BRNO UNIVERSITY OF TECHNOLOGY

FACULTY OF ELECTRICAL ENGINEERING
AND COMMUNICATION

DEPARTMENT OF TELECOMMUNICATIONS

Ing. Pavel Mašek

HETEROGENEOUS CONNECTIVITY OF MOBILE DEVICES IN 5G WIRELESS SYSTEMS

HETEROGENNÍ PROPOJENÍ MOBILNÍCH ZAŘÍZENÍ V BEZDRÁTOVÝCH SYSTÉMECH 5. GENERACE

SHORTENED VERSION OF PH.D. THESIS

Specialization:	Teleinformatics
Supervisor:	doc. Ing. Jiří Hošek, Ph.D.
Opponents:	doc. Ing. Rastislav Róka, Ph.D.
	doc. Ing. Radek Martinek, Ph.D.
Date of Defense:	5. 9. 2017

KEYWORDS
5G, Cellular Systems, Device-to-Device (D2D) communication, Dynamic Spectrum Allocation, Internet of Things, Licensed Shared Access, Machine-to-Machine, Networking, Proximity-based communication.

KLÍČOVÁ SLOVA
5. Generace mobilních sítí, Dynamické alokování spektra, Internet věcí, Komunikace mezi stroji, Komunikace zařízení v bezprostřední blízkosti, Přímá komunikace mezi zařízeními, Sdílený přístup k licencovanému spektru, Síťová komunikace.

ARCHIVED IN
Dissertation is available at the Science Department of Dean's Office FEEC, Brno University of Technology, Technická 10, Brno, 616 00

MÍSTO ULOŽENÍ PRÁCE
Disertační práce je k dispozici na Vědeckém oddělení děkanátu FEKT VUT v Brně, Technická 10, Brno, 616 00

ISBN 80-214-
ISSN 1213-4198

CONTENTS

1 INTRODUCTION

It was not more than a few decades since mobile wireless communications were introduced with the first generation (1980's) – voice-only communication systems. Over the last couple of decades the world has witnessed gradual, yet steady evolution of mobile wireless communications towards second (circuit switching systems), third (utilizing both circuit and packet switching systems), fourth and now fifth generation (all-IP) wireless networks, which are using packet switching. Along with mentioned factors, approaches also differentiate between licensed spectrum and unlicensed spectrum [1]. Looking back to the history, all four recent generations of cellular systems have been evolved over approximately 10-year intersecting cycles and therefore, many expect that the next major evolution in wireless communications – the 5th generation – will be implemented and fully deployed around 2020 and beyond [2].

Any massive real-world deployment is always preceded by thorough research stage and since we, as a research society, are currently around four years prior to the expected roll out of next-generation 5G mobile systems, it is not surprising that 5G is the most turbulent topic among research community nowadays with hundreds of scientific articles indexed by international databases every year. The key of such extreme interest lies in the anticipated revolutionary character and high heterogeneity of the future 5G network's architecture combining the aspects of emerging ultra-high-frequency spectrum access, hyper-connected vision, new application-specific requirements and much more [1].

1.1 Research Motivation

A quick look into recent wireless network statistics reveal that global mobile traffic grew 63 % in 2016 and almost half a billion (429 million) mobile devices and connections were added in 2016 [3]. Globally, smart devices represented 46 % of the total mobile devices and connections in 2016; they accounted for 89 % of the mobile data traffic. Another interesting finding is that smartphones represented only 45 % of total mobile devices and connections in 2016, but represented 81 % of total mobile traffic. Cisco's Visual Networking Index (VNI) forecasts that by 2021, nearly three-quarters of all devices connected to the mobile network will be "smart" [3].

Based on the above mentioned facts, 5G is introducing a breakthrough shift, barely comparable with previous generations, based on completely new technologies and brilliant innovations, which go beyond our current imagination. Inspired by the above, this dissertation thesis discusses key elements which have been recently introduced as the potential 5G enablers. As each of previously mentioned mobile generations has been driven by new application use-cases and the associated users' demands, we assume the same also for upcoming 5G networks. Due to that, this dissertation work is, as well, giving an insight into end user's perception of newly arriving technological changes and applications which is an essential performance indicator of overall adoption of 5G. One of the most rapidly developing applications, where the 5G is expected to fulfill its needs, is Internet of Things (IoT). Moreover, researchers are exploring new applications in directions of augmented reality (AR) and virtual reality (VR), Internet of Vehicles (IoV), Tactile Internet (TI), Device-to-Device (D2D) communications, Machine-to-Machine (M2M) communications, named also Machine-Type Communications (MTC) [4].

Since it is already clear that not all these challenges can be accommodated even by current wireless networks, the next-generation (5G) networks should take this role and act as an enabler for upcoming communication use-cases. It is more than evident that 5G is a very broad research topic with many technical, social and business aspects [4], [5], [6]. Therefore, it is not an objective of present dissertation thesis "*Heterogeneous Connectivity of Mobile Devices in 5G Wireless Systems*" to cover all of them, but rather explore in more detail selected domains.

1.2 State-of-the-art Issues of 5G Wireless Systems

Over the last few years, it was not even clear what the "5G" really stands for and what kind of technologies, communication protocols and applications will be the biggest drivers of this new cellular infrastructure. As the technology pillars in the architecture of future 5G mobile networks were identified, a diversity of wireless technologies will collaborate to support the 5G communication networks with their demanding applications and services. Despite decisive progress in many enabling solutions, next-generation cellular deployments may still suffer from a glaring lack of bandwidth due to inefficient utilization of radio spectrum, which calls for immediate action.

As technical envelope of 5G vision, there are several broadly discussed performance criteria which are expected to be delivered by fifth-generation (5G) systems. In this shortened version of dissertation thesis, the most important of them are described and in the full version of dissertation thesis, they are covered in detail:

- Virtually unlimited capacity and ubiquitous coverage introducing the "anytime & anywhere" connectivity. Serious increase of network throughput (1 – 10 Gbps).
- High degree of flexibility & network intelligence of all involved technology components to deliver most of the services "on-demand" – with respect to meet agreed Quality of Service (QoS) and Quality of Experience (QoE).
- Significantly lower end-to-end latency (below 1 ms) to enable new application scenarios e.g., Tactile Internet (TI).
- Unrestricted mobility to enable the mobile broadband even for very fast moving objects (up to 500 km/h) e.g., controlled dynamic spectrum sharing at the airport.
- Energy efficient communication to reduce power consumption at the side of end users and telecommunication operators.

Despite very active research during last couple years resulting in a variety of promising solutions created across academia and industry as well, the true 5G landscape is still not there yet. However, main essential is already known - all technical and user requirements can be barely fulfilled by a single Radio Access Technology (RAT). Therefore, as fundamentally different to previous generations of cellular systems, the 5G networks will not be just an incremental advance of 4G, but rather constructed as a set of directly bounded communication technologies and protocols [7].

While in recent cellular systems, the selected wireless technologies have been developing and operating individually, the 5G needs significant increase in network capacity and throughput. Therefore, it requires a tighter interconnection and cooperation between different types of RATs. As a result, it becomes unavoidable to aggregate different radio technologies as part of a common converged radio network – to be transparent to the end users, and develop techniques that can efficiently utilize the radio resources available across different spectral bands [8]. Following this vision, the Heterogeneous Networks (HetNets) represent a key building block of next-generation 5G systems, where different RATs operating in licensed (LTE) as well as unlicensed spectrum (WiFi) are collectively providing the multiplied performance [9], [10].

With respect to the convergence of various RATs, the telecommunication operators increase the density of their mobile networks by deploying new cells with different coverage – to boost the overall network capacity [11] – since the multi-RAT concept together with continuing network densification are still not providing satisfactory outputs (from 5G perspective), especially due to limited space and narrow frequency bandwidth of all legacy wireless technologies.

Therefore, the heterogeneous deployments (infrastructures) have to be provided by novel wireless communication technologies – utilizing extremely high frequency millimeter-wave (mmWave) band ranging from 3 to 300 GHz [12]. Of course, the mentioned mmWave communications are naturally not suitable for long-range use-cases since the wave length can not infiltrate from dense materials efficiently. Therefore, it can be easily dispersed by rain drops, gases, and flora. Nevertheless, mmWave and Visible Light Communication (VLC) technologies can improve the transmission data rates for indoor setups, because they have come up with large bandwidth. This, in fact, particularly supports one of the key ideas of designing the 5G cellular architecture – the outdoor and indoor scenarios should be physically separated, so that penetration loss through building walls can be limited or even fully avoided [13].

2 MOBILE NETWORKS IN 5G ERA

It has been more than a few decades since mobile wireless communications were introduced with the first generation, voice-only mobile systems. Going further, over the last couple of decades, the society has witnessed gradual, yet steady evolution of mobile wireless communications towards second, third and fourth generation wireless networks. Introduction of digital modulations, effective reuse of frequencies, penetration of packet-based Internet communications and rapid advancement in physical layer technologies e.g., Wideband Code Division Multiple Access (WCDMA), Orthogonal Frequency-Division Multiple Access (OFDMA), MIMO, and Hybrid Automatic Repeat Request (HARQ) have played a significant role towards this gradual evolution [14], [15].

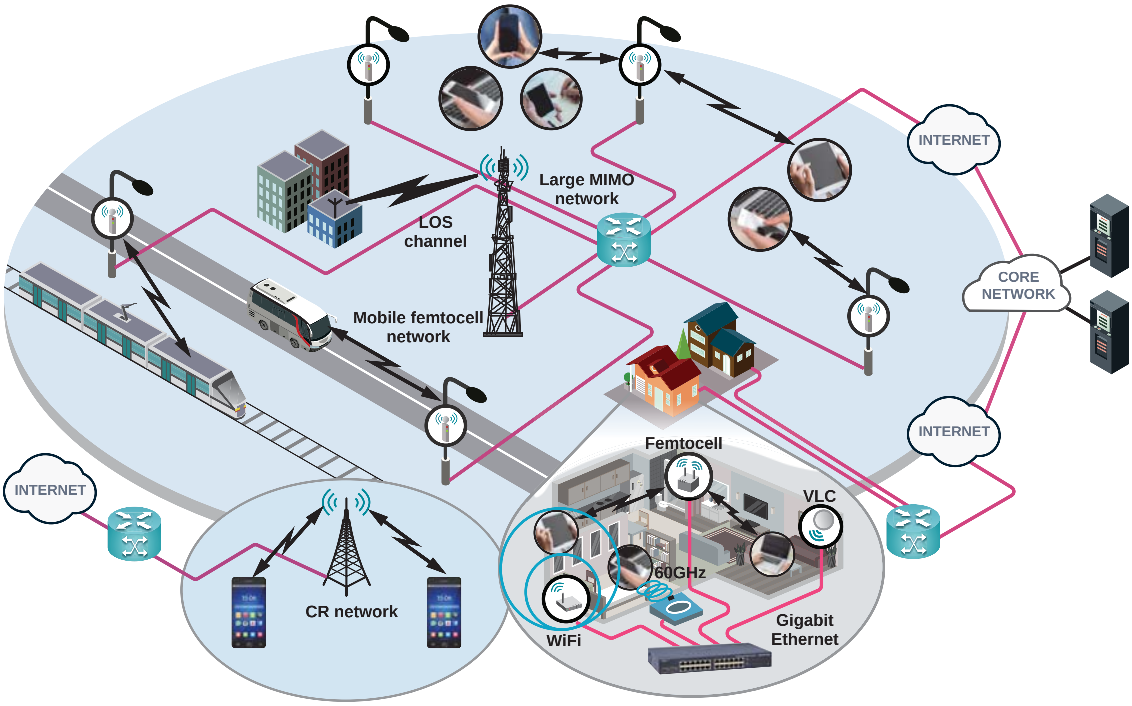


Fig. 2.1: Considered 5G heterogeneous wireless cellular architecture.

On top of this, with the ever increasing popularity of smart devices, currently all-IP based fourth generation (LTE) mobile networks have become a part of everyday life – providing a set of new, user-oriented mobile multimedia applications (e.g., mobile video conferencing, streaming video, and online gaming). These new applications are not only satisfying users' requirements (QoS, QoE), but also opening up new business horizons for telecommunication operators to increase their revenue, see Fig. 2.1 [6].

2.1 Existing Cellular Networks – Issues and Challenges

A quick look into recent wireless network statistics reveal that global mobile traffic grew 63 % in 2016 and almost half a billion (429 million) mobile devices and connections were added in 2016 [3]. Globally, smart devices represented 46 % of the total mobile devices and connections in 2016; they accounted for 89 % of the mobile data traffic. Another interesting finding is that smartphones represented only 45 % of total mobile devices and connections in 2016, but represented 81 % of total mobile traffic. Cisco's Visual Networking Index (VNI) forecasts that by 2021, nearly three-quarters of all devices connected to the mobile network will be "smart". In fact, since 2012 video traffic is more than half of the global mobile traffic [3]. An average mobile user is expected to download around 1 terabyte of data annually by 2020 [16]. Moreover, researchers are exploring new application use-cases in directions of Augmented / Virtual Reality (AR / VR), Internet of Things, Internet of Vehicles, Device-to-Device communications and Machine-to-Machine communications [1], [4] [14].

Tab. 2.1: Differences between human-based and M2M communications [17], [18], [19]

	Human-based	Machine-to-Machine
Traffic Direction	Mostly downlink; although uplink traffic is increasing over the last years due to interactive applications such as social networking, humans still download more than they upload.	Mainly uplink data to report sensed information. For some applications, symmetric uplink and downlink capacity is needed in order to allow for the dynamic interaction between sensors and actuators.
Message Size	The size of the messages is generally big, motivated by demanding applications such as multimedia and real-time transmissions, including video streaming.	The size of the messages is generally very short (e.g.very few bits of the reading of a meter, or even just 1 bit to inform of the existence or absence of a given event).
Connection and Access Delay	Human-based applications tend to be very demanding once a connection has been established. However, although not desirable, longer connection delays are typically well tolerated.	Many M2M applications will be based on duty-cycling, i.e., having devices sleeping and just waking up from time to time to transmit data. For some applications, the connection delays should be very short to ensure quick access to the network when waken up.
Transmission Periodicity	Human-based data traffic is very random and asynchronous in nature. In addition, the frequent transmission of control information is required to ensure high throughput and good delay performance.	Very wide range of alternatives. For many applications, transmissions will be very sparse in time. In addition, many applications will have known periodic patterns (e.g., programmed tasks).
Mobility	Mobility management and exchange of location information are constantly required to ensure seamless connectivity and allow for roaming.	For most of M2M applications, mobility is not a major concern. Some applications may have no mobility at all.
Information Priority	In general, there is no major differentiation between users in terms of priority, but only between applications for each user.	Some M2M applications may transmit critical information and thus require very high priority with a detailed level of granularity.
Number of Devices	At most, hundreds of devices per connection point. Typically, tens of devices per connection point.	Higher than in human-based communications. Hundreds or thousands of devices per connection point.
Security and Monitoring	Humans can raise an alert in the case of troubleshooting or tampering.	M2M devices cannot raise an alert in the case of malfunctioning or tampering.
Lifetime and Energy Efficiency	Humans can recharge batteries in a daily manner.	Once an M2M network has been deployed, some devices may require to operate for years or decades without maintenance.

Indeed, different research studies have been launched over the last years to understand, how cellular systems need to evolve to be able to provide efficient access to M2M / IoT networks. It is important to highlight that M2M principles are fundamentally different from humand-based (Human-to-Human; (H2H)) communications. These differences require a mentality shift on the way that cellular systems have been designed. A summary of the main differences between M2M and H2H traffic is shown in Table 2.1 [17], [18].

Based on the above, supporting this enormous and rapid increase of data usage and connectivity is an challenging task in present 4G (LTE) cellular systems. Furthermore, while currently deployed LTE networks were originally designed to support up to 600 Radio Resource Control (RRC)-connected users per cell [20], M2M communications and IoT requires supporting of tens of thousands of connected (smart) devices within a single cell. Therefore, LTE cellular network is exploring avenues of different research and development e.g., MIMO, HetNets, data offloading techniques and dynamic spectrum sharing to enhance network capacity and data rates [16].

2.2 Emerging 5G Use Cases and Applications

A wide variety of new emerging 5G applications put pressure on the commercial roll out of 5G wireless systems. 5G network architecture is expected to provide network solutions for a wide range of public and private sectors i.e., energy, agriculture, city management, health care, manufacturing and transport, with significantly improved user experience [21]. Aside from the enormous number of connections, 5G networks also have to support diverse nature of devices and their associated service requirements.

Although research and development in some of these applications are already underway in 4G wireless, original 4G LTE standards, 3GPP LTE Release 8 [22] did not include support to any of these applications. Rather, these applications were spawned later, and started explosive increase in wireless data usage, thereby imposing additional utilization of resource constrained 4G wireless networks. Naturally, later releases of 4G LTE networks, often named as "LTE Advanced", gradually started to include these applications. On the other hand, it is expected that massive bandwidth of 5G mmWave communications will provide a native support for these emerging applications. In this section, some of the demanding applications i.e., D2D communications, M2M communications, IoV, IoT and healthcare are discussed in detail.

2.3 5G Technology Enablers

In order to enable the ubiquitous connectivity required for many of the 5G-IoT applications, broad set of features and functionalities will need to be integrated to the currently predominantly broadband approach. This inherently leads us to a strong heterogeneous networking (HetNets) paradigm with multiple types of wireless access nodes (with different MAC/PHY, coverage, backhaul connectivity, QoS design parameters, etc.) [23], [24]. HetNets will offer the required seamless connectivity for the emerging IoT through a complex set of mechanisms for coordination and network management [25], [26]. Evolved 4G and emerging 5G networks will thus be characterized by interoperability and integration between various radio access networks, including those working with unlicensed frequencies. The aim of this section is to review recently finished 4G, currently ongoing 4G-Evolution as well as emerging 5G design efforts to provide an overall picture towards accommodating a heterogeneous networking [5].

2.3.1 5G Radio Access Network Enablers

To provide a common connected platform for a variety of applications and requirements for 5G, the main technology components will be described on following lines within this section [4], [27], [28], [29].

Massive MIMO

Massive MIMO represents an evolving technology that has been upgraded from the current MIMO technology. The massive MIMO system uses arrays of few hundred antennas which are at the same moment in one time / frequency slot serving tens of user terminals. The main objective of this technology is to extract all the benefits of "classic" MIMO but on a significantly larger scale. In general, massive MIMO is an evolving technology of next-generation (5G) networks, which is energy efficient, robust, secure and spectrum efficient [30].

The massive MIMO depends on spatial multiplexing, which further relies on the base station – to have channel state information, both on the uplink as well as on the downlink. In case of downlink channel, it is not easy, but in case of uplink, it is feasible, as the end-terminals send pilot waveforms. On the basis of pilots, the channel response of each terminal is estimated. In conventional MIMO systems, the base station sends the pilot waveforms to the terminals and based on these, the terminal estimates the channel, quantizes it and sends feedback them to the base station.

This process is not implementable for massive MIMO systems, especially in high mobility conditions because of two reasons: (i) the downlink pilots from the base station must be orthogonal among the antennas, due to which the requirement of time, frequency slots for the downlink pilots increases with the increase in the number of antennas; (ii) as the number of base station antennas increases the number of the channel estimates also increases for each terminal which in turn needed hundred times more uplink slots to feedback the channel responses to the base station. A general solution to this problem is to work in Time Division Duplex (TDD) mode and depend on the reciprocity between the uplink and downlink channels [31].

Therefore, massive MIMO technology depends on phase coherent signals from all the antennas at the base station, but the computational processing of these signals is feasible to do within the 5G-ready infrastructure.

Millimeter Wave Technologies

As the demand for capacity in mobile broadband communications increases dramatically every year [3], wireless carriers must be prepared to support up to a thousand-fold increase in total mobile traffic by 2020. This situation encourages researchers to seek greater capacity and to find new wireless spectrum beyond the 4G standard. Recent studies suggest that mmWave frequencies could be used to augment the currently saturated 700 MHz to 2.6 GHz radio spectrum bands for wireless communications.

Today's microwave cellular systems, which are originally targeted to different application use-cases i.e., voice / data / video calling, have precious little spectrum: around 600 MHz are currently in use, divided among operators. Currently, there exist two ways to gain access to more microwave spectrum [32]:

- Repurpose or refarm radio spectrum – this option has occurred worldwide with the repurposing of terrestrial TV spectrum for applications such as rural broadband access. Unfortunately, refarming has not freed up that much spectrum, only about 80 MHz and at a high cost associated with moving the incumbents.
- To share spectrum utilizing e.g., cognitive radio techniques – the high hopes initially placed on cognitive radio have been dampened by the fact that an incumbent not fully willing to cooperate is a major obstacle to spectrum efficiency for secondary users.

Altogether, it appears that doubling the current cellular bandwidth is the best case scenario at microwave frequencies. Alternatively, there is an enormous amount of spectrum at mmWave frequencies ranging from 3 GHz to 300 GHz. Many bands therein seem promising, including most recently the local multipoint distribution service at 28–30 GHz, the license-free band at 60 GHz, and the E-band at 71–76 GHz, 81–86 GHz, and 92–95 GHz. Foreseeably, several tens of gigahertz could become available for 5G. Needless to say, work needs to be done on spectrum policy to render these bands available for mobile (cellular) communication. Signal propagation is not an insurmountable challenge. Recent measurements indicate similar general characteristics as at microwave frequencies, including distance-dependent path loss and the possibility of non-line-of-sight (NLOS) communication. A main difference between microwave and mmWave frequencies is the sensitivity to blockages – the information in [33], for instance, indicate a path loss exponent of two for line-of-sight (LOS) propagation but four (plus additional power loss) for non-line-of-sight. Therefore, mmWave cellular research will need to incorporate sensitivity to blockages and more complex channel models in the analysis, and also study the effects of enablers such as higher density infrastructure and relays. Another enabler is the separation between control and data planes.

Antenna arrays are a key feature in mmWave systems. Large arrays can be used to keep the antenna aperture constant, eliminating the frequency dependence of path loss relative to omnidirectional antennas (when utilized at one side of the link) and providing a net array gain to counter the larger thermal noise bandwidth (when utilized at both sides of the link). Adaptive arrays with narrow beams also reduce the impact of interference, meaning that mmWave systems could more often operate in noise-limited rather than interference-limited conditions. Since meaningful communication might only happen under sufficient array gain, new random access protocols are needed to work when transmitters can only emit in certain directions and receivers can only receive from certain directions. Adaptive array processing algorithms are required that can adapt quickly when beams are blocked by people or some device antennas become obscured by the user's own body [33].

Direct Communications

D2D communications represent a turning point in cellular systems. They entail the possibility that two devices can exchange data without the involvement of the BS or with just a partial aid from the base station [34]. As the basic principles of D2D communication were described in Section 2.2.1 (in the full version of dissertation thesis), this section deals with the description tightly connected with the offloading scenarios – future application use-cases within 5G systems. In contrast to WiFi / WiFi-Direct and Bluetooth technologies, which provide D2D capabilities in the unlicensed band, with D2D communications the QoS and QoE are controllable because of the use of the licensed spectrum. A new generation of scenarios and services in 5G wireless systems can hence be enabled, including: (i) device relaying, (ii) context-aware services, (iii) mobile cloud computing, (iv) offload strategies, and (v) disaster recovery. Based on above text, four different types of D2D communications can be distinguished [35]:

- Device relaying with telecommunication operator controlled link establishment – any device can broaden the coverage of the BS, by acting as a relay node.
- Direct D2D communication with operator controlled link establishment – any pair of network nodes can directly interact due to a D2D link, which is set up under the control of the telecommunication operator.
- Device relaying with device controlled link establishment – the endpoints of a data session are in charge of setting up a relaying infrastructure made of one or more relaying devices.
- Direct D2D communication with device controlled link establishment – any pair of devices can exchange messages thanks to a D2D link, which is established without any operator control.

Enabling D2D communication use-cases in the licensed band is essential to strengthening the support to mission critical applications and group communications. Besides these advantages of D2D communications, security, trust, interference management, resource discovery, and pricing issues should be addressed to capitalize the potential of this promising technology. These issues become very challenging when the D2D link is set up without any involvement of the base station. Moreover, new business models are required to answer the “pay for what” question. In fact, devices that act as relays will deplete their own resources (as battery, storage, communication, and processing) to assist the D2D model. Further, the cross-operator D2D capability is an open challenge complicated by the fact that frequency division duplex (FDD) spectrum bands are different for different telecommunication operators [34], [36], [35].

The RATs’ principles (illustrated in first half of this section), will need to be supported by a suitable RAN technologies. Therefore, interesting propositions, which have been made with respect to the RAN, are discussed within the scope of remaining paragraphs of this section.

Licensed Assisted Access

Recently, the 3GPP commenced a work item on License Assisted Access (LAA), where licensed and unlicensed carriers are aggregated. LAA uses licensed frequency spectrum for control-related transmissions while sending data over both licensed and license-exempt carriers. Whilst mainly designed for high-capacity applications, the approach could be beneficial in the context of a ever-increasing amount of smart devices with increasing data rate demands. Notably, all non-critical IoT traffic could be transmitted via the license-exempt band whilst being controlled from the licensed band [37], [38].

Licensed Shared Access

In light of the above, it appears that the shared use of spectrum becomes unavoidable even for those who have conventionally enjoyed exclusive access rights [39]. However, the currently existing options of spectrum sharing (in primary licensed or unlicensed spectrum) do not offer much of the requested interference protection, therefore resulting in insufficient reliability, QoS guarantees, and predictability of operation. By contrast, the emerging Licensed Shared Access (LSA) regulatory concept allows for more advanced spectrum sharing between a limited number of entities with carefully defined usage rights, combining the benefits of command-and-control spectrum management with a flexible and innovative market-friendly approach.

Broadly, LSA approach (framework) enables authorized spectrum sharing by allowing at least two users, the incumbent (i.e., the current holder of spectrum rights) and the LSA licensee (i.e., the temporary user of spectrum) to access the same frequency bands in a licensed predetermined manner following a well defined mutual agreement [40]. In other words, LSA guarantees that the incumbent retains spectrum access rights anytime, anywhere, and the LSA licensee(s) will refrain from using this spectrum when needed by the incumbent (or at least will not disrupt the incumbent's operation).

2.3.2 5G Core Network Enablers

Finally, this section discusses the emerging network enablers which are pertinent to supporting the above 5G RAN enablers. Traditionally, core networks have been designed as a single network architecture serving multiple purposes, addressing a range of requirements, and supporting backward compatibility and interoperability. This "one-size-fits-all" approach has kept costs at a reasonable level, given that one set of vertically integrated nodes has provided all functionality. Technology has, however, evolved.

Virtualization, advanced automation and orchestration make it possible to build networks in a more (i) scalable, (ii) flexible, and (iii) dynamic way. From several possible enablers, the focus in this section is given on Software Defined Networking (SDN) and Network Function Virtualization (NFV).

Software Defined Networking

The SDN paradigm initially designed for wired networks (e.g., data centers), has recently gained a lot of attention into the wireless systems [41], and it is seen as a key technology enabler for 5G networks [42], [43]. SDN separates the data plane (i.e., the traffic forwarding between network devices, such as switches, routers, end terminals) from the control plane (i.e., the decision making about the routing of traffic flow) [44]. SDN merges network control into a logical entity, namely SDN controller, and allows programmability of the network, by external applications. With its centralized view of the network i.e., topology, active flows, etc., SDN provides dynamic, flexible, and automated reconfiguration of the network. Moreover, SDN will be able to address flexibility and interoperability challenges of future multi-vendor, multi-tenant 5G scenarios. In fact, with SDN it will be possible to deploy a vendor-independent service delivery platform, able to proactively respond to the changing business, end-users and market needs.

Therefore, SDN will simplify network design, management and maintenance in heterogeneous networked environments [45]. With the explosion of devices connected within the 5G infrastructure, traditional network architectures will not be able to manage both the volume of devices, and the amount of data they will be dumping into the network. There will be need to effectively manage the load of traffic and the network resources in the 5G era, to avoid possible collapse of the network, and allow the coexistence of different services with different QoS and QoE requirements [46].

Network Function Virtualization

The NFV stands for a complementary technology of SDN, destined to impact future 5G networks. Nowadays, NFV aims to virtualize a set of network functions, by deploying them into software packages, which can be assembled and chained to create the same services provided by legacy networks [47]. The NFV concept comes from the classical service, whereby many virtual machines running different operating systems, software and processors, can be installed on the same server. By moving network functions from dedicated hardware into general purpose computing/storage platforms (e.g., servers), NFV technologies will allow to manage many heterogeneous 5G / IoT-ready devices. Moreover, by implementing the network functions in software packages that can be deployed in virtualized infrustructure, NFV offers scalability and large flexibility in operating and managing mobile devices.

With NFV, it will be possible to reduce both CAPEX and OPEX. Currently, the use of NFV is under discussion in the context of virtualizing the core network, and centralizing the base band processing within RAN [48], [49].

3 PROXIMITY-BASED (CELLULAR-ASSISTED) NETWORKING

While telecommunication operators have finally started to deploy fourth generation broadband technology, many believe it will still be insufficient to meet the anticipated demand in mobile traffic over the coming years [3]. Generally, the natural way to cope with traffic acceleration is to reduce cell size, and this can be done in many ways. The most obvious method is via cells' densification, but this requires additional Capital Expenditures (CAPEX) and Operating Expenses (OPEX) to install and manage these new base stations. On the other hand, another approach, which avoids this additional CAPEX / OPEX, involves offloading cellular traffic onto direct D2D connections whenever the users involved are in proximity. Given that, most end devices are capable of establishing concurrent cellular and WiFi connections today, we expect the majority of immediate gains from this approach to come from the use of the unlicensed bands.

However, despite its huge commercial success, WiFi-based direct connectivity may suffer from stringent session continuity limitations, excessive user contention, and cumbersome manual setup / security procedures. In this chapter, the proposed, implemented and successfully tested integration of managed D2D communication links into current cellular technology is comprehensively described. Provided outcomes are based on a simulation tools and especially on showcasing the proposed principles in real 3GPP LTE-A wireless system located at Brno University of Technology, Czech Republic.

3.1 Experimental Evaluation of 3GPP LTE-Assisted D2D Communication

Today's cellular network deployments can be considered as nearly ubiquitous – they provide a wide range of advanced services and are constantly evolving to serve the increasing subscriber populations. However, their ability to meet user expectations (QoS and QoE) is at risk given the constant influx of new customers and the increasing numbers of multimedia service requests [3]. Another problem for network users is the limited battery budget of their mobile devices; this becomes a serious limitation for many throughput-hungry applications. Both of these problems require significant changes in the way we approach wireless content delivery.

Traditional solutions primarily focus on increasing density of cell deployment. Owing to shorter radio links, smaller cells provide higher bit rates and require less energy for uplink transmission, especially in dense urban environments [50]. However, deploying larger numbers of smaller cells can become prohibitively expensive, and the complexity of interference management may increase much over what is desired.

In contrast, proximity-based communication between client devices enables shorter radio links without the cost of additional infrastructure [51], [52]. With such shorter and lower to the ground links, their interference is easier to manage when compared to standard infrastructure small cells. Thus, in this section, it is proposed that whenever possible, client devices use their direct connectivity capabilities, instead of cellular links. In particular, the corresponding benefits of WiFi-Direct technology [53] are highlighted, since it is already available on most client devices and does not reuse the more expensive cellular bands. With respect to this, the use of infrastructure network-assisted device discovery (by example of 3GPP LTE), which expedites the discovery process resulting in greater success probability and improved energy efficiency [51], is explored.

With this section, it is demonstrated how cellular traffic can be effectively offloaded onto WiFi-Direct D2D links and provide the estimated gains in energy efficiency and capacity from such offloading. According to existing research, certain applications, in particular Peer-to-Peer (P2P) and socially-oriented ones, could utilize such links effectively [54], [55]. This approach is based on developed advanced System-Level Simulation (SLS) toolkit, capable of capturing dynamic multi-radio scenarios, with realistic user mobility and traffic arrival patterns. It is shown, how the shorter range and simpler protocol of network-assisted D2D communications allow users to benefit from higher data rates over short distances without compromising their battery life – in this particular research, one of the forms of network assistance, when the network informs the clients only about potential D2D partners, is utilized [52].

3.1.1 LTE-Assisted WFD D2D System Implementation

With finished preliminary research [50], [56], [57], [58], [59], we have identified numerous benefits that could become available if a coordinated, network-assisted D2D technology is deployed by network operators. On the other hand, introducing D2D technology within today's network infrastructure poses a number of challenges and requires updates to the current longstanding cellular architecture. Therefore, to conduct a comprehensive study and reveal the practical promises of D2D communications, we have designed a trial development and deployment program.

Our trial was aimed at demonstrating how the direct connectivity paradigm could be seamlessly integrated into a real-world, operator-grade cellular network with minimal modifications and overheads, as well as within a reasonable time frame. Our secondary goal was to quantify gains that could be achieved by a fully-functional, operator-supported D2D system. As a equipment basis for our trial, we have utilized the experimental LTE-A network of Brno University of Technology (BUT), Czech Republic, which supports most of the functionality expected of LTE Release 10 systems. During the implementation phase, we have upgraded the LTE network of BUT with our own implementation of Proximity Services (ProSe) functionality as envisioned by the 3GPP specifications [60], [61]. This has allowed us to perform live D2D integration trials, along with corresponding performance evaluations.

D2D Standardization Activities

The key enabling technologies for D2D communications have been around for years, since practically any IP-ready mobile device is tentatively capable of direct connectivity. However, the development of the necessary supporting standards and user interfaces has extremely lagged behind any developments in related technologies, with the first D2D activity started by 3GPP in 2012 within the framework of ProSe. 3GPP TR 22.803 [60] is known as the initial document to specify what exactly is to be understood by the term "proximity-based services" and D2D naturally fits into this category. However, 3GPP ProSe specifications and associated work items cover, in fact, a much broader area. They address the phenomena in both society and economy that drive the need for proximity-based communications, including such examples as social networking; disaster relief and emergency service operations; advertising; and so on [62], [63]. As a result of this work, the decision was made that ProSe functionality has become a part of future LTE releases.

Consequently, the technical side of providing ProSe mechanisms demanded attention. In particular, a requirement has emerged on establishing proximity in an efficient way without revealing personal user information. Indeed, proactively requesting any content via short-range radio links discloses the type of the desired content, whereas broadcasting content advertisements discloses what is available to share. Either way, users of a proximity-based discovery service, both content providers and consumers, may prefer to remain anonymous or impose their customized policies on a set of targeted peers. With decentralized solutions, such as those offered by WFD, meeting the aforementioned requirements is clearly impossible. This, in turn, has led to research on a potential Evolved Packet Core (EPC)-level discovery procedure, where a 3GPP network would act as a trusted intermediary and implement all of the necessary policies on behalf of users. The report TR 33.833 [64] quantifies specific goals, which have been respectively targeted by 3GPP.

Currently, since the overall vision of how the ProSe function is to be implemented has already taken shape, follow-up work has begun on supporting infrastructure, such as billing [65]. As of 2015, most of the architectural progress on ProSe and D2D communications is summarized in the TS 23.303 document [66]. It is likely, however, that additional amendments will be made when the activities on the technical side of unlicensed LTE commence [67]. Today, some of the related ideas on potential license-assisted communication (RP-140770) and on LTE-based short-range radio within licensed bands [68] have already been documented.

However, aside from their in-house efforts on short-range radio, 3GPP also supports alternative non-3GPP radio technologies, including WFD, for ProSe radio communication. The integration between 3GPP and WiFi solutions has been a long-standing effort, with specifications TS 23.234 [69], TS 23.327 [70], and TS 23.402 [71] outlining how LTE devices that are connected over non-3GPP access technology, could still receive access to all of the 3GPP services. Following this lengthy integration effort, all current ProSe architectures support the use of IEEE 802.11 family, as a link layer for most of the ProSe functions, with the exception of public safety services [66].

Architectural Considerations

Envisioned implementation of the generic D2D system concept offered by the standards has naturally met a number of deployment challenges that made us deviate from the reference solutions. Current architectural considerations of LTE networks preclude us from deploying the network assistance functions in the way that would have been the most “natural” from an engineering standpoint; thus, within the novel implementation introduced in this work, it was necessary to adopt several clever workarounds to develop a workable system with today’s technology.

In what follows, the key decisions made in each step of the deployment process are discussed. The experimental cellular network setup installed at the Department of Telecommunications, BUT (see Fig. 3.1) stands for a complete commercial-grade implementation of all the crucial subsystems comprising contemporary 4G mobile networks.

Utilized trial deployment is configured to provide the necessary packet-switched data access services and their derivatives, such as VoIP communications over converged LTE and WiFi radio access infrastructure. The EPC is dimensioned to enable high data rate services with appropriate QoS and QoE provisions, as well as to support up to 100000 concurrently served users. For voice and video calls, the switching capabilities are implemented by employing the high capacity IP Multimedia Subsystem and its related components, for both mobile and fixed access users, as well as for connectivity to external telephone networks and teleconferencing systems. The general goal behind BUT’s LTE test network deployment has been achieving the synergy of a complete and customizable experimental mobile network that allows for rapid implementation and prototyping of novel concepts and technologies, such as the D2D communications paradigm discussed at length in this section. We thus effectively used this asset to showcase that the LTE-assisted WFD technology has matured enough for an example implementation in a commercial-grade LTE/WiFi network.

However, several LTE mechanisms expected by the ultimate network-assisted D2D architecture were not available at the time of our early implementation. This included the evolved Serving Mobile Location Centre (eSMLC) server, for which an alternative interface to obtain device location information directly from a Mobile Management Entity (MME) has been developed as a substitute. In addition, the D2D server functionality has been implemented as a virtualized appliance, whereas the final commercial operator-grade implementation should preferably use a more robust technology. Further design choices are explained below [50], [57], [58].

Impact of Radio Access Network on D2D Communications

A radio access network (RAN) is crucial for any kind of service in a mobile network. To this end, RAN is typically required to transport signaling and data. That being said, D2D communications are not particularly demanding with respect to RAN capacity, as D2D messages are only few bytes in size, and ultimately D2D communications reduces the RAN load. The D2D signaling should, however, be prioritized to decrease service response times. Correspondingly, the impact of RAN loading on the observed service times has been measured extensively in performed experiments. Other RAN functionality required for LTE-assisted D2D communications is positioning. Providing user location information is achieved via cooperation between eNBs and other EPC elements, and is described thoroughly at length of following sections.

Notably, neither special configuration nor non-standard service is needed from the RAN side for the D2D system to function. A deployed RAN is part of our test-bed installation. Generally, commercial-grade equipment

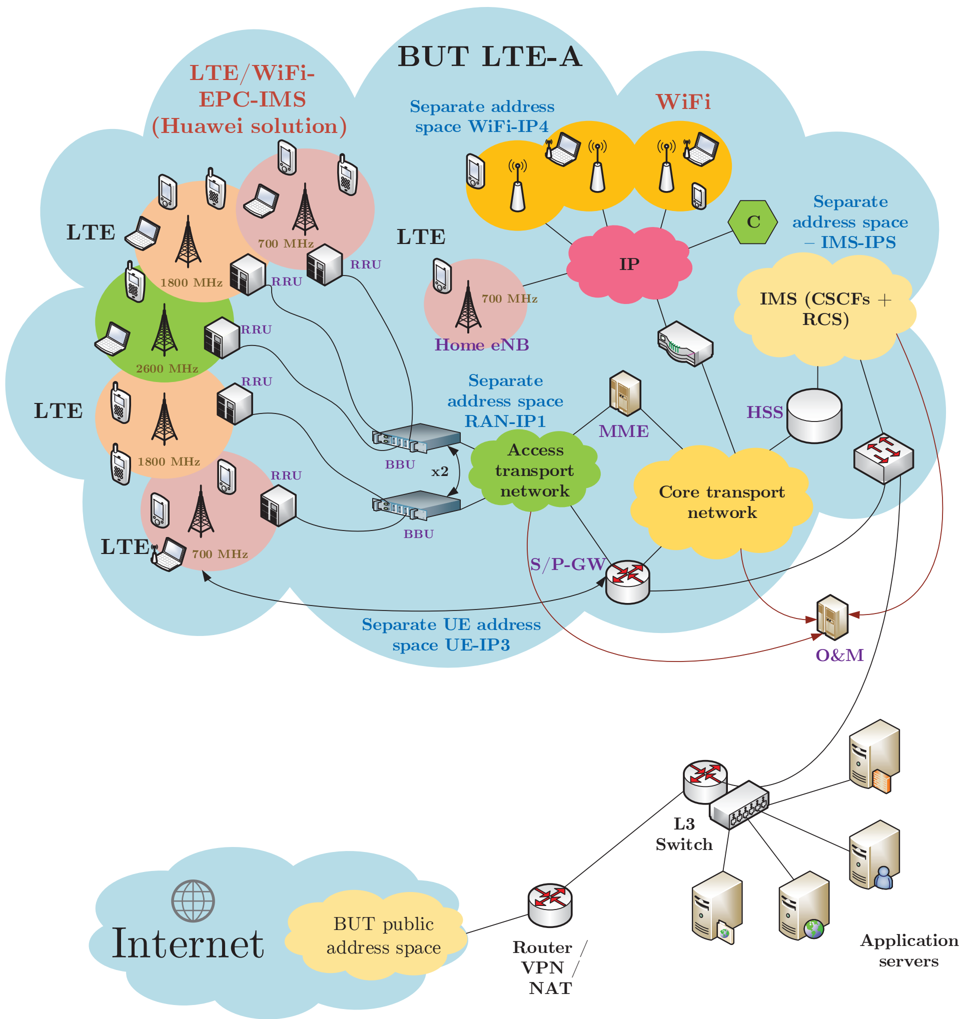


Fig. 3.1: System topology of experimental 4G cellular network deployed at BUT, Czech Republic [50], [72].

is used, but the RAN itself is not a part of the radio access subsystem of a mobile network operator. Hence, we have complete control over the RAN loading conditions. Finally, as our RAN solution is deployed as real application, external interference is present.

Providing Unrestricted IP Connectivity

One of the key requirements for any direct user connectivity is the ability to communicate between devices without any intermediary hosts operating at the transport layer or above. Typically, in cellular networks, the associated devices acquire their IP addresses from private ranges (for example, 192.168.0.0/16). As long as a D2D link is established within a single operator's network, this does not impose any constraints. However, the effective firewall policies deployed in the core network, which deny direct access between user devices to enhance security, actually cause difficulties. The original purpose of the aforementioned firewall policies could have been prevention of undesired incoming connections and P2P data transfer between different mobile network users. As a result, they deny any P2P connectivity over cellular networks [73].

For the purposes of our network-assisted D2D trial, we needed to circumvent the firewall by making the D2D server capable to open direct communication paths for selected connections whenever necessary. In our case, implementing such functionality has been fairly straightforward, since the D2D server can reconfigure the firewall on a per-connection basis. In our implementation, the firewall is located inside the Unified Gateway (UGW) entity; logically composed of a Serving Gateway (SGW) and Packet Data Network Gateway (PGW).

Communication between D2D Server and Users

By design, D2D network assistance relies on a network's ability to communicate with those user devices that are engaged in direct connectivity with the network. This inherent ability must be augmented by an efficient means of initiating such communications. For example, the said connectivity could be straightforwardly enabled with a session initiation protocol, but this would require an active radio bearer. Having an active LTE radio bearer for only several packets is naturally not conducive to system efficiency; hence, this is to be avoided. In LTE, there are multiple ways to transfer short messages between a network and those individual users of the network that do not require a dedicated bearer setup, such as non-access stratum (NAS) signaling, which is typically used to set up the bearers themselves.

As a result, the practical deployment of the D2D functionality requires an implementation, where the D2D server would be positioned outside of the system core as a conventional IP service, but with a capability to access certain core network functions (see Fig. 3.2). While this may not be the optimal solution in final commercial deployments, it enabled us to move forward promptly with regard to D2D system implementation. Although the proposed location of the D2D server does not follow the 3GPP guidelines exactly, we believe that our modification does not produce any negative impact with respect to latency, as the connection between the SMLC and the D2D server is implemented via a tunnel over a fiber channel.

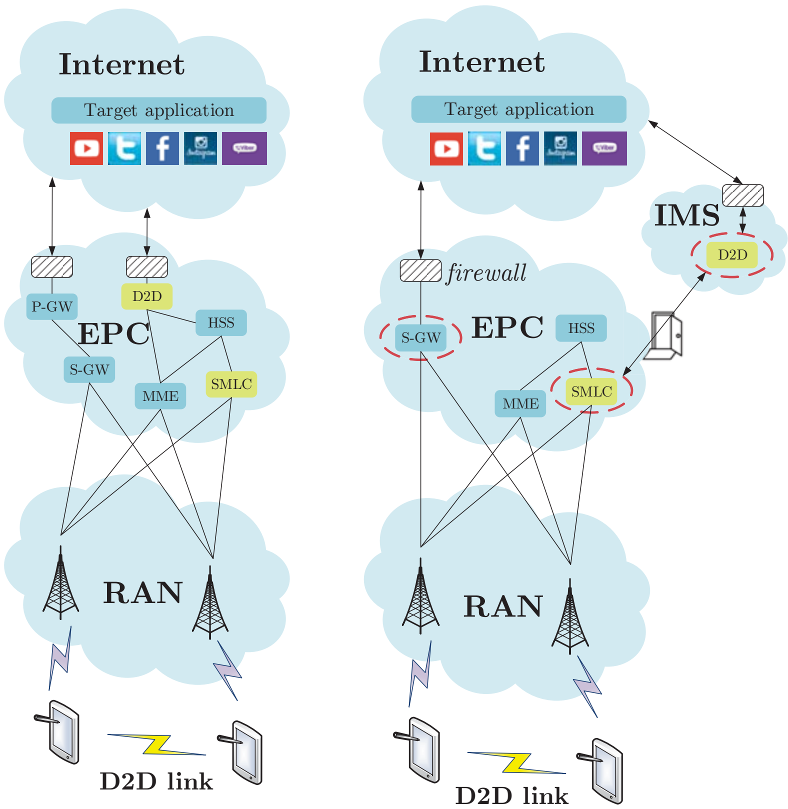


Fig. 3.2: Comparison between (a) hypothetical/standardized and (b) our deployed implementations of D2D system in 4G mobile networks [50].

Integrating Location Services

For the network-assisted direct connectivity to operate efficiently, the D2D server needs regular updates on the current locations of users. In LTE, such information is conventionally aggregated by the eSMLC entity to be then made available for the user devices via Secure User Plane Location (SUPL) bearers. A copy of the location information in question is typically stored inside the MME for its internal usage. Whereas the exact means of how this information is obtained may vary, the general procedure is such that a phone's GPS location will be used if available. Alternatively, the positioning reference symbols within the LTE frame will be used to pinpoint the location of user equipment (UE) through triangulation.

Either way, the coordinates are obtained in line with the standard techniques outlined by 3GPP and are enabled in most modern equipment (post Release 9). As a result, enabling location-based services is not a major challenge in contemporary LTE networks, and most of the time such functionality is already provided by the operators for the mobile devices to use.

In the considered proximal scenario, the D2D server accesses the location information on behalf of the UE, and then draws conclusions on whether other UEs are sufficiently close to initiate direct communication. An example of such decision logic is presented in Fig. 3.3. With the help from location services, the UE can thus power on its radio only when the intended contact is in proximity, hence saving battery and network resources. The specific signaling used in the trial is further discussed in following sections. Naturally, one would need to select the thresholds triggering various decisions, but those are largely hardware-specific and have not been the core subject of this research work.

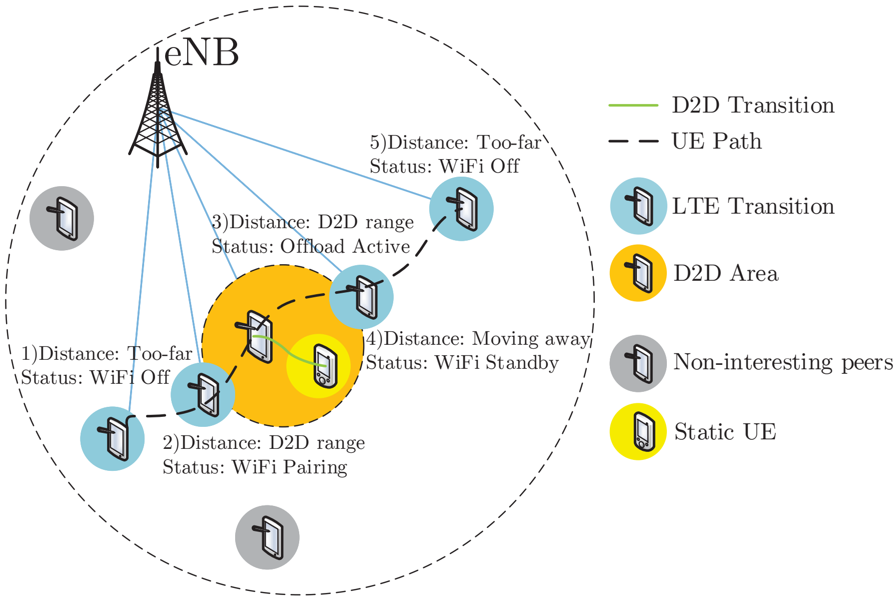


Fig. 3.3: Switching logic between LTE and proximal D2D connections, as mobile UE is moving along its route and meeting static D2D peer [50].

Linking D2D Server and Core Network Components

Recall that the location information is typically made available in LTE only for MME and the UEs themselves. Similarly, the interfaces to control the firewall policies at the UGW are also core-local. In our implementation, we had to create a secure connection to the core to allow the D2D server to communicate with the MME and UGW, as to extract the location information and configure firewall policies via the maintenance interfaces, which is, of course, not the preferred final solution to go with.

In reality, one would require enabling the D2D server to connect to the MME/SMLC and UGW via an efficient and secure application programming interface (API), without exposing the entire management console. Due to the time limitations imposed by the purposes of this trial and to the vendor-specific nature of the firmware, we have left the development of such API for future work.

D2D Connection Control

Efficient control of D2D connectivity between UEs deserves a separate dedicated research, which can be generally split into signaling and executive components. The signaling may be handled by a service that is running as a service on the mobile device, whereas the executive part is essentially integrated into the kernel drivers of the operating system (OS) of the UE.

While implementation of the signaling part as an application is rather trivial for most of today's mobile phones, actually forcing control over WiFi and cellular connections (together with the associated routing table) requires significant modifications to the permission levels of the UE. In practice, for most platforms (Android, the majority of Linux-based systems, and so on), this means employing custom-built firmware for the phone, or obtaining the administrative privileges by virtue of hacks and exploits. For closed platforms, such as iOS and Windows Phone, these solutions are nearly impossible without cooperation from the platform vendors.

Performance Evaluation of Implemented D2D System

The primary goals of intended performance evaluation are as follows:

- Indicating bottlenecks that could potentially hinder the adoption of D2D connectivity in future wireless technologies.
- Establishing appropriate performance bounds and limitations for D2D technology, as well as outlining what services could be most suited for direct communications in contemporary and near-future markets.

To achieve these diverse goals, we follow the measurement procedures as outlined in the remainder of this section.

Measurement Methodology It is important to note that in this work (by contrast to numerous past publications) we are not interested in the performance of the D2D link itself, since that would largely depend on the current channel and user contention levels. In this study, we specificaly concentrate on assessing signaling performance and network assistance logic, as the latter can be reliably measured in our controlled trial environment. On the other hand, such metrics as D2D throughput are extremely difficult to assess conclusively in practice due to the high variability in wireless environments.

Based on the above reasoning, the single most important parameter of D2D signaling is the connection setup time. This latency is crucial, as lengthy connection setup times may delay the transfer of the flow to an alternative radio link, thus affecting other parameters such as energy efficiency and user experience. Due to small message sizes, in the order of a few tens of bytes, other QoS requirements on the D2D signaling are easy to fulfill and can be provided by any access network; therefore, we are not considering these as key performance indicators for this technology. Other aspects of D2D link performance, which are not directly related to mobile network infrastructure, do not affect service setup response times and, as such, are out of the scope of this work. The connection setup time may in turn be decomposed into several components, as described below.

First, by looking at the proposed D2D protocol signaling illustrated in Fig. 3.4, we learn that before any WFD connection is actually set up, there are several important actions to be taken from the network side to observe user proximity. Therefore, we decomposed the signaling procedure in question into several distinct stages. In stage 1, the responsiveness of the system is not of particular interest, since the procedures taking effect there are constrained primarily by the human user input.

However, once stage 1 is completed and the proximity is established, the actual D2D signaling is triggered. Ultimately, we have configured for stages 2 (D2D link negotiation), 3 (link setup), and 4 (actual flow switching) to occur as quickly as possible to maximize the effective time during which we can take advantage of the D2D link, especially if the communicating peers are highly mobile. These stages also constitute natural measurement checkpoints, where the resultant system agility would straightforwardly depend on how long it takes to transit from one checkpoint to another. Therefore, we further commit to measure an aggregate latency at each of the above state transitions. To provide the best available accuracy, we have additionally decomposed the considered D2D protocol into individual messages and performed our measurements on a statistically large sample set. The latency has been measured between the client device (laptop Lenovo T430 with the USB LTE modem Huawei E392) and the D2D server (located behind the UGW). Utilizing the modem E392, the laptop has been connected

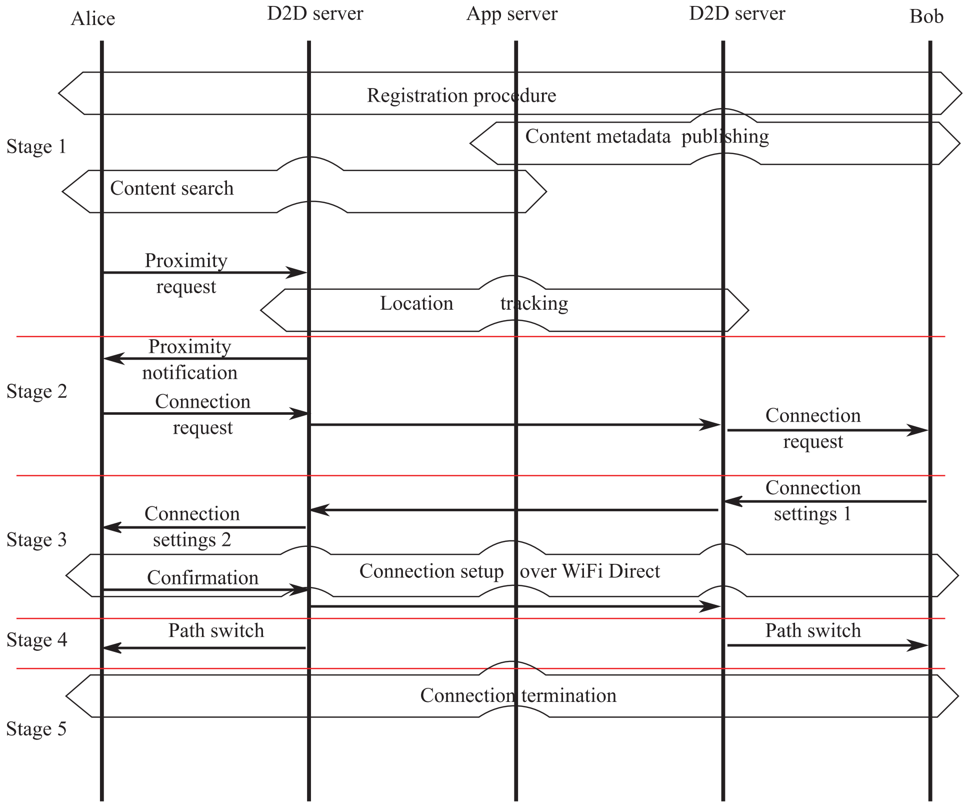


Fig. 3.4: Considered D2D protocol signaling including all five stages required for proximity discovery, D2D link establishment, and termination [50].

to our experimental BUT LTE network. The distance between the client and the eNodeB has been fixed to 5 meters throughout the entire performance evaluation campaign.

As a measurement tool, Iperf [74] has been employed to assess the maximum throughput of a particular radio access solution. Later, Iperf has also been used to create predefined loading on both uplink and downlink channels to thus emulate specific conditions on a "deployed" network. The client has been configured in the role of a sink (receiving data traffic from the D2D server) and a generator (generating the constant User Datagram Protocol (UDP) data bit stream for Iperf, which runs on the D2D server). The WFD D2D link setup times have been measured between two Sony Xperia ZL phones, running custom Cyanogenmod 10.2 aftermarket firmware. The link setup time has been assessed following the known-channel pre-shared key (PSK) authentication procedure (essentially, employing the WiFi Protected Access II (WPA2) protocol). The WPA2 authentication enables legacy devices (which do not support WFD extensions) to still join the networks initiated by other devices. The measurement for the connection setup time has been done with the *wpa_cli* interface, by monitoring the time between when a new network has been enabled and when a new connection has been completed. A connection was considered completed once the first IP packet was successfully acknowledged.

Numerical Results and Discussion Following the above methodology, the measured values of the LTE network latency have been obtained to compose Table 3.1, which summarizes the estimated delays introduced by various connection setup stages. Here, the stage 3 delay includes Wi-Fi link setup, as discussed further on. The fractional loads on WiFi and LTE networks have been matched. For LTE, our results indicate that if the overall loading on the LTE network does not exceed 90%, the roundtrip times (RTTs) between the client and the D2D server are generally below 30 ms. Hence, cell loads of up to 90% do not have any evident effect on the LTE signaling procedure when compared with the WiFi link setup time. Under higher loads, the WiFi setup time becomes comparable with the LTE latencies, thus making LTE signaling optimization a concern. This is especially critical since offloading is most needed under high loads.

Tab. 3.1: Network latency measurements for different cell loads

Cell and WiFi load	Idle (10 %)	50 %	90 %	99 %
Measured RTT (ms)	18	25	27	60
Stage 2 (ms)	36	50	54	120
Stage 3 (ms)	750+36	850+50	1000+54	1150+120
Stage 4 (ms)	9	13	13	30
Full procedure (ms)	849	988	1148	1480

In the case of 99% load, the network queues have been overfilled; we can assume this particular scenario to be highly relevant for an offloading operation. The resulting LTE RTT values range in the order of half a second, and without the appropriate QoS support in LTE the obtained latency values would dramatically impact the overall performance of the considered D2D technology.

Further, Fig. 3.5 clarifies the delay values for the different types of cell load. Whereas the values remain negligible for up to 90% loading, the "full-buffer" condition yields a major increase in the monitored delay, which becomes hardly acceptable (in the range of seconds). Therefore, we paid close attention to the measurements in the range 91% to 99% of cell loading. In this extreme case, the implementation of QoS on the UE side is crucial to prioritize the D2D signaling messages, see Fig. 3.6.

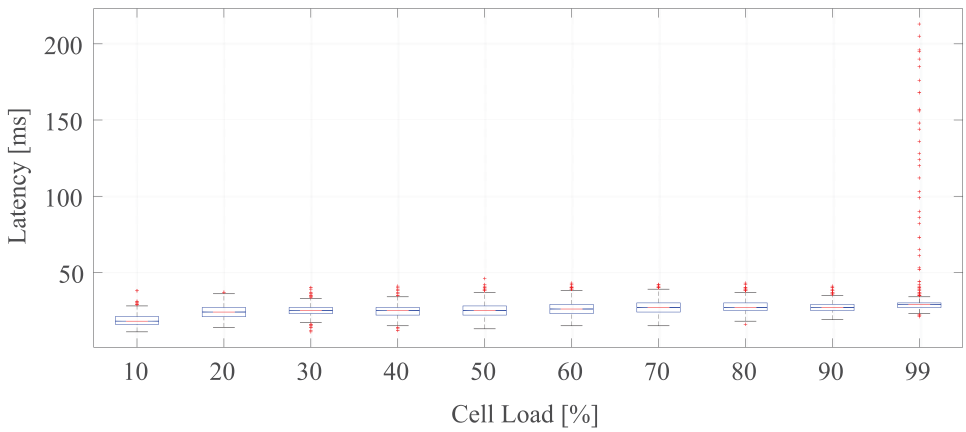


Fig. 3.5: Network latency for alternative LTE cell loads [50].

It is evident that the network conditions remain reasonably stable for up to an extremely loaded state (corresponding to a cell load of 99%).

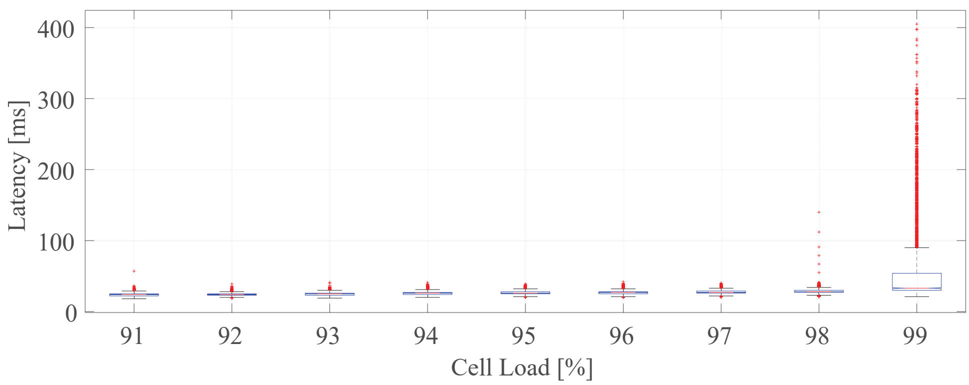


Fig. 3.6: Network latency for extreme LTE cell loads [50].

Under extreme loads, the average latencies remain largely acceptable; however, due to the high variance, the perceived quality of user experience suffers significantly. Interestingly, out of all the D2D connection stages, by far the most costly ones are connection negotiation and the actual setup of the link. In practice, those require several RTTs to complete and may add up to several seconds of signaling time in highly congested cells.

Naturally, this necessitates further improvements to how the signaling messages are carried, since the primary goal of D2D is to actually reduce the load on the cellular links; moreover, the longer it takes to perform D2D offloading, the worse the resultant congestion levels will be. It is noteworthy that the D2D-link tear-down procedure at stage 5 is not instant. However, since the direct links are removed only when they are not needed anymore, this tear-down procedure is not as time-critical. Hence, it is not evaluated explicitly by our present study.

Performance Summary

Based on the performance numbers obtained above, one can conclude the following on the current bottlenecks of the LTE-assisted WFD implementation:

- The WFD link setup time could be dramatically shortened by removing the OS-introduced delays. While this does not significantly affect user experience, it does delay the actual start of offloading. Under most conditions, this overhead is nearly constant.
- The D2D connection negotiation stage is the primary bottleneck in an idle LTE system, as it involves multiple RTT cycles to complete, and can grow exceptionally under heavy cell load.
- In a loaded network, any communication with the D2D server may take a considerable amount of time since the signaling traffic shares the same channel with the data traffic due to the absence of dedicated bearer support in the mobile OS API.
- If the users are highly mobile, then their location tracking may become inaccurate, thus resulting in false proximity notifications or late link establishments due to relatively short durations of effective proximity.

At this point, it has to be reiterated that while the conventional performance metrics typically focus on throughput, in this particular case of offloading, the subject of specific interest is signaling procedure, see Fig. 3.7, rather than the gains ultimately attainable by its use – video overview of performed trial is available[1].

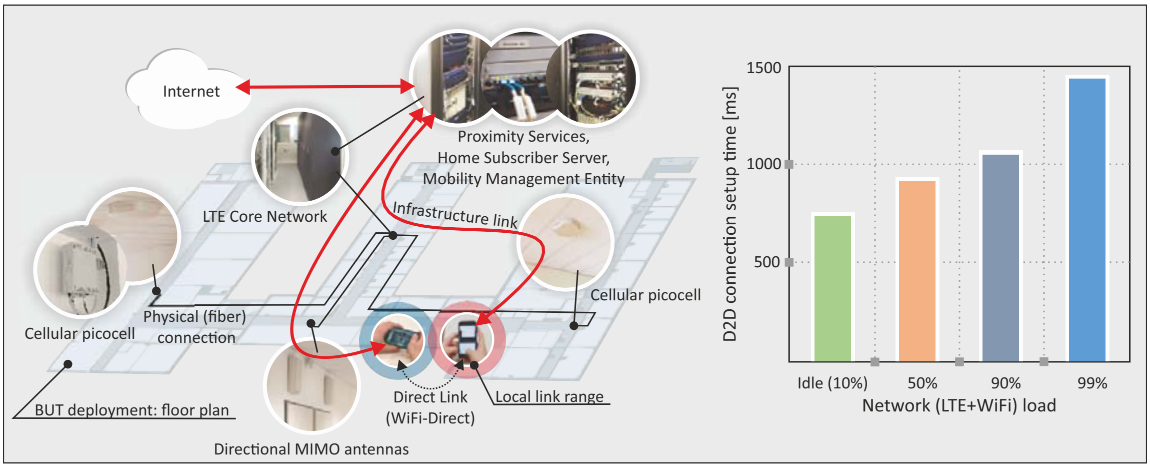


Fig. 3.7: Trial of LTE-assisted WiFi-Direct technology [59].

In summary, as far as future proximity-based services are concerned, we are convinced that it is entirely feasible to utilize our considered network-assisted D2D logic for mobile devices moving at least at pedestrian speeds. Whereas there is some additional delay introduced by network assistance signaling, in almost all cases the connection will be established on time for the users to reliably benefit from it before they come into physical contact, even when walking toward each other.

[1]http://wislab.cz/our-work/lte-assisted-wifi-direct

4 DYNAMIC ALLOCATION OF SPECTRUM RESOURCES

Next-generation (5G) communication networks will provide high bandwidth to mobile users via heterogeneous wireless architectures and dynamic spectrum access techniques. The inefficient usage of the existing spectrum can be improved through opportunistic access to the licensed bands without interfering with the existing users. Upcoming 5G networks, however, impose several research challenges due to the broad range of available spectrum as well as diverse QoS and QoE requirements of 5G applications. These heterogeneities must be captured and handled dynamically as mobile terminals roam between wireless architectures and along the available spectrum pool.

Since the historical fragmentation in spectrum access models accentuates the need for novel concepts that allow for efficient sharing of already available but underutilized spectrum, the emerging LSA regulatory framework is expected to enable more advanced spectrum sharing between a limited number of users while guaranteeing their much needed interference protection. However, the ultimate benefits of LSA may in practice be constrained by space-time availability of the LSA bands. Hence, more dynamic LSA spectrum management is required to leverage such real-time variability and sustain reliability when, for example, the original spectrum user suddenly revokes the previously granted frequency bands as they are required again.

We outline the functionality that is required by the LSA system to achieve the much needed flexible operation as well as report on the results of our respective live trial that employs a full-fledged commercial-grade cellular network deployment. Our practical results become instrumental to facilitate more dynamic bandwidth sharing and thus promise to advance on the degrees of spectrum utilization in future 5G systems without compromising the service quality of their users.

4.1 Highly Dynamic 5G-Ready Spectrum Management with Licensed Shared Access

The 5G wireless systems aim to decisively advance on the levels of spectral and energy efficiency, user-experienced throughput, as well as communication latency and reliability. They are expected to rely on leveraging extremely high frequency (i.e., mmWave) spectrum bands, employing massive Multiple-Input Multiple-Output (MIMO) techniques, as well as deploying increased numbers of small cells with various sizes and across different frequencies. However, the use of mmWave radios is costly and the key enabling technology is still under standardization, whereas massive MIMO requires complex and expensive coordination that is difficult to achieve in practice. Therefore, the main feasible method to offer larger capacity on existing cellular deployments is via extreme network densification [75].

The mobile network operators (5G) are however struggling to deploy a higher density of small cells due to the need of extra investment that is not compensated by the actual revenues [76]. On the other hand, multiple field measurement campaigns strongly show that the conventional spectrum below 6 GHz may be substantially underutilized across space, time, and frequency [77]. This is a consequence of the legacy "command-and-control" spectrum management approach that used to create static and overprotective allocations[2]. Hence, as a viable alternative to deploying additional small cells, the MNOs may quickly augment capacity on their deployments – more *dynamic* and market-friendly spectrum management mechanisms should be made available in the emerging 5G systems [78].

With more dynamic spectrum management, the expensive frequency bands could be shared between different stakeholders flexibly, as opposed to exclusive use of licensed spectrum. This may go far beyond opening up unlicensed frequencies for collective uncontrolled use and promises to unlock the much needed additional bandwidth, that is currently employed sparsely by its existing incumbents. It can also improve the utilization

[2]Spectrum Efficiency Working Group, Report of the spectrum efficiency working group, Federal Communications Commission, 2002.

of presently allocated spectrum across its various dimensions (space, time, frequency), which is essential to support the throughput-hungry 5G applications. To this effect, powerful spectrum sharing technologies emerged recently, such as LTE in unlicensed spectrum (LTE-U), License Assisted Access (LAA), MulteFire, Citizens Broadband Radio Service (CBRS), and Licensed Shared Access (LSA) [79].

The LSA framework is an evolution of the industry-driven Authorized Shared Access (ASA) technology for controlled spectrum sharing between the incumbent holding the rights to use the frequency bands and the licensee (e.g., the MNO), who is utilizing such spectrum temporarily[3]. This concept has been taken forward by the European Commission (EC) to develop a new "individual licensing regime" for authorized spectrum sharing [80]. According to the EC's Radio Spectrum Policy Group (RSPG), the LSA framework enables a limited number of licensees to operate in a frequency band already assigned to one or more incumbents in accordance with well-defined sharing rules. As a result, all of the authorized users, including the incumbents, can maintain their desired Quality of Service (QoS) requirements (RSPG 13-538) [75].

Ever since its introduction several years ago, the LSA concept has spawned an avalanche of engineering, business, and regulatory work that focused on adapting it promptly to practical applications [81]. This development has been facilitated by the Conference Europeenne des Postes et des Telecommunications (CEPT) as it had established two project teams, PT52 and PT53 (ETSI TS 103 113), to ensure that there are no barriers to the adoption of LSA in 2.3 – 2.4 GHz bands from a regulatory perspective[4] [82]. In parallel, ETSI has been targeting to outline the LSA system architecture in their respective technical specifications (ETSI TS 103 154) and (ETSI TS 103 235). In the US, a Notice of Proposed Rulemaking (NPRM) in 3.5 GHz band was introduced by the FCC[5] [83].

As a result of this concentrated effort, the LSA has soon been ready for practical demonstrations, which took place in Spain (2015), Italy (2016), France (2016), Finland (2016), Czech Republic (2016), and the Netherlands (2017). However, many of these past activities considered near-static LSA operation with longer-term allocations since they primarily addressed the technical feasibility of LSA implementation [84]. Building on our previous conceptual work in [85] and technology groundwork in [86], we here complement these earlier initiatives with a new perspective on highly dynamic spectrum management within the LSA framework. In this section, we specifically emphasize the QoS aspects and the corresponding service reliability performance as discussed work unveils the limits of dynamic LSA operation based on a practical trial in a live LTE-A system [75].

4.1.1 Dynamic Spectrum Management with LSA

With the emerging LSA framework, flexible and more dynamic spectrum sharing may be enabled, which becomes increasingly valuable for demanding 5G applications [87]. The advanced services that can benefit from cross-band spectrum aggregation are those that require massive bandwidths, but have difficulty to be supported by the existing MNO deployments (e.g., augmented and virtual reality) [88]. Another category that may take advantage of highly dynamic LSA operation is industrial Internet of Things (IoT) applications across different verticals, especially those requiring reliable operation and dedicated QoS guarantees (e.g., automotive). Finally, LSA can improve a wide range of local broadband services, such as those where the MNOs do not have a possibility to deploy exclusive licensed spectrum (e.g., enterprise) [89], [90].

Therefore, we expect that as LSA technologies mature, an increasing variety of 5G applications and services will be capable of taking advantage of more efficient geographic-temporal spectrum management. In [85], the LSA functionality required to enable truly dynamic spectrum sharing at the timescales of seconds is outlined. These capabilities are crucial for incumbent systems with high-speed mobility (e.g., express trains and airplanes) and may offer much improved performance compared to rigid and near-static LSA implementations. Given that LSA is an example of vertical sharing (see Fig. 4.1a), multiple spectrum users across the same geographical area

[3]Frequency Management Working Group (WGFM), "Report on ASA Concept, FM(12)084 Annex 47", 2012.
[4]EC Mandate on MFCN for 2.3 - 2.4 GHz, 2014.
[5]Enabling Innovative Small Cell Use In 3.5 GHz Band NPRM & Order, 2012.

can operate at different priority tiers. For instance, the LSA licensee (e.g., a commercial LTE network) may avoid causing interference to the LSA incumbent (e.g., an air traffic control system).

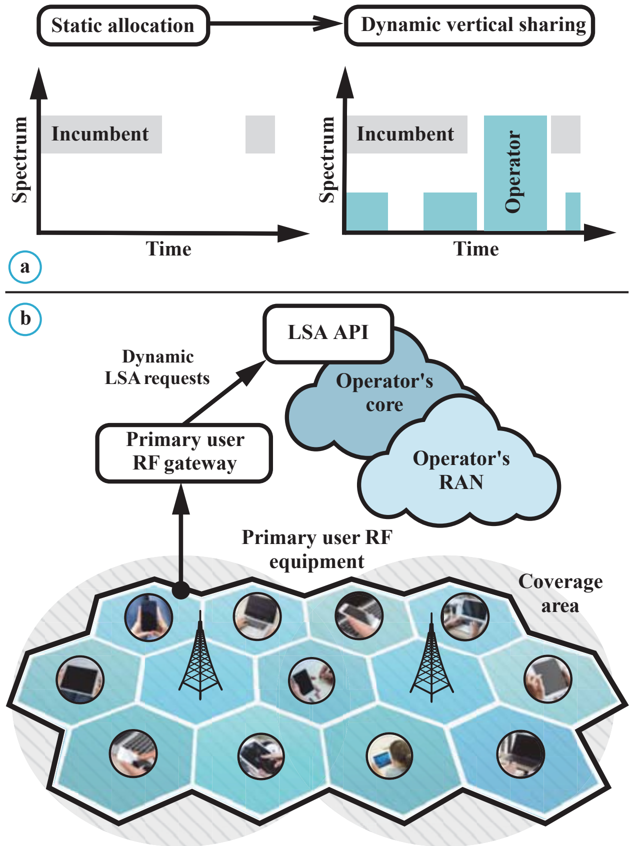


Fig. 4.1: (a) Improved spectrum utilization with dynamic sharing; (b) High-level architecture of dynamic LSA operation [75].

Our envisaged dynamic LSA system is intended to operate according to the high-level architecture captured in Fig. 4.1b. The primary user of the spectrum (i.e., the incumbent) identifies its target constraints (across space, time, and frequency), where the wireless interference constraints have to be met by the secondary user (i.e., the LSA licensee, such as the MNO). Further, a corresponding LSA request is issued and transferred to the operator's dedicated API, where it is then converted into specific RAN instructions (e.g., interference estimation, transmit power reduction, frequency band change, LSA spectrum usage policy, etc.). Once these commands are received by the operator's cloud, its RAN executes the required actions as instructed.

Considering the key mechanisms and constraints in place for dynamic LSA systems, the proposed functionality as discussed below enables the RAN to respond to the received LSA-specific requests issued by the incumbent in near real time. The primary benefit of our approach is in flexible bandwidth segmentation, which is much more fine-grained and adaptive than in previous LSA implementations. With such dynamic and on-demand network configuration, substantial radio resources can be made available to both the incumbent(s) and the LSA licensee(s), since the proposed logic tightly applies to the time domain. Accordingly, the timescale of radio resource utilization has higher granularity than in static LSA approaches. At the time instant, when the incumbent does not utilize its bandwidth resources, they are released automatically, without any additional administrative overheads or delays associated with the LSA database updates.

In addition, our approach efficiently leverages the spatial dimension of the shared spectrum resources. That is, the locations, where the bandwidth has to be released by the LSA licensee (the MNO) back to the incumbent, can be obtained with higher precision than what is possible with the conventional LSA setups, in which the geographical blocks are typically represented as a coarse grid on the map. In the following lines, we continue by providing a systematic perspective on our development efforts to implement a dynamic LSA system. We also highlight our crucial design choices with regards to the communication chain that facilitates near real time LSA operation within a practical LTE network deployment.

4.1.2 Principles of Dynamic LSA Implementation

Based on the above considerations, we expect that upcoming LSA implementations will require software and hardware modifications within the existing cellular network infrastructure as well as, potentially, on the side of the UE. Despite substantial ongoing efforts to evolve the LTE system as one of the 5G cornerstone technologies, a support for highly dynamic LSA operation in practical MNO deployments calls for a dedicated technology development effort. If not reflected comprehensively in further LTE releases, the LSA spectrum sharing mechanisms may be slow to enter the market, where they are much needed at this time. In this section, we address this important demand by exposing the key system functionality required to support highly dynamic LSA in a 3GPP LTE system.

Proposed components and functionality

We remind that wireless technology standardization conventionally begins by defining the functional elements and interfaces between them. Aiming to lay the groundwork for this, we first identify the main functions and interfaces necessary for implementing dynamic LSA mechanisms (see Fig. 4.2). Given that we have recently outlined the core principles behind the dynamic LSA framework as a proof-of-concept study in [86], we build the present system architecture proposal on our rich hands-on experience acquired then. To this end, we rely on a full-fledged cellular system deployment at Brno University of Technology (BUT), Czech Republic, which offers an excellent example of a contemporary 3GPP LTE-A target system.

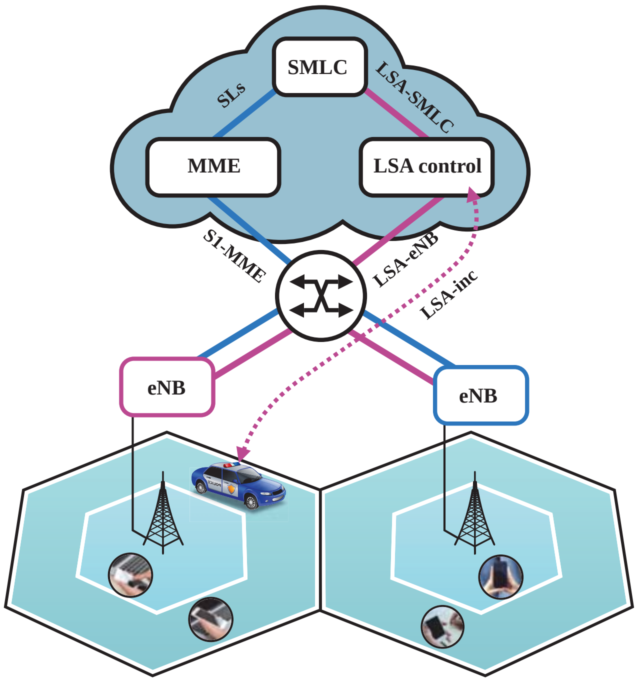


Fig. 4.2: Key elements of proposed dynamic LSA architecture [75].

Control is responsible for accepting the incumbent's requests and managing the power allocation across the network. In conventional LSA systems, this function could be performed by the LSA controller, while the LSA repository acts as a proxy. For the intended highly dynamic operation, we advocate for the need of a direct interface between the incumbent and the LSA controller to reduce control-plane latency (LSA-inc).

Positioning is used for locating the sources of interference in the network, which is key to efficient power allocation. In cases when the LSA band is utilized for downlink (DL) communication, the positions of all LTE eNBs are typically well-known, and may thus be programmed into the controller. However, in cases when the LSA is applied to uplink (UL) or TDD bands, it is important to have the position estimates of the UEs as well, since they become the main sources of interference. Currently, such information is not available to the LSA controller, and we propose to introduce the LSA-SMLC interface, which would allow the LSA controller to directly access the LTE positioning data (as calculated based on the cellular signals or reported by the UEs themselves).

Policy Assignment offers the capability to issue commands to the individual cells, which is at the very core of the LSA concept. In our dynamic LSA system, this implies setting transmit power constraints for the corresponding radio interfaces – in the limit up to full cell shutdown. However, a commonly overlooked issue is that shutting down a particular cell does not necessarily eliminates the interference coming from its service area. This is because some UEs may associate (or remain associated) with a neighboring cell and thus continue their transmissions on the LSA bands for the extended periods of time. The use of OA&M interface (as proposed by 3GPP and utilized in static LSA trials) may therefore become insufficient to implement the control decisions with the required levels of responsiveness. The LTE-eNB interface indicated in our diagram (see Fig. 4.2) aggregates (i) functions normally accessible via OA&M (set the transmit power, shut the cell down, etc.) as well as (ii) certain instructions given directly to the UE via the RRC interface (e.g., initiate handover to a specific cell, switch the band, etc.).

Envisioned LSA System Operation our proposed system design targets to ensure that the dynamic LSA framework can achieve the highest degree of control accuracy. To this end, we enable the UEs to know their precise location (which can even be acquired indoors if desired), and make them report it to the SMLC. This information can then be extracted from the SMLC via a proprietary system monitoring interface, which acts as LSA-SMLC. The control interface LSA-inc can be implemented as a JSON structure (via socket), over which the current coordinates and the threshold power settings could be reliably reported with the minimal delay by each of the incumbent's users, thus defining a constraint in the controller's power allocation algorithm. Unlike in the static LSA cases, dynamic reports are transient in nature and time out on their own. Hence, the network reverts to its default operation whenever no more reports are being sent. For our below test measurements, such reports were triggered manually.

Most importantly, to address the increased control granularity in the dynamic LSA system, the LSA-eNB interface needs to be implemented as a combination of the OA&M and the direct UE control. One of our core proposals, that are instrumental to the dynamic LSA operation, is to reduce the transmit power instead of a complete cell shutdown, which considerably improves system capacity. While it is relatively easy to lower the UL or DL power limit in a cell instead of shutting it down, actually ensuring it in the UEs that are thus forced outside of the cell's coverage area is much more difficult. In our setup, the DL power control has a few possible settings that match the cell coverage area in the DL with its intended service area in the UL.

In a commercial-grade test deployment, it may be cumbersome to send the RRC control signals from the core network that would enforce the UE handover out of the LSA band (as UEs prefer to handover inside their current band when a cell is shut down). To mimic this functionality, the UEs may employ a user-space program that would shut them down instead. Further, the UEs that fall out of the service area of a particular cell – but still able to receive its DL signaling – may, according to their protocol, attempt a RACH procedure. These RACH transmissions use power ramping if unsuccessful (and they will be unsuccessful as the cell is instructed not to accept the initiating UEs), and may thus violate the interference constraints. While RACH packets are only short bursts, they may still cause issues in certain cases.

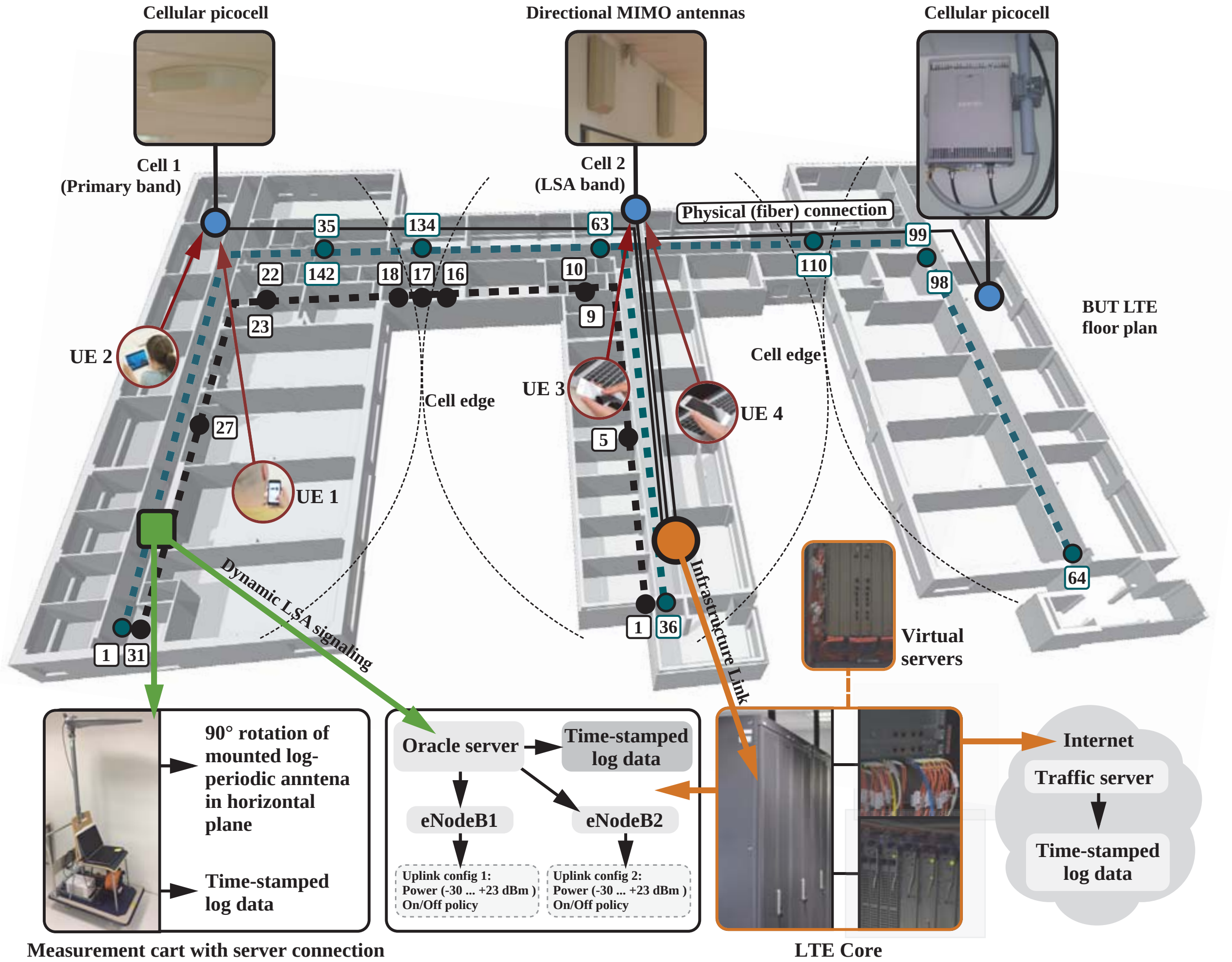


Fig. 4.3: Implemented dynamic LSA setup (architectural components) in our test 3GPP LTE infrastructure [75].

4.1.3 Important practical considerations

Out of the three interfaces identified for the needs of dynamic LSA operation (see the previous section), only two are deployment-ready today. Indeed, reporting the desired interference constraints may be readily achieved with existing IP-based protocols, while connecting the SMLC with the LSA controller is fairly straightforward. On the contrary, ensuring that all of the UEs follow the dynamic power allocation policy in a predictable manner is complicated due to certain limitations in current cellular signaling:

- The UEs cannot be forced to handover away from their current serving eNB without a deep integration into the proprietary code inside the MME. While the required functionality may be made available by some of the core network vendors, it is presently not a part of any standard specification. Similarly, UE's handovers between the individual cells under the same eNB may not be even reported to the MME, which translates into the need for more proprietary interfaces as of today.
- The service area of a cell with the reduced UL power level may not be easily predicted by the UEs, thus often resulting in futile RACH attempts by the devices that are relatively far away from the cell center (and given its UL power limits). Adding the relevant information elements into the beacon signal could certainly allow to improve on this, but it is neither supported by the current beacon formats nor available in the practical eNBs utilized for testing.

Our LSA controller implementation utilizes all of the needed interfaces together with the relevant power allocation policies that are discussed in the following section. Its current version employs a heuristic iterative search procedure to locate the optimal power assignment across all cells in order to match certain performance targets (e.g., maximize the throughput or minimize the number of users that lose service). For the most unsatisfied constraint, a simple algorithm is applied in a loop: a vector of the expected reduction in interference for a 1-dB reduction of power in cell i is computed, and the most impactful cell is chosen. The power in this cell is then reduced by 1 dB and the vector is updated (with its sorted order restored). Once the constraint at hand is no longer the most unsatisfied, another constraint is chosen to proceed further. When all of constraints are satisfied, the search is complete.

The asymptotic complexity of our proposed algorithm is $O(C \cdot R)$, where C is the number of cells and R is the number of constraints (e.g., incumbent's devices) in the system. The resultant power allocations may be then stored and used to initialize the subsequent runs, thus further reducing the time needed for reaction. This means that proposed solution can be scaled up to hundreds of cells if desired, without much sacrifice in the response times. To summarize, our proposed system implements all of the dynamic LSA functionality. While some of its components may not be production-ready as of yet (i.e., the LSA-eNB interface) due to the limitations in the underlying LTE subsystems, it clearly confirms that the considered system is not only feasible, but may also be deployed in larger cellular networks with reasonable effort [75].

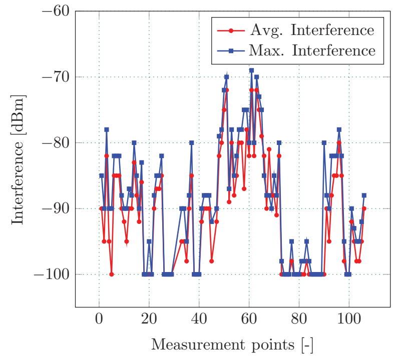
4.1.4 Measurement Methodology and Obtained Results

To systematically demonstrate the outlined principles of highly dynamic LSA operation in practice, we conducted a full-scale real-world implementation of our capable LSA-based spectrum sharing system in a commercial-grade 3GPP LTE channel and evaluated its availability over the LSA frequency bands. The UL system has been preferred due to its significantly higher implementation complexity as compared to the DL LTE channel, which also led to more interesting observations. The primary system-level parameters are summarized in Table 4.1, while the composition of our trial implementation is detailed in Fig. 4.3.

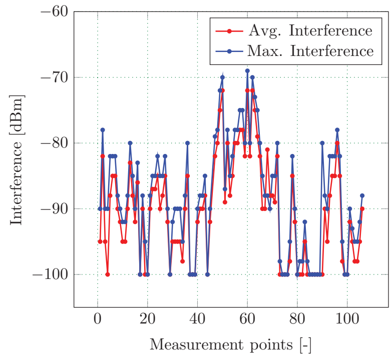
Tab. 4.1: Main system-level parameters.

Description	**Value**
3GPP LTE system baseline	Release 10
Division multiplexing	FDD
Number of cells (eNBs)	3
Frequency band	17 (700 MHz)
Bandwidth	5 MHz
Number of resource blocks (RBs)	25
Max. eNB power level	0 dB
Min. eNB power level	-30 dB
Interference threshold	-85 dBm
Path loss coefficients	5 dBm (concrete) 2 dBm (gypsum)
Pathloss model	Enhanced FSPL
Number of terminals (UEs)	4
Transmission data rate	512 kbps
Frequency analyzer	R&S TSMW
Antenna	HL040 log-periodic broadband

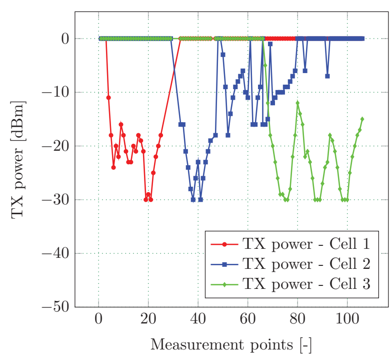
The UEs under test were continuously communicating with the data server located in the Internet, which recorded their effective UL and DL bit-rate values over time. Prior to taking measurements, the UEs were configured to target a constant bit-rate (CBR) transmission at 512 kbps, if sufficient radio resources were available; otherwise, they utilized all of the remaining UL resources subject to the current transmit power restrictions. The trial focused on analyzing the LSA band and demonstrating its highly dynamic operation. Hence, the UEs were forced to close the connection whenever they were supposed to switch over to the non-LSA frequencies.



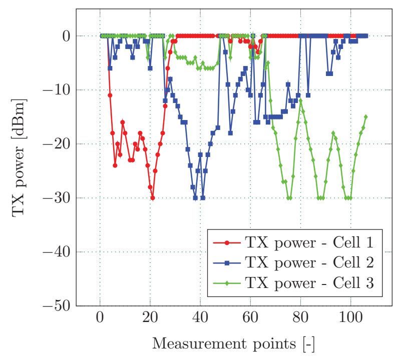
(a) Average/maximum interference for Derivative algorithm.



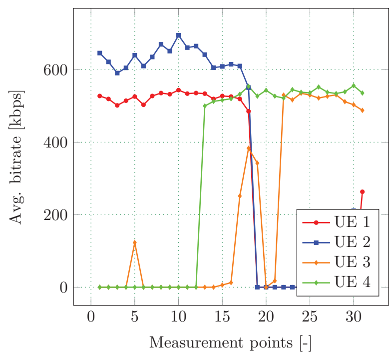
(b) Average/maximum interference for Shannon algorithm.



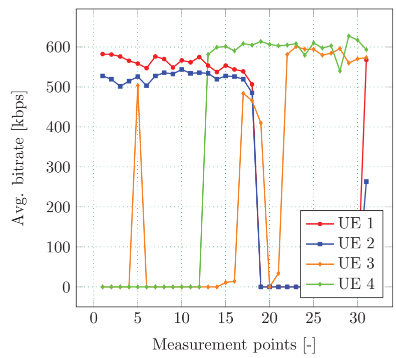
(c) TX power allocation for each cell according to Derivative algorithm.



(d) TX power allocation for each cell according to Shannon algorithm.



(e) Average bit-rate for deployed UEs (black path in Fig. 4.3).



(f) Maximum bit-rate for deployed UEs (black path in Fig. 4.3).

Fig. 4.4: Measurement-based results produced in the conducted dynamic LSA trial [75].

Power allocation policies

Two representative power allocation policies based on heuristic iterative search were implemented as part of the LSA controller: (i) "Derivative" algorithm and (ii) "Shannon" algorithm. Both methods assign the maximum allowed uplink power across the cells such that the total interference from the network towards the moving measurement cart (which represents the incumbent) does not exceed the given threshold. Our trial scenario utilizes four UEs; they remain stationary and their positions are known. The interference from a cell (eNB) can typically be approximated by the maximum interference from all the UEs in this cell. Hence, for the three utilized cells (see Fig. 4.3), the total interference on the measurement cart is estimated as a sum of the interference levels produced by the UEs closest to the cart, taken across all cells.

Derivative algorithm lowers the maximum allowed power across the cells in such a way that the decrease in interference per dB of power reduction is maximal. It chooses the cell in which the interfering UE is located closest to the cart and lowers the power in that cell – thus gaining the maximal loss in interference against the minimal loss in the overall power. If the power in a cell is reduced below its minimum feasible level, it is then shut down.

Shannon algorithm calculates the potential decrease in the total interference against the power reduction in each cell; then, using the capacity-estimation formula it also evaluates the changes in the user's effective transmission rate. Importantly, if the user's bit-rate is below its required value (subject to the current power restriction), there is no change in the effective transmission rate. Further, the algorithm selects the cell for power decrease to maximize the interference reduction subject to the minimal loss in the total effective transmission rate of its users. In effect, it attempts to maintain the UE connectivity to the cells instead of providing them with the highest reliability of the service. Therefore, this algorithm prioritizes keeping the users connected.

A user is considered to be connected if its effective bit-rate is higher than a certain threshold. Similar to Derivative algorithm, Shannon algorithm first lowers the power in the closest cell, thus reducing its interference towards the measurement cart, but only until the lowest transmission rate in this cell reaches the threshold. While the interference threshold is still not reached, the algorithm moves to the second closest cell. If the power level is such that the bit-rate of the "worst" UE has reached the threshold, but the interference still remains above the target value, the algorithm returns to the first closest cell, drops the "worst" user and again lowers the power in a similar manner, by comparing the second "worst" user's bit-rate against the threshold. Varying the transmission rate threshold, one can control the minimal guaranteed bit-rate. On the other hand, setting the threshold too high can cause more user drops from the network, since a UE is considered dropped when it cannot maintain its threshold bit-rate. Therefore, a sensible solution would be to set the QoS-guaranteed bit-rate equal to the threshold transmission rate [75].

Key Hands-on Observations

Our primary objectives in the conducted dynamic LSA trial were to (i) compare the above power allocation policies and (ii) verify whether the corresponding algorithms operate as intended, i.e., do not breach the interference threshold while meeting their respective optimization targets. The measurements were collected on a predetermined set of the measurement cart locations, see Fig. 4.3, where two evaluation scenarios are illustrated: (i) the black path indicates the measurements for the purposes of LSA assessment while (ii) the blue path stands for the overall testing of the implemented heuristic iterative search logic. Note that the system was notified on the movement of the measurement cart, while at every time instance the interference was recorded. All of the relevant data was continuously logged for further analysis. Accordingly, Fig. 4.4a and Fig. 4.4b report on the average (red line) and the maximum (blue line) interference levels received by the measurement cart at each position for both algorithms.

Comparing Fig. 4.4c and Fig. 4.4d, that illustrate the power allocated according to Derivative and Shannon algorithms respectively, one can observe, that when the cart is located in the first cell the Derivative algorithm only alters the power allocation there, while the Shannon algorithm adjusts all three cells in order to raise power and transmission rate in the closest cell. Hence, it can be concluded that Shannon algorithm operates more flexibly over the entire network, while Derivative algorithm mostly concentrates on the nearest cells. From Fig. 4.4a and Fig. 4.4b, where the average and the maximum interference received by the measurement cart across the check points are reported, we learn that unlike the Derivative algorithm the Shannon algorithm not only lowers the average interference, but also attempts to maintain it as close to the threshold as possible, so increasing power and transmission rate in the network.

Importantly, for most of the considered locations our implemented heuristic iterative search logic is ensuring that the interference threshold level is not breached. Further, analyzing the service rates displayed in Fig. 4.4e and Fig. 4.4f, we note that the cells can support the necessary data rates of 512 kbps for their cell-center UEs at any power level, until they have to be shut down. Clearly, the cell-edge UEs n.2 and n.3 suffer from less consistent service even when the cells remain up. However, we generally observe that while the iterative search logic may not always provide consistent performance, the interference received by the measurement cart in most of the cases remains below the defined threshold. This is true even when cells (Cell 1 and Cell 2) are turned on and the measurement cart is located directly between them [75].

4.1.5 Main Outcomes and Perspectives

This work accentuates the importance of highly dynamic spectrum sharing to leverage additional bandwidth that may be lightly used by its original incumbents. To further improve upon spectrum utilization in demanding 5G systems, we focus on the emerging LSA framework for vertical sharing, where the incumbent(s) and the licensee(s) operate over the same geographical area by utilizing common frequencies in a carefully controlled manner. This concept has been coined in 2013 and since then rapidly took flight with many hands-on demonstrations across Europe (see Fig. 4.5), primarily in 2016. Supported by visible research initiatives, such as ADEL [77] and CORE++ [79], [81], [83], the LSA functionality has been tested in a number of countries with the emphasis on the feasibility of its early implementation.

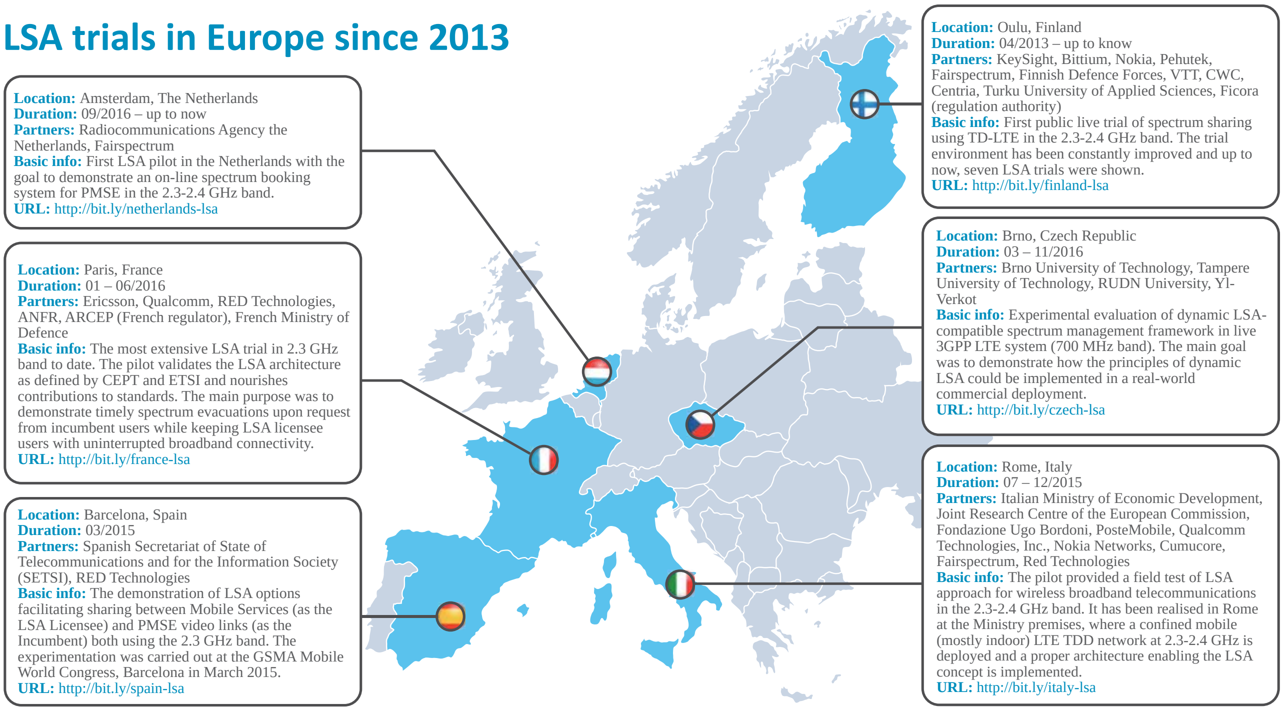


Fig. 4.5: Geography of existing LSA trials in Europe [75].

Complementing these important efforts, the present study relies on our rigorous past research[6] to advance the state-of-the-art on LSA by outlining its additional functionality required for highly dynamic operation. We therefore elaborate the key principles of system implementation as well as contribute our unique practical methodology based on a live cellular network deployment. The obtained measurements support the rich capabilities of highly dynamic LSA operation in a commercial-grade network as well as report on the crucial performance indicators related to QoS and service reliability.

Going further, we believe that our results will become instrumental to fully reap the benefits of LSA-based highly dynamic spectrum management in 5G networks, whether in 2.3 – 2.4 GHz bands or at alternative frequencies, such as 3.4 – 3.8 GHz and possibly up to 4.2 GHz in perspective. This will require further demonstration efforts that may rely on our methodology proposed in this work, which could also be useful for other spectrum sharing initiatives across the globe, including Citizens Broadband Radio Service (CBRS) in the US as well as dynamic spectrum utilization at mmWave frequencies. Ultimately, our work advances the efficient utilization of LSA-based spectrum management at high granularity and facilitates its adoption in commercial cellular network deployments.

[6] Please, refer to a summary on our prior system-level evaluations of dynamic LSA operation here: http://winter-group.net/dyn-lsa-sim-res/

5 CONCLUDING SUMMARY

Rapid penetration of wireless connectivity, almost exponential increase in wireless data (multimedia) usage and proliferation of feature-rich smart devices are gradually setting the stage for next major cellular evolution towards 5G. Next-generation (5G) wireless systems are already promising a manifold increase in data rate, ubiquitous connectivity, Quality of Service, and Quality of Experience. A plethora of new applications, like Machine-to-Machine communication, Internet of Things, Tactile Internet, Augmented and Virtual reality, and Smart Grids are expected to be supported under the umbrella of 5G wireless communication systems.

Within the full version of the dissertation thesis, the mobile networks in the 5G era are comprehensively described in the Chapter 2, where a wide variety of new emerging 5G applications (substantiating the commercial roll out of 5G wireless systems) are discussed thoroughly. Having summarized the key requirements given by the 5G applications, the second part of the Chapter 2 details the 5G technology enablers. To provide the complete picture, this part is further divided to: (i) 5G radio access network enablers (containing both, Radio Access Technology and Radio Access Network parts of the 5G access network i.e., Massive MIMO, Device-to-Device communication, millimeter wave technology, Licensed Shared Access and Radio Access Network as a Service), and (ii) 5G core networks enablers, which are pertinent to support the above 5G RAN enablers i.e., Software Defined Networks and Network Function Virtualization.

At the length of the full dissertation thesis, author has focused on two promising research areas targeting the research questions connected with the lack of bandwidth: (i) proximity-based (cellular-assisted) networking (Chapter 3), and (ii) dynamic allocation of spectrum resources (Chapter 4). Both chapters are structured with the goal to provide at the first the problem statement followed by the design of proposed novel mechanisms (communication principles), targeting to reach the requirements given by the upcoming 5G applications and the **comprehensive evaluation of implemented principles within the experimental LTE deployment at Brno University of Technology, Czech Republic.**

As the direct communication was recognized as the one of the key communication mechanisms to provide end users with the virtually unlimited capacity and ubiquitous coverage introducing the “anytime & anywhere” connectivity, the proximity-based (cellular-assisted) networking is detailed within the Chapter 3. This novel form of networking is enabled by D2D communication between / among the laptops, smartphones, and wearables of people in proximity of each other. Unfortunately, it has remained limited by the fact that most people are simply not aware of the many potential virtual opportunities in their proximity at any given time. This is complemented by the topical digital privacy and security concerns surrounding direct communication between “stranger’s” devices. Fortunately, these concerns can be mitigated with the help of a centralized trusted entity, such as a cellular service provider, which can not only authenticate and protect the privacy of devices involved into D2D communication, but also facilitate the discovery of device capabilities and their available content. Based on the above, Chapter 3 offers an extensive research work behind this type of “cellular-assisted” D2D communication, detailing the key aspects of this technology as well as its implementation, relevant usage scenarios, security challenges, and user experience observations from large-scale LTE deployment located at Brno University of Technology. With the proposed novel prototype implementation, current 3GPP LTE networks may supply mobile devices with effective means to discover, connect, and communicate with their desired proximal partners over high-rate WiFi-Direct channels. What is even more important, such connectivity can be made seamless and automatic, taking advantage of reliable, secure, and optimized direct links. **As a main output of this particular research (Chapter 3), the obtained findings (supported by the realized 3GPP LTE-Assisted Wi-Fi Direct trial implementation[7]) were included as a part of the 3GPP Release 12 specification [91].**

Chapter 4 further builds on the studied requirements given by the upcoming 5G applications, when the attention is given to the dynamic allocation of spectrum resources. First part of the fourth chapter is dedicated

[7] http://wislab.cz/our-work/lte-assisted-wifi-direct

to a practical evaluation of RAN modeling in indoor LTE deployment, where we benefit from the full access to the real LTE-A deployment and complete building documentation. The measured values of Signal to Interference plus Noise Ratio and network throughput are compared with the results from the novel radio signal propagation model, implemented in Python. The obtained values are in consequence used for the transmission power configuration to realize the coverage and capacity optimization in cellular systems. In the process of configuration the TX power optimization mechanism, a number of challenges regarding the used LTE-A deployment have been addressed. As mentioned at length in the Chapter 4 (in full version of dissertation thesis), utilized LTE test-bed is provided by Huawei company. **Since the configuration procedure highly depends on the vendor (Huawei, Ericsson, etc.), we primarily aimed to identify key elements of TX power management mechanisms, which are common for all LTE deployments.**

The second part of the Chapter 4 (in full version of dissertation thesis) deals with the thorough description of the 5G-Ready Highly Dynamic Spectrum Management with LSA framework. Despite decisive progress in many enabling solutions, fifth-generation cellular deployments may still suffer from a glaring lack of bandwidth due to inefficient utilization of radio spectrum resources, which calls for immediate action. To this end, several capable frameworks have recently emerged to help the mobile network operators leverage the abundant frequency bands that are lightly utilized by other incumbents. The recent Licensed Shared Access regulatory framework allows for controlled sharing of spectrum between the incumbent and licensee, such as the MNOs, which coexist geographically. This powerful concept has been subject to several early technology demonstrations that confirm its implementation feasibility. However, the full potential of LSA-based spectrum management can only become available if it is empowered to operate dynamically and at high space-time-frequency granularity. Complementing the prior efforts, in this work, the functionality that is required by the LSA system to achieve the much needed flexible operation as well as report on the results of respective live trial that employs a full-fledged commercial-grade cellular network deployment is detailed. **Obtained practical results (supported by a novel mathematical analysis of the LSA policy in a characteristic scenario featuring a highly-dynamic incumbent (the airport)) become instrumental to facilitate more dynamic bandwidth sharing and thus promise to advance on the degrees of spectrum utilization in future 5G cellular systems without compromising the service quality of their users.**

BIBLIOGRAPHY

[1] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.

[2] M. J. Marcus, "5G and" IMT for 2020 and beyond"[Spectrum Policy and Regulatory Issues]," *IEEE Wireless Communications*, vol. 22, no. 4, pp. 2–3, 2015.

[3] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015–2020 White Paper," *Cisco Public Information*, 2017.

[4] E. Hossain and M. Hasan, "5G cellular: key enabling technologies and research challenges," *IEEE Instrumentation & Measurement Magazine*, vol. 18, no. 3, pp. 11–21, 2015.

[5] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, 2014.

[6] P. Masek, J. Hosek, K. Zeman, M. Stusek, D. Kovac, P. Cika, J. Masek, S. Andreev, and F. Kröpfl, "Implementation of true IoT vision: Survey on enabling protocols and hands-on experience," *International Journal of Distributed Sensor Networks*, 2016.

[7] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?," *IEEE Journal on selected areas in communications*, vol. 32, no. 6, pp. 1065–1082, 2014.

[8] T. Shuminoski and T. Janevski, "Radio network aggregation for 5G mobile terminals in heterogeneous wireless networks," in *11th International Conference on Telecommunication in Modern Satellite, Cable and Broadcasting Services (TELSIKS), 2013*, vol. 1, pp. 225–228, IEEE, 2013.

[9] S. Talwar, S.-p. Yeh, N. Himayat, K. Johnsson, G. Wu, and R. Q. Hu, "Capacity and coverage enhancement in heterogeneous networks," *Heterogeneous Cellular Networks*, pp. 51–65, 2013.

[10] S. Andreev, O. Galinina, A. Pyattaev, J. Hosek, P. Masek, H. Yanikomeroglu, and Y. Koucheryavy, "Exploring synergy between communications, caching, and computing in 5G-grade deployments," *IEEE Communications Magazine*, vol. 54, no. 8, pp. 60–69, 2016.

[11] I. Hwang, B. Song, and S. S. Soliman, "A holistic view on hyper-dense heterogeneous and small cell networks," *IEEE Communications Magazine*, vol. 51, no. 6, pp. 20–27, 2013.

[12] F. Khan, Z. Pi, and S. Rajagopal, "Millimeter-wave mobile broadband with large scale spatial processing for 5G mobile communication," in *50th Annual Allerton Conference on Communication, Control, and Computing (Allerton), 2012*, pp. 1517–1523, IEEE, 2012.

[13] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 122–130, 2014.

[14] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE access*, vol. 3, pp. 1206–1232, 2015.

[15] D. Raychaudhuri and N. B. Mandayam, "Frontiers of wireless and mobile communications," *Proceedings of the IEEE*, vol. 100, no. 4, pp. 824–840, 2012.

[16] T. S. Rappaport, W. Roh, and K. Cheun, "Wireless engineers long considered high frequencies worthless for cellular systems. They couldn't be more wrong," *IEEE SPECTRUM*, vol. 51, no. 9, pp. 34–+, 2014.

[17] M. Beale, “Future challenges in efficiently supporting M2M in the LTE standards,” in *IEEE Wireless Communications and Networking Conference Workshops (WCNCW), 2012*, pp. 186–190, IEEE, 2012.

[18] K. Zheng, F. Hu, W. Wang, W. Xiang, and M. Dohler, “Radio resource allocation in LTE-advanced cellular networks with M2M communications,” *IEEE communications Magazine*, vol. 50, no. 7, 2012.

[19] A. Laya, L. Alonso, and J. Alonso-Zarate, “Is the Random Access Channel of LTE and LTE-A Suitable for M2M Communications? A Survey of Alternatives,” *IEEE Communications Surveys and Tutorials*, vol. 16, no. 1, pp. 4–16, 2014.

[20] H. Holma, A. Toskala, and J. Reunanen, *LTE Small Cell Optimization: 3GPP Evolution to Release 13*. John Wiley & Sons, 2016.

[21] G. I. P. Association *et al.*, “5G Vision-The 5G Infrastructure Public Private Partnership: the next generation of communication networks and services,” *White Paper, February*, 2015.

[22] 3GPP Release 8.0, “Overview of 3GPP Release 8 v0.3.3 (2009-2014).” `http://www.3gpp.org/specifications/releases/72-release-8`. [Online].

[23] J. G. Andrews, “Seven ways that HetNets are a cellular paradigm shift,” *IEEE Communications Magazine*, vol. 51, no. 3, pp. 136–144, 2013.

[24] H. Ali-Ahmad, C. Cicconetti, A. De la Oliva, M. Draxler, R. Gupta, V. Mancuso, L. Roullet, and V. Sciancalepore, “CROWD: an SDN approach for DenseNets,” in *Second European Workshop on Software Defined Networks (EWSDN), 2013*, pp. 25–31, IEEE, 2013.

[25] B. Soret, K. I. Pedersen, N. T. Jørgensen, and V. Fernández-López, “Interference coordination for dense wireless networks,” *IEEE Communications Magazine*, vol. 53, no. 1, pp. 102–109, 2015.

[26] M. Condoluci, M. Dohler, G. Araniti, A. Molinaro, and K. Zheng, “Toward 5G densenets: architectural advances for effective machine-type communications over femtocells,” *IEEE Communications Magazine*, vol. 53, no. 1, pp. 134–141, 2015.

[27] O. N. Yilmaz, Z. Li, K. Valkealahti, M. A. Uusitalo, M. Moisio, P. Lundén, and C. Wijting, “Smart mobility management for D2D communications in 5G networks,” in *IEEE Wireless Communications and Networking Conference Workshops (WCNCW), 2014*, pp. 219–223, IEEE, 2014.

[28] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, “5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1201–1221, 2017.

[29] G. Wunder, M. Kasparick, S. ten Brink, F. Schaich, T. Wild, I. Gaspar, E. Ohlmer, S. Krone, N. Michailow, A. Navarro, *et al.*, “5GNOW: Challenging the LTE design paradigms of orthogonality and synchronicity,” in *IEEE 77th Vehicular Technology Conference (VTC Spring), 2013*, pp. 1–5, IEEE, 2013.

[30] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, 2014.

[31] J. Nam, J.-Y. Ahn, A. Adhikary, and G. Caire, “Joint spatial division and multiplexing: Realizing massive MIMO gains with limited channel state information,” in *46th Annual Conference on Information Sciences and Systems (CISS), 2012*, pp. 1–6, IEEE, 2012.

[32] Z. Pi and F. Khan, “An introduction to millimeter-wave mobile broadband systems,” *IEEE communications magazine*, vol. 49, no. 6, 2011.

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

[34] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 86–92, 2014.

[35] M. R. Palattella, M. Dohler, A. Grieco, G. Rizzo, J. Torsner, T. Engel, and L. Ladid, "Internet of things in the 5G era: Enablers, architecture, and business models," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 3, pp. 510–527, 2016.

[36] A. Ometov, P. Masek, J. Urama, J. Hosek, S. Andreev, and Y. Koucheryavy, "Implementing secure network-assisted D2D framework in live 3GPP LTE deployment," in *IEEE International Conference on Communications Workshops (ICC), 2016*, pp. 749–754, IEEE, 2016.

[37] H.-J. Kwon, J. Jeon, A. Bhorkar, Q. Ye, H. Harada, Y. Jiang, L. Liu, S. Nagata, B. L. Ng, T. Novlan, *et al.*, "Licensed-Assisted Access to Unlicensed Spectrum in LTE Release 13," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 201–207, 2017.

[38] S.-Y. Lien, C.-C. Chien, H.-L. Tsai, Y.-C. Liang, and D. I. Kim, "Configurable 3GPP Licensed Assisted Access to Unlicensed Spectrum," *IEEE Wireless Communications*, vol. 23, no. 6, pp. 32–39, 2016.

[39] K. Chatzikokolakis, P. Spapis, A. Kaloxylos, and N. Alonistioti, "Toward spectrum sharing: opportunities and technical enablers," *IEEE Communications Magazine*, vol. 53, no. 7, pp. 26–33, 2015.

[40] K. Buckwitz, J. Engelberg, and G. Rausch, "Licensed shared access (LSA)– Regulatory background and view of administrations," in *9th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), 2014*, pp. 413–416, IEEE, 2014.

[41] F. Granelli, A. A. Gebremariam, M. Usman, F. Cugini, V. Stamati, M. Alitska, and P. Chatzimisios, "Software defined and virtualized wireless access in future wireless networks: scenarios and standards," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 26–34, 2015.

[42] H.-H. Cho, C.-F. Lai, T. K. Shih, and H.-C. Chao, "Integration of SDR and SDN for 5G," *IEEE Access*, vol. 2, pp. 1196–1204, 2014.

[43] A. Hakiri and P. Berthou, "Leveraging SDN for the 5G networks: trends, prospects and challenges," *arXiv preprint arXiv:1506.02876*, 2015.

[44] D. Kreutz, F. M. Ramos, P. E. Verissimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-defined networking: A comprehensive survey," *Proceedings of the IEEE*, vol. 103, no. 1, pp. 14–76, 2015.

[45] M. Mendonca, K. Obraczka, and T. Turletti, "The case for software-defined networking in heterogeneous networked environments," in *Proceedings of the 2012 ACM conference on CoNEXT student workshop*, pp. 59–60, ACM, 2012.

[46] M. Amani, T. Mahmoodi, M. Tatipamula, and H. Aghvami, "SDN Based Data Offloading for 5G Mobile Networks," *ZTECOMMUNICATIONS*, vol. 12, p. 34, 2014.

[47] B. Han, V. Gopalakrishnan, L. Ji, and S. Lee, "Network function virtualization: Challenges and opportunities for innovations," *IEEE Communications Magazine*, vol. 53, no. 2, pp. 90–97, 2015.

[48] P. Demestichas, A. Georgakopoulos, D. Karvounas, K. Tsagkaris, V. Stavroulaki, J. Lu, C. Xiong, and J. Yao, "5G on the horizon: key challenges for the radio-access network," *IEEE Vehicular Technology Magazine*, vol. 8, no. 3, pp. 47–53, 2013.

[49] D. Wubben, P. Rost, J. S. Bartelt, M. Lalam, V. Savin, M. Gorgoglione, A. Dekorsy, and G. Fettweis, "Benefits and impact of cloud computing on 5G signal processing: Flexible centralization through cloud-RAN," *IEEE signal processing magazine*, vol. 31, no. 6, pp. 35–44, 2014.

[50] A. Pyattaev, J. Hosek, K. Johnsson, R. Krkos, M. Gerasimenko, P. Masek, A. Ometov, S. Andreev, J. Sedy, V. Novotny, *et al.*, "3GPP LTE-assisted Wi-Fi-direct: Trial implementation of live D2D technology," *Etri Journal*, vol. 37, no. 5, pp. 877–887, 2015.

[51] J. Lehtomäki, I. Suliman, J. Vartiainen, M. Bennis, A. Taparugssanagorn, and K. Umebayashi, "Direct communication between terminals in infrastructure based networks," *Proc. ICT Mobile Wireless Commun. Summit*, pp. 1–8, 2008.

[52] G. Fodor, E. Dahlman, G. Mildh, S. Parkvall, N. Reider, G. Miklós, and Z. Turányi, "Design aspects of network assisted device-to-device communications," *IEEE Communications Magazine*, vol. 50, no. 3, 2012.

[53] W.-F. Alliance, "Wi-Fi Alliance Technical Committee P2P Task Group," *Wi-Fi Peer-to-Peer (P2P) Technical Specification, Version*, vol. 1, pp. 1–159, 2010.

[54] L. Zhang, J. K. Muppala, and W. Tu, "Exploiting proximity in cooperative download of large files in peer-to-peer networks," in *Second International Conference on Internet and Web Applications and Services, 2007. ICIW'07.*, pp. 1–1, IEEE, 2007.

[55] A. Berl and H. De Meer, "Integration of mobile devices into popular peer-to-peer networks," in *NGI'09 Next Generation Internet Networks, 2009*, pp. 1–9, IEEE, 2009.

[56] P. Masek, K. Zeman, J. Hosek, Z. Tinka, N. Makhlouf, A. Muthanna, N. Herencsar, and V. Novotny, "User performance gains by data offloading of LTE mobile traffic onto unlicensed IEEE 802.11 links," in *38th International Conference on Telecommunications and Signal Processing (TSP), 2015*, pp. 117–121, IEEE, 2015.

[57] A. Pyattaev, K. Johnsson, S. Andreev, and Y. Koucheryavy, "Proximity-based data offloading via network assisted device-to-device communications," in *IEEE 77th Vehicular Technology Conference (VTC Spring), 2013*, pp. 1–5, IEEE, 2013.

[58] A. Pyattaev, K. Johnsson, S. Andreev, and Y. Koucheryavy, "3GPP LTE traffic offloading onto WiFi direct," in *IEEE Wireless Communications and Networking Conference Workshops (WCNCW), 2013*, pp. 135–140, IEEE, 2013.

[59] S. Andreev, J. Hosek, T. Olsson, K. Johnsson, A. Pyattaev, A. Ometov, E. Olshannikova, M. Gerasimenko, P. Masek, Y. Koucheryavy, *et al.*, "A unifying perspective on proximity-based cellular-assisted mobile social networking," *IEEE Communications Magazine*, vol. 54, no. 4, pp. 108–116, 2016.

[60] 3GPP TR 22.803, "Feasibility Study for Proximity Services (ProSe)." `https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=653`. [Online].

[61] 3GPP SP-110638, "LTE Proximity-Based Services Study Item." `https://www.cept.org/Documents/fm-49/5920/FM49(12)032_Standardisation-of-Public-Safety-in-3GPP`. [Online].

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

[63] T. Doumi, M. F. Dolan, S. Tatesh, A. Casati, G. Tsirtsis, K. Anchan, and D. Flore, "LTE for public safety networks," *IEEE Communications Magazine*, vol. 51, no. 2, pp. 106–112, 2013.

[64] 3GPP TR 33.833, "Study on Security Issues to Support Proximity Services." `http://www.3gpp.org/DynaReport/33833.htm`. [Online].

[65] 3GPP TS 32.277, "Telecommunication Management; Charging Management; Proximity-Based Services (ProSe) Charging." `http://www.3gpp.org/DynaReport/32277.htm`. [Online].

[66] 3GPP TS 23.303, "Proximity-Based Services (ProSe); Stage 2." `http://www.3gpp.org/DynaReport/23303.htm`. [Online].

[67] F. M. Abinader, E. P. Almeida, F. S. Chaves, A. M. Cavalcante, R. D. Vieira, R. C. Paiva, A. M. Sobrinho, S. Choudhury, E. Tuomaala, K. Doppler, *et al.*, "Enabling the coexistence of LTE and Wi-Fi in unlicensed bands," *IEEE Communications Magazine*, vol. 52, no. 11, pp. 54–61, 2014.

[68] 3GPP TR 36.843, "Study on LTE Device to Device Proximity Services; Radio Aspects." `http://www.3gpp.org/dynareport/36843.htm`. [Online].

[69] 3GPP TS 23.234, "3GPP System to Wireless Local Area Network (WLAN) Interworking; System Description." `http://www.3gpp.org/DynaReport/23234.htm`. [Online].

[70] 3GPP TS 23.327, "Mobility between 3GPP-Wireless Local Area Network (WLAN) Interworking and 3GPP Systems." `http://www.3gpp.org/DynaReport/23327.htm`. [Online].

[71] 3GPP TS 23.402, "Architecture Enhancements for Non-3GPP Accesses." `http://www.3gpp.org/DynaReport/23402.htm`. [Online].

[72] V. Novotny, R. Krkos, and J. Sedy, "Možnosti využití technologie mobilní sítě 4. generace firmy Huawei," pp. 1–34, 2014.

[73] E. Gelenbe, G. Gorbil, D. Tzovaras, S. Liebergeld, D. Garcia, M. Baltatu, and G. Lyberopoulos, "Security for smart mobile networks: The NEMESYS approach," in *International Conference on Privacy and Security in Mobile Systems (PRISMS), 2013*, pp. 1–8, IEEE, 2013.

[74] A. Tirumala, F. Qin, J. Dugan, J. Ferguson, and K. Gibbs, "Iperf: The TCP/UDP bandwidth measurement tool," *htt p://dast. nlanr. net/Projects*, 2005.

[75] P. Masek, E. Mokrov, K. Zeman, A. Ponomarenko-Timofeev, A. Pyattaev, J. Hosek, S. Andreev, Y. Koucheryavy, and K. Samouylov, "A Practical Perspective on 5G-Ready Highly Dynamic Spectrum Management with LSA," *IEEE Communications Magazine; Under the review (Submitted 07/2017)*, vol. XX, no. X, pp. 1–8, 2017.

[76] M. Mustonen, M. Matinmikko, M. Palola, S. Yrjölä, and K. Horneman, "An evolution toward cognitive cellular systems: licensed shared access for network optimization," *IEEE Communications Magazine*, vol. 53, pp. 68–74, May 2015.

[77] A. Morgado, A. Gomes, V. Frascolla, K. Ntougias, C. Papadias, D. Slock, E. Avdic, N. Marchetti, N. Haziza, H. Anouar, Y. Yang, M. Pesavento, F. Khan, and T. Ratnarajah, "Dynamic LSA for 5G networks the ADEL perspective," in *2015 European Conference on Networks and Communications (EuCNC)*, pp. 190–194, June 2015.

[78] M. Hafeez and J. M. H. Elmirghani, "Green Licensed-Shared Access," *IEEE Journal on Selected Areas in Communications*, vol. 33, pp. 2579–2595, Dec 2015.

[79] S. Yrjölä, V. Hartikainen, L. Tudose, J. Ojaniemi, A. Kivinen, and T. Kippola, "Field Trial of Licensed Shared Access with Enhanced Spectrum Controller Power Control Algorithms and LTE Enablers," *Journal of Signal Processing Systems*, pp. 1–14, 2016.

[80] M. Mustonen, T. Chen, H. Saarnisaari, M. Matinmikko, S. Yrjölä, and M. Palola, "Cellular architecture enhancement for supporting the european licensed shared access concept," *IEEE Wireless Communications*, vol. 21, pp. 37–43, June 2014.

[81] M. Matinmikko, H. Okkonen, M. Palola, S. Yrjölä, P. Ahokangas, and M. Mustonen, "Spectrum sharing using licensed shared access: the concept and its workflow for LTE-advanced networks," *IEEE Wireless Communications*, vol. 21, pp. 72–79, April 2014.

[82] D. Guiducci, C. Carciofi, V. Petrini, E. Spina, G. D. Sipio, D. Massimi, and P. Chawdhry, "Sharing under licensed shared access in a LTE real test network at 2.3-2.4 GHz," in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–6, Sept 2016.

[83] J. Kalliovaara, T. Jokela, R. Ekman, J. Hallio, M. Jakobsson, T. Kippola, and M. Matinmikko, "Interference measurements for Licensed Shared Access (LSA) between LTE and wireless cameras in 2.3 GHz band," in *IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN), 2015*, pp. 123–129, Sept 2015.

[84] M. Palolo, T. Rautio, M. Matinmikko, J. Prokkola, M. Mustonen, M. Heikkilä, T. Kippola, S. Yrjölä, V. Hartikainen, L. Tudose, A. Kivinen, J. Paavola, J. Okkonen, M. Mäkeläinen, T. Hänninen, and H. Kokkinen, "Licensed Shared Access (LSA) trial demonstration using real LTE network," in *9th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), 2014*, pp. 498–502, June 2014.

[85] A. Ponomarenko-Timofeev, A. Pyattaev, S. Andreev, Y. Koucheryavy, M. Mueck, and I. Karls, "Highly dynamic spectrum management within licensed shared access regulatory framework," *IEEE Communications Magazine*, vol. 54, pp. 100–109, March 2016.

[86] P. Masek, E. Mokrov, A. Pyattaev, K. Zeman, A. Ponomarenko-Timofeev, A. Samuylov, E. Sopin, J. Hosek, I. A. Gudkova, S. Andreev, *et al.*, "Experimental Evaluation of Dynamic Licensed Shared Access Operation in Live 3GPP LTE System," in *IEEE Global Communications Conference (GLOBECOM), 2016*, pp. 1–6, IEEE, 2016.

[87] Q. Zhao and B. M. Sadler, "A Survey of Dynamic Spectrum Access," *IEEE Signal Processing Magazine*, vol. 24, pp. 79–89, May 2007.

[88] A. A. Daoud, G. Kesidis, and J. Liebeherr, "Zero-Determinant Strategies: A Game-Theoretic Approach for Sharing Licensed Spectrum Bands," *IEEE Journal on Selected Areas in Communications*, vol. 32, pp. 2297–2308, November 2014.

[89] M. M. Buddhikot, "Understanding Dynamic Spectrum Access: Models, Taxonomy and Challenges," in *2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2007*, pp. 649–663, April 2007.

[90] R. H. Tehrani, S. Vahid, D. Triantafyllopoulou, H. Lee, and K. Moessner, "Licensed Spectrum Sharing Schemes for Mobile Operators: A Survey and Outlook," *IEEE Communications Surveys Tutorials*, vol. 18, pp. 2591–2623, Fourthquarter 2016.

[91] 3GPP Release 12, "Overview of 3GPP Release 12." `http://www.3gpp.org/specifications/releases/68-release-12`. [Online].

ABSTRACT

This dissertation thesis addresses the "Heterogeneous Connectivity of Mobile Devices in 5G Wireless Systems". Despite decisive progress in many enabling solutions, next-generation cellular deployments still suffer from a glaring lack of bandwidth due to inefficient utilization of radio spectrum, which calls for immediate action. The main aim of this Ph.D. work is to propose novel mechanisms providing proximity-based (cellular-assisted) networking and communication algorithms for dynamic allocation of spectrum resources in fifth-generation (5G) systems. Proposed communication mechanisms are comprehensively evaluated by the developed simulation tools (calibrated with the 3GPP data sets) as well as within the experimental 3GPP LTE-A cellular deployment at Brno University of Technology (BUT), Czech Republic. Obtained practical results (supported by novel mathematical analysis in characteristic scenarios) become instrumental to facilitate more dynamic bandwidth sharing. Thus, they promise to improve the degrees of spectrum utilization in future 5G systems without compromising the service quality and user experience (QoS and QoE) across different applications. As the main output of this thesis, particular research findings contributed in part to the 3GPP Release 12 specifications.

ABSTRAKT

Předkládaná disertační práce je zaměřena na "heterogenní propojení mobilních zařízení v bezdrátových systémech 5. generace". Navzdory nepochybnému pokroku v rámci navržených komunikačních řešení postrádají mobilní sítě nastupující generace dostatečnou šířku pásma a to hlavně kvůli neefektivnímu využívání rádiového spektra. Tato situace tedy v současné době představuje řadu otázek v oblasti výzkumu. Hlavním cílem této disertační práce je proto návrh nových komunikačních mechanismů pro komunikaci mezi zařízeními v bezprostřední blízkosti s asistencí mobilní sítě a dále pak návrh a implementace algoritmů pro dynamické přidělování frekvenčního spektra v nastupujících mobilních sítích 5G. Navrhnuté komunikační mechanismy a algoritmy jsou následně komplexně vyhodnoceny pomocí nově vyvinutých simulačních nástrojů (kalibrovaných s využitím 3GPP trénovacích dat) a zejména pak v experimentální mobilní síti LTE-A, která se nachází v prostorách Vysokého učení technického v Brně, Česká Republika. Získané praktické výsledky, které jsou podpořeny zcela novou matematickou analýzou ve speciálně navržených charakteristických scénářích, představují řešení pro vlastníka spektra v případě požadavků na jeho dynamické sdílení. Tato metoda tedy představuje možnost pro efektivnější využití spektra v rámci mobilních sítí 5G bez degradace kvality služeb (QoS) a kvality zážitků (QoE) pro koncové uživatele. Vědecký přínos dosažených výsledků dokazuje fakt, že některé z principů představených v této disertační práci byly zahrnuty do celosvětově uznávaného standardu (specifikace) 3GPP Release 12.