation and protocols in situations similar in terms of node count, transmit power and attenuations. Measurements were performed in a way to be as D2D protocol agnostic as possible. Furthermore, data was recorded in terms of SINR and successful decoding rates. Both of these metrics are broadly applicable to many protocols. Two factors must still however be taken into account when transferring elsewhere the results presented in this chapter.

Firstly, the measurement set-up simulated a very lightly loaded cell wherein the base station was able to place only one static interfering UE in proximity to the D2D pair. This leads to lower levels of interference and therefore presents an easier task for the D2D receiver in terms of Turbo decoding. Real-world UEs would be mobile. This will introduce more variance in the received interference power (e.g., Doppler effects) that the attenuation measurements here cannot capture.

Secondly, the test code was added to a software-defined radio (SDR) platform without hard real-time guarantees (see sections 5.1 and 5.5). As such, some subframes may have been lost due to software issues or processing delays instead of being lost to interference. A production quality system would not experience as many such late subframes and would consequently produce better performance. These late subframes are here treated as measurement errors.

Finally, the velocities considered in section 7.2 should be applicable to many situations. Urban situations rarely permit travelling much in excess of 30m/s. Industrial applications will likely see very low velocities as vehicle speeds in and around factories are typically limited for safety reasons. Indoor office and commercial environment velocities would presumably be lower still.

8 Conclusions

Maturation of commodity wide-coverage wireless communication systems has led to a situation where data rates on offer exceed the needs of many applications at economically feasible prices. Yet, the main use case remains the same as it was at the dawn of the cellular era: communication between humans. Enabling the extension of networks to facilitate latency and reliability critical applications requires upcoming standards to account for new types of requirements and offer new types of solutions in return. Device-to-device (D2D) communication offers the prospect of enabling these new types of uses by tailoring the protocol details to the needs of a particular D2D application while simultaneously retaining the benefits of controlled access to the spectrum by the licensed network operator.

Regardless of the use case, all network-controlled D2D schemes integrating into cellular systems require careful design to prevent new sources of interference degrading system performance instead of the desired gain. The measurements performed in the work were selected to establish bounds within which designs must fit in order to mitigate negative impacts caused by interference. Analysis of the results obtained highlighted the need to consider the balance between increased spectral efficiency through resource re-use and protection offered to high-reliability links. High-speed cellular transmitters in particular impose the need to send frequent channel condition reports from the D2D cluster to the base station (BS). These D2D standard agnostic measurements constitute a broad framework to guide the development of standards for various specific use cases each having its own requirements. The issue of node velocities is likely to prove a much more critical consideration in the development of vehicle-to-vehicle (V2V) applications than it is for industrial automation where machinery installed in a factory is unlikely to move far or often.

Beyond straightforward interference measurements, the extension of the ARF platform performed for this work has demonstrated the feasibility of D2D communication even within today's Long Term Evolution (LTE) standard designed without specific support for it. The selected software based approach to building a demonstrator possesses the additional advantage of serving to highlight one of the benefits of D2D, namely flexibility. In the process of implementing the test-bed for this work, many iterations of the D2D functionality were created without modification of the LTE portion of the code base or incurring additional hardware costs. In a real world deployment this would represent the ability to update the functioning of a particular cluster of co-operating nodes without need to modify equipment owned by the network operator or to obtain regulatory clearance for any revised hardware.

Adapting the protocols used in a part of the spectrum to the needs of each group of communicating nodes results in improving the feasibility of integrating multiple types of telecommunication services into single devices. By sharing components and code, development cost and time as well as additional hardware costs can be kept low. The obtained measurement results and the bounds derived from them using a distance-attenuation model in chapter 7 demonstrate the co-existence of multiple node types to be possible under at least the conditions used for this work. Networks can offer new and improved services without the need to adapt network core infrastructure. For high-reliability links this means the ability to utilise the established authentication and authorisation framework of the cellular network as well as its ability to relay messages over long distances as a back-up to direct communication. Should the use case require the nodes' hardware to be moved, it can be done without further work since the cellular technology likely covers a large geographic area or even the whole World, thus ensuring the need to support only a single standard. Similarly, in case the application's requirements evolve, the intra-cluster D2D protocol can be updated independently of the cellular side implementation. In fact, hardware could conceivably be manufactured in bulk such that only the D2D part of the device need be implemented and tailored to a particular application's needs.

Enabling the above mentioned benefits, more research needs to be performed in order to fully establish the bounds of what type of D2D design patterns work and which do not. Testing system functionality with an increased number of nodes operating using antennas as opposed to cables would serve to confirm and further refine the findings of this work in terms of interference levels experienced by the various receivers in the network. Mobile nodes in particular would be important to validate the numerical predication based on a simple path-loss model presented in chapter 7. Beyond these, developing a larger scale system will require algorithms to make decisions concerning the use of D2D for both terminals and network equipment. For BSs this involves developing scheduling algorithms capable of efficiently arbitrating the needs of cellular and D2D users and to further be able to effect guarantees of reliability for critical links. On the other hand, terminals will need to make the decision as to whether they will request cellular or direct communication resources as well as how to handle outages in the latter. For some application reverting back to BS relayed communications, along with its associated increase in latency, constitutes a viable fall-back while not for others. Effective scheduling choices constitute a key part of the networks' ability to offer the protected resource for nodes to build high-reliability communication links amongst themselves. These scheduling algorithms will aim to maximise system performance within the constraints and bounds identified in this work.

References

[1] 3GPP. LTE. http://www.3gpp.org/technologies/keywords-acronyms/98-lte, Date of publication unknown. [Online; accessed 08/08/2015].

[2] Rysavy Research/4G Americas. Beyond LTE: Enabling the mobile broadband explosion, August 2015.

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

[4] M. Bauer, G. May, and V. Jain. A wireless gateway approach enabling industrial real-time communication on the field level of factory automation. In *IEEE Emerging Technology and Factory Automation (ETFA)*, pages 1–8, September 2014.

[5] Alan Bensky. *Communications Engineering : Short-range Wireless Communication: Fundamentals of RF System Design and Application (2nd Edition)*. Newnes, December 2003.

[6] M. Brau, Y. Corre, and Y. Lostanlen. Assessment of 3D network coverage performance from dense small-cell LTE. In *IEEE International Conference on Communications (ICC)*, pages 6820–6824, June 2012.

[7] Qingchun Chen, Jiaxiang Li, Pingzhi Fan, and Vahid Tarokh. On the performance estimate for Turbo codes. In *International Conference on Wireless Communications and Signal Processing (WCSP)*, pages 1–5, October 2010.

[8] Kae Won Choi and Zhu Han. Device-to-device discovery for proximity-based service in LTE-Advanced system. *IEEE Journal on Selected Areas in Communications*, 33(1):55–66, January 2015.

[9] T. Chrysikos, G. Georgopoulos, and S. Kotsopoulos. Attenuation over distance for indoor propagation topologies at 2.4 GHz. In *IEEE Symposium on Computers and Communications (ISCC)*, pages 329–334, June 2011.

[10] T. Chrysikos, S. Kotsopoulos, and G. Karagiannidis. Attenuation over distance and excess path loss for a large-area indoor commercial topology at 2.4 GHz. In *19th International Conference on Telecommunications (ICT)*, pages 1–6, April 2012.

[11] Cisco. Cisco Visual Networking Index: Global mobile data traffic forecast update, 2014–2019, February 2015.

[12] U. Datta and S. Kundu. Packet level performance of a CDMA wireless sensor networks in presence of correlated interferers. In *4th International Conference on Computers and Devices for Communication (CODEC)*, pages 1–4, December 2009.

[13] Kenneth J. Duda and David R. Cheriton. Borrowed-virtual-time (BVT) scheduling: Supporting latency-sensitive threads in a general-purpose scheduler. *SIGOPS Oper. Syst. Rev.*, 33(5):261–276, December 1999.

[14] A. Duel-Hallen. Fading channel prediction for mobile radio adaptive transmission systems. *Proceedings of the IEEE*, 95(12):2299–2313, December 2007.

[15] Ericsson. Ericsson Mobility Report, February 2013.

[16] ETSI. *LTE; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2 Release 8.* TS 36.300 version 8.12.0, April 2010.

[17] ETSI. *LTE; Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Architecture description Release 8.* TS 36.401 version 8.8.0, June 2010.

[18] ETSI. *LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification Release 8.* TS 36.321 version 8.12.0, March 2012.

[19] ETSI. *LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception Release 8.* TS 36.101 version 8.26.0, February 2015.

[20] 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*, 50(3):170–177, March 2012.

[21] X. Gelabert, Guohua Zhou, and P. Legg. Mobility performance and suitability of macro cell power-off in LTE dense small cell HetNets. In *IEEE 18th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pages 99–103, September 2013.

[22] A. Goel, J. Walpole, and M. Shor. Real-rate scheduling. In *Proceedings of the 10th IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS)*, pages 434–441, May 2004.

[23] Ashvin Goel, Luca Abeni, Charles Krasic, Jim Snow, and Jonathan Walpole. Supporting Time-sensitive Applications on a Commodity OS. *SIGOPS Oper. Syst. Rev.*, 36(SI):165–180, December 2002.

[24] A. Gupta and R.K. Jha. A survey of 5G network: Architecture and emerging technologies. *Access, IEEE*, PP(99):1–1, July 2015.

[25] Olivier Hersent, David Boswarthick, and Omar Elloumi. *Internet of Things : Key Applications and Protocols (2nd Edition).* Wiley, January 2012.

[26] V. Hinostroza, H. Garces, and G. Flores. Review of frequency selectivity parameters for broadband signals. In *European Conference on Wireless Technologies*, pages 173–176, October 2007.

[27] Jongwoo Hong, Seungil Park, and Sunghyun Choi. Neighbor device-assisted beacon collision detection scheme for D2D discovery. In *International Conference on Information and Communication Technology Convergence (ICTC)*, pages 369–370, October 2014.

[28] Jian Huang, Hong Zhang, Wei Xu, and Hua Zhang. Grouping based inter-cell interference coordination in LTE-A dense small-cell networks. In *IEEE 5th International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications (MAPE)*, pages 78–83, October 2013.

[29] IEEE. *Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs)*. IEEE Std 802.15.4-2011, September 2011.

[30] National Instruments. *NI USRP-2932*, December 2012.

[31] ITU-R. *REPORT ITU-R M.2134 Requirements related to technical performance for IMT-Advanced radio interface(s)*, November 2008.

[32] ITU-R. *Recommendation ITU-R M.2012-1*, February 2014.

[33] Lianghai Ji, A. Klein, N. Kuruvatti, R. Sattiraju, and H.D. Schotten. Dynamic context-aware optimization of D2D communications. In *IEEE 79th Vehicular Technology Conference (VTC Spring)*, pages 1–5, May 2014.

[34] Kalle Ruttik Riku Jäntti Jussi Kerttula, Nicolas Malm and Olav Tirkkonen. Implementing TD-LTE as software defined radio in general purpose processor. In *The Software Radio Implementation Forum*, August 2014.

[35] Hu Kai and Cui Chen. An improved method for generation of multiple uncorrelated Rayleigh fading waveforms. In *International Conference on Communications and Mobile Computing (CMC)*, volume 2, pages 72–76, April 2010.

[36] Donghan Kim, Yongjun Kwak, Jinyoung Oh, Younsun Kim, and Juho Lee. Discovery resource grouping for D2D discovery for mitigation of in-band emission in LTE- Advanced. In *Globecom Workshops (GC Wkshps)*, pages 869–874, December 2014.

[37] Long Le, A. Festag, R. Baldessari, and Wenhui Zhang. Vehicular wireless short-range communication for improving intersection safety. *IEEE Communications Magazine*, 47(11):104–110, November 2009.

[38] Jacob Leverich and Christos Kozyrakis. Reconciling high server utilization and sub-millisecond quality-of-service. In *Proceedings of the Ninth European Conference on Computer Systems*, EuroSys '14, pages 4:1–4:14, New York, NY, USA, April 2014.

[39] Jialin Li, Naveen Kr. Sharma, Dan R. K. Ports, and Steven D. Gribble. Tales of the Tail: Hardware, OS, and Application-level Sources of Tail Latency. In *Proceedings of the ACM Symposium on Cloud Computing*, SOCC '14, pages 9:1–9:14, New York, NY, USA, November 2014.

[40] Y.J. Liang and B. Girod. Network-adaptive low-latency video communication over best-effort networks. *IEEE Transactions on Circuits and Systems for Video Technology*, 16(1):72–81, January 2006.

[41] K. Lidström and T. Larsson. A spatial QoS requirements specification for V2V applications. In *IEEE Intelligent Vehicles Symposium (IV)*, pages 548–553, June 2010.

[42] Akhilesh Lunge and Prashant Borkar. A review on improving traffic flow using cooperative adaptive cruise control system. In *2nd International Conference on Electronics and Communication Systems (ICECS)*, pages 1474–1479, February 2015.

[43] R.G. Meyer and A.K. Wong. Blocking and desensitization in RF amplifiers. *IEEE Journal of Solid-State Circuits*, 30(8):944–946, August 1995.

[44] B. Panigrahi, H.K. Rath, and A. Simha. Interference-aware discovery and optimal uplink scheduling for D2D communication in LTE networks. In *Twenty First National Conference on Communications (NCC)*, pages 1–6, February 2015.

[45] CROWD Project. Final architecture design 1.1, March 2015.

[46] METIS Project. Deliverable D6.2 initial report on horizontal topics, first results and 5G system concept, April 2014.

[47] METIS Project. Deliverable D6.6 final report on the METIS 5G system concept and technology roadmap, April 2015.

[48] Kishore Kumar Pusukuri, Rajiv Gupta, and Laxmi N. Bhuyan. Shuffling: A Framework for Lock Contention Aware Thread Scheduling for Multicore Multiprocessor Systems. In *Proceedings of the 23rd International Conference on Parallel Architectures and Compilation*, PACT '14, pages 289–300, New York, NY, USA, August 2014.

[49] A. Pyattaev, K. Johnsson, A. Surak, R. Florea, S. Andreev, and Y. Koucheryavy. Network-assisted D2D communications: Implementing a technology prototype for cellular traffic offloading. In *IEEE Wireless Communications and Networking Conference (WCNC)*, pages 3266–3271, April 2014.

[50] Wei Quan, Changqiao Xu, Jianfeng Guan, Hongke Zhang, and L.A. Grieco. Scalable name lookup with adaptive prefix Bloom filter for named data networking. *IEEE Communications Letters*, 18(1):102–105, January 2014.

[51] A.S. Ragab and M.M. Mokhtar. A new iterative method for improving equalizer performance in mobile fading channels. In *Thirteenth National Radio Science Conference (NRSC)*, pages 555–562, March 1996.

[52] B. Raghothaman, E. Deng, R. Pragada, G. Sternberg, T. Deng, and K. Vanganuru. Architecture and protocols for LTE-based device to device communication. In *International Conference on Computing, Networking and Communications (ICNC)*, pages 895–899, January 2013.

[53] Alessandro Ricci and Andrea Santi. Programming Abstractions for Integrating Autonomous and Reactive Behaviors: An Agent-oriented Approach. In *Proceedings of the 2Nd Edition on Programming Systems, Languages and Applications Based on Actors, Agents, and Decentralized Control Abstractions*, AGERE! 2012, pages 83–94, New York, NY, USA, October 2012.

[54] Luigi Rizzo, Giuseppe Lettieri, and Vincenzo Maffione. Speeding Up Packet I/O in Virtual Machines. In *Proceedings of the Ninth ACM/IEEE Symposium on Architectures for Networking and Communications Systems*, ANCS '13, pages 47–58, Piscataway, NJ, USA, October 2013.

[55] Max Robert. *Bluetooth: A Short Tutorial.* Springer, December 2000.

[56] P.M. Santos, T.E. Abrudan, A. Aguiar, and J. Barros. Impact of position errors on path loss model estimation for device-to-device channels. *IEEE Transactions on Wireless Communications*, 13(5):2353–2361, May 2014.

[57] Rohde & Schwarz. *R&S RSC Step Attenuator Specifications*, December 2012.

[58] Rohde & Schwarz. *R&S SMBV100A Vector Signal Generator Specifications*, May 2015.

[59] Bluetooth SIG. *Specification of the Bluetooth System Core Package v4.2.* Specification of the Bluetooth System, December 2014.

[60] B. Sklar. Rayleigh fading channels in mobile digital communication systems .I. Characterization. *IEEE Communications Magazine*, 35(7):90–100, July 1997.

[61] Minying Sun, Yuan Li, and Sumei Sun. Impact of SNR estimation error on Turbo code with high-order modulation. In *IEEE 59th Vehicular Technology Conference (VTC)*, volume 3, pages 1320–1324 Vol.3, May 2004.

[62] D. Sweeney. An introduction to Bluetooth a standard for short range wireless networking. In *15th Annual IEEE International ASIC/SOC Conference*, pages 474–475, September 2002.

[63] Tektronix. *RSA6000 Series Data Sheet*, November 2011.

[64] A. Thanos, S. Shalmashi, and Guowang Miao. Network-assisted discovery for device-to-device communications. In *IEEE Globecom Workshops (GC Wkshps)*, pages 660–664, December 2013.

[65] M. Tjiong and J. Lukkien. On the false-positive and false-negative behavior of a soft-state signaling protocol. In *International Conference on Advanced Information Networking and Applications (AINA)*, pages 971–979, May 2009.

[66] N.L.M. van Adrichem and F.A. Kuipers. Globally accessible names in named data networking. In *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 345–350, April 2013.

[67] Peng Wang, Wu Wei, and Li Zhuoming. System performance of LTE-advanced network with D2D multi-hop communication. In *3rd International Conference on Consumer Electronics, Communications and Networks (CECNet)*, pages 645–648, November 2013.

[68] Ye Wang, Wen-Jun Lu, and Hong-Bo Zhu. Propagation characteristics of the LTE indoor radio channel with persons at 2.6 GHz. *IEEE Antennas and Wireless Propagation Letters*, 12:991–994, July 2013.

[69] Lili Wei, R.Q. Hu, Yi Qian, and Geng Wu. Enable device-to-device communications underlaying cellular networks: challenges and research aspects. *IEEE Communications Magazine*, 52(6):90–96, June 2014.

[70] Osman N.C. Yilmaz, Y.-P.Eric Wang, Niklas A. Johansson, Nadia Brahmi, Shehzad A. Ashraf, and Joachim Sachs. Analysis of ultra-reliable and low-latency 5g communication for a factory automation use case. In *IEEE International Conference on Communication Workshop (ICCW)*, pages 1190–1195, June 2015.

[71] Jisheng Zhao, Jun Shirako, V. Krishna Nandivada, and Vivek Sarkar. Reducing Task Creation and Termination Overhead in Explicitly Parallel Programs. In *Proceedings of the 19th International Conference on Parallel Architectures and Compilation Techniques*, PACT '10, pages 169–180, New York, NY, USA, September 2010.

[72] Liang Zhou, K. Ruttik, O. Tirkkonen, and R. Berry. Cell-edge inversion by interference cancellation for downlink cellular systems. In *IEEE International Conference on Communication Systems (ICCS)*, pages 127–131, November 2014.

[73] Yuming Zhu, Xiuming Shan, and Yong Ren. A power efficient adaptive modulation scheme over fading channel. In *IEEE International Conference on Communications, Circuits and Systems and West Sino Expositions*, volume 1, pages 257–261 vol.1, June 2002.

[74] E. Zihan, Kae Won Choi, and Dong In Kim. Distributed random access scheme for collision avoidance in cellular device-to-device communication. *IEEE Transactions on Wireless Communications*, 14(7):3571–3585, July 2015.

[75] Ewaldo Zihan and Kae Won Choi. Random access protocol for collision avoidance in cellular device-to-device communication. In *IEEE International Conference on Communications (ICC)*, pages 2369–2374, June 2014.