CHAPTER

5G and beyond networks

6

Silvia Ruiz[a,b], **Hamed Ahmadi**[c,b], **Gordana Gardašević**[d], **Yoram Haddad**[e], **Konstantinos Katzis**[f], **Paolo Grazioso**[g], **Valeria Petrini**[g], **Arie Reichman**[h], **M. Kemal Ozdemir**[i], **Fernando Velez**[j], **Rui Paulo**[j], **Sergio Fortes**[k], **Luis M. Correia**[l], **Behnam Rouzbehani**[m], **Mojgan Barahman**[m], **Margot Deruyck**[n], **Silvia Mignardi**[o], **Karim Nasr**[p], **and Haibin Zhang**[q]

[a] *Universitat Politècnica de Catalunya, Barcelona, Catalonia, Spain*
[c] *University College Dublin, Dublin, Ireland*
[d] *University of Banja Luka, Banja Luka, Bosnia and Herzegovina*
[e] *Jerusalem College of Technology, Jerusalem, Israel*
[f] *European University Cyprus, Nicosia, Cyprus*
[g] *Fondazione Ugo Bordoni, Rome, Italy*
[h] *Ruppin Academic Center, Hadera, Israel*
[i] *Istanbul Medipol University, Istanbul, Turkey*
[j] *Universidade Beira Interior, Covilhã, Portugal*
[k] *Universidad de Malaga, Malaga, Spain*
[l] *INESC-ID / IST, University of Lisbon, Lisbon, Portugal*
[m] *Universidade de Lisboa, Lisbon, Portugal*
[n] *Ghent University, Ghent, Belgium*
[o] *University of Bologna, Bologna, Italy*
[p] *University of Greenwich, London, United Kingdom*
[q] *TNO, The Hague, Netherlands*

6.1 Introduction

The aim of this chapter is to investigate the Network Layer aspects that will characterize the merger of the cellular paradigm and the IoT architectures, in the context of the evolution towards 5G-and-beyond.

In recent years, researchers have been working to identify and assess the network architecture of 5G-and-beyond systems; studying the impact of the "fog" networking/computing approach foreseen for 5G on the evolution of the radio access technologies; evaluating RRM approaches compatible to the new requirements; proposing new concepts and paradigms to take into account the plethora of new applications arising from the IoT context. The different chapter sections structure and summarize the main contributions on these topics.

[b] Chapter editors.

Inclusive Radio Communications for 5G and Beyond. https://doi.org/10.1016/B978-0-12-820581-5.00012-2

In this chapter we first will have a look at the issues from previous generations carried to 5G and then introduce more recent challenges mainly concerning 5G and beyond. We first review recent contributions on ad-hoc networks and then continue with spectrum management, resource management, and scheduling.

Small cell networks, Heterogeneous Networks, and network densification are key issues of 5G. We look at different aspect concerning them and summarize the theoretical and hands-on contributions in these areas. Cloud-based densification is a 5G technology which has attracted a lot of attention in recent years. We have a section on Cloud Radio Access Network (C-RAN) which will be followed by a section on Software Defined Network (SDN) and Network Function Virtualization (NFV). In the later section we discuss network virtualization and softwarization, its challenges and recent contributions.

Finally the chapter includes two sections devoted to promising technologies, and emerging services and applications that will become more and more important in next years as: the use of Unmanned Aerial Vehicles (UAVs) as Unmanned Aerial Base Stations (UABSs), real-time smart grid state estimation, vehicle to vehicle and V2X communications to improve road safety and efficiency, or new public protection and disaster relief wireless networks standards.

6.2 Ad-hoc and V2V networks

One of the critical questions, when it comes to V2X communication, is which communication technology to use. With the rise of autonomous vehicle this is a critical issue which requires the best possible communication means both in term of throughput and latency. There has been many standards proposed for V2X, but among the most famous we can mention IEEE 802.11p which is a dedicated standard for wireless access in vehicular environments (WAVE) in the licensed 5.9 GHz band. Another possibility could be the use of the legacy overlaying LTE cellular technology. However, due to possible overload in classical cellular networks this should be done in a wise manner. One possible implementation is to combine both the flexibility and low latency of the ad-hoc direct vehicular communication of IEEE 802.11p with the wide range coverage and high data rate which can be potentially realized with LTE. In [DMM+16], a dynamic data traffic algorithm to enable hybrid vehicular communication is proposed. To steer the traffic between the cellular LTE network and the ad-hoc 802.11p communication a novel steering algorithm was developed that is based on an optimized utilization of the dedicated spectrum of each system. The goal of this algorithm is to react on an impending 802.11p channel congestion as well as on an overloaded LTE cell to which the vehicle is connected to. The proposed solution was implemented in a system-level simulator and evaluated using a realistic simulation scenario set in the city of Doha, Qatar where a road congestion at an urban intersection is emulated.

When considering V2V, one frequent question arises about the way it can be simulated. In [BBG+18] an accurate simulator for VANET is proposed. Classical

platforms rely on many assumptions that fail to catch the right mobility of the vehicles and some basic physical layers parameters. For instance one of the most popular approaches is to treat the vehicles as a Poisson Point Process (PPP) along a straight line, resulting in uniformly distributed vehicles along the line whereas we need to account for inherently inhomogeneous vehicle distributions. Thus, in [BBG+18] the mobility simulator SUMO is used in combination with the vehicular network extension VEINS of the network simulator OMNET++. In addition, Karedal's pathloss model is used to calculate communication ranges.

Within the framework of V2V and ad-hoc networks, the topic of Delay/Disruption-Tolerant Network (DTN) is frequently mentioned and considered as a challenge since they have dynamic topology with frequent disconnections. Due to this sparse and intermittent connectivity, inference and learning over DTNs is much more complicated than in traditional networks. In [LBC+16], the authors consider the problem of distributed defective node detection in DTNs. A node is considered as defective when one of its sensors frequently reports erroneous measurements. The identification of such defective nodes is very important to save communication resources and to prevent erroneous measurements polluting estimates provided by the DTN. It is assumed that nodes are not aware of the status (good or defective) of their sensors. Distributed Fault Detection (DFD) for Wireless Sensor Networks (WSNs) has been well-investigated; however, the networks considered in most of the literature are dense and have a static topology whereas DFD in DTNs has been under-investigated. In [LBC+16], the authors present a fully distributed and easily implementable algorithm that allows each node of a DTN to determine whether its own sensors are defective. The theoretical results provide guidelines to properly choose the parameters of the algorithm. In the simulations, a jump motion model and a Brownian motion model are considered. The results show a good match with theory.

6.2.1 Prediction and reliability

We are all interested in ubiquitous connectivity but without interruption and disconnection. To avoid these kinds of situations the cellular operators can perform some important tasks. First, the network has to be designed to be reliable. For instance, a single point of failure should be avoided. Then prediction should be used to know ahead of time that overload and lost of coverage can occur in order to trigger in advance a handover and/or strengthen the deployment.

Network topologies are sometimes assessed according to their resilience to failure. For instance, a star topology with a central node is considered as very vulnerable whereas a tree topology is better while some aggregation links should be doubled (also known as 1+1 links) in order to ensure connectivity if one link fails. One of the most resilient topologies is the ring topology. The reason is that even if a link fails the node is not completely disconnected since there is always a second link to which the node is connected. However, one important and complex issue is how to split the nodes in the network into several rings to be connected to the aggregation node. This NP-hard problem is considered in [EHD17]. The authors considered two different

cases. First, when the probability of failure is the same for each link, the authors propose an optimal solution where the main condition is that the empirical variance of the ring size distribution should be as small as possible. Then, when it is assumed that links may have various failure probabilities, the problem is further split into two different cases. The first one is when fixed 3-nodes ring sizes are assumed for each ring. An optimal solution is proposed based on a search of a perfect matching in a weighted graph. While in the more general cases with ring of different sizes, which is proved to be NP-complete, some approximation methods are proposed.

In [MHRA18,RHMA18], it is proposed to use Machine Learning theory to predict the wireless coverage in any given point in space. Up to now, cellular coverage is predicted by cellular operators thanks to propagation model and radio planning tool based on ray-tracing. These tools are however limited when networks are ultra dense. Moreover if we consider the M2M scenario to be incorporated with future 5G, this kind of planning is almost impossible. In these works, the authors proposed to gather sample SINR measurements from smartphones and/or dedicated low cost IoT devices. Then, thanks to Artificial Neural Networks based on radial basis function activation, the SINR can be predicted in other places. This kind of methods proved to be very accurate and can be run inside a smartphone, since machine learning processing is now incorporated in the chip of smartphones. Another kind of prediction is considered in [Bit19] with a mmWave mesh networks. Cognitive Radio capabilities are assumed where the goal is to predict SINR at the primary users where interference is generated by secondary users. Thanks to SINR prediction, some pre-processing over the interference (such as interference cancellation) can be performed. The SINR can be predicted using a Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) method.

Given the availability of sources that can provide detailed information about social events, and the interest of using this in the management of cellular networks, in [FPSB18b] a novel general framework for the automatic acquisition of events data, its association with cellular network information, and its application in Self Organizing Networks (SON) is proposed.

Recently, the direct communication between mobile terminals, also called D2D, has attracted a lot of attention as an extension of the conventional connectivity. The aim is to increase the range of services, making it possible to lower the load offered to networks and to provide coverage extension (via relay or other means). [IOS+19] addresses a time-dependent analysis of the Signal-to-Interference Ratio (SIR) in D2D communications scenarios in order to build a uniform theory of mobility-dependent characterization of wireless communications systems. Analytical expressions for the variance and Coefficient of Variation (CoV) of SIR for a link between two communicating devices in D2D scenarios have been derived which use the kinetic-based mobility model capable of capturing a wide variety of users' movement pattern. Moreover, the effect of the variance and CoV on link stability in terms of the number of outages for non-stationary mobility patterns of users has been investigated.

6.2.2 Network simulation/emulation platforms

The use of tactical radios to communicate in case of emergency and rescue situation is a typical real use case of Mobile Ad-Hoc Networks (MANETs). Due to the operating environment, these radios experience specific characteristics such as limited bandwidth, in addition to high delay and packet losses. In order to assess the value of the optimizations and the performance of the new generation radios, many open-source and commercial network simulators/emulators exist to evaluate the Quality of Experience (QoE) of the different services. Authors in [NPJ15] provide a survey on best known emulation testbeds for MANETs. They outline differences between simulators and emulators and the complexity of deploying and managing emulators for research purposes. A comprehensive list of frameworks can also be found in [Ope]. However, most of these frameworks do not offer out-of-the-box features that allows for the use of customize proprietary applications in the system or means for quickly reproducing emulated experiments in the field. Authors in [RGW+18] present the work achieved to develop, evaluate and test a novel application-layer routing algorithm specifically designed for tactical MANETs. In order to test and evaluate the proposed system, they developed two different platforms. The simTAKE platform has been developed to facilitate the development and analysis of the results in laboratory conditions, without the need to deploy the system on the field, which is very time-consuming and might produce almost no reproducible results. The platform provides the possibility to set the number of communication systems (called nodes) deployed in the scenario, their topology and the parameters of the network such as data bandwidth, packets delay, and packet drop rate as well as the distribution of messages emitted by each node during a simulation. The emulTAKE platform offers a system intended for the next level of tests, namely under field conditions. It allows to control the emulation of users for each node. With this platform, the field test manager can define scenarios, deploy them to each node, in order to generate messages according to specific distributions. Finally, results can be collected once a simulation is completed from a single web page. In order to evaluate the proposed algorithm in real conditions, two platforms and specific QoE metrics have been developed. This contribution might be of interest to those involved in V2V communications, especially the emulation of such networks. In order to emulate the MANET's radio performance in a more realistic way, authors in [NWB+18] used the Extendable Mobile Ad hoc Network Emulator (EMANE) framework to reproduce the behavior of real tactical radio benchmarked in a laboratory. This was done by measuring the throughput and round-trip time of real tactical radios using wideband or narrowband Time Division Multiple Access (TDMA)-based waveforms. The wideband and narrowband measured performances were reproduced very closely using the same 2 second multiframe (M1D7) TDMA structure as the radios under test. The rate adaptation featured by the real radios could also be reproduced by changing the slot data rate using a Python agent external to EMANE. All contributions to EMANE and OSLRd1 required to obtain the presented results are available in a gitlab repository [Pre]. Authors in [NBB19] extend their previous research in order to investigate the performance results (Round Trip Time and message Completion Rate) of a 24-nodes MANET using TDMA radios and OLSR

in a real time emulator based on the open-source EMANE framework. This contribution presents EMANE based emulation results using realistic TDMA radios including the use of fixed data rate or a rate adaptation scheme. The open source OLSRd1 or OLSRd2 implementations have been used.

Detection of social relations ties among nodes is important because of mobile devices in the MANET-DTN are carried by humans which are social creatures that live and move in social groups. For those reasons, mobility models should follow human behavior. Authors in [HMDP18] propose tools for evaluation of social relations from movement in mobility models which are described as new evaluation methods. Based on the proposed evaluation methods, it is possible to decide when the used mobility model shows signs of social behavior. The proposed evaluation method was created based on the Louvain method for community detection and other network graph parameters (average weighted degree). Simulations of evaluation methods were made as a comparison between two random mobility models and one social based mobility model. All models were simulated with a different number of nodes and radio ranges and evaluated by proposed method and other existing protocol-dependent and independent methods. Results of simulations prove that social based mobility models should have modularity quality values above 0.3, which is a good indicator of significant community structure. On the other hand, small values of modularity quality were obtained when random mobility models were used.

In the last years WiFi has seen many improvements which gave rise to new IEEE 802.11 standards such IEEE 802.11ah specification (also called "Wi-Fi HaLow") [IEE17] which plays a key role in the integration of a Wi-Fi oriented technology to operate in TV white space spectrum (TVWS) in the UHF band. Most of existing simulators have been developed with specific networks protocols in mind. In [MPSK17] a novel and universal link-level simulator for well known IEEE 802.11g/n/ac and also for upcoming 802.11 ah/af/ax has been developed. Basic functions of the simulator for the IEEE 802.11n standard were verified by simulations. The functionalities of the proposed simulator were demonstrated by several simulation outputs. The advantage of the proposed WLAN simulator, besides supporting wide range of IEEE 802.11 technologies, is also that is compatible with Vienna LTE/LTE-A simulator [Vie] which allows to investigate possible coexistence scenarios which can occur between these systems in shared ISM bands (2.4 GHz and 5 GHz), which will likely happen in the forthcoming 5G networks.

6.2.3 Energy harvesting

Typically, a Wireless Sensor Network (WSN) comprises a number of sensor nodes deployed in a given geographical area which collect and send measurements to a Fusion Center (FC). Current technological advances make it feasible to deploy inexpensive sensors in large numbers. In this context, the problem of optimally selecting a subset of sensors to perform a given task naturally arises. This often stems from resource (e.g., bandwidth), interference level or energy consumption constraints, which make massive sensor-to-FC communications barely recommended (or simply not possible).

In the literature, this is referred to as the sensor selection problem. Besides, Energy Harvesting (EH) is becoming a promising technology capable of extending the operational lifetime (or even allowing self-sustainable operation) of WSNs. Authors in [CMA16] have proposed a novel EH-aware sensor selection policy that outperforms a well-known EH-agnostic policy. A selection is needed due to the reduced number of available sensor-to-FC channels. The goal is to minimize the distortion in the reconstruction of the underlying source at the FC, subject to the causality constraints imposed by the EH process. Performance in terms of reconstruction distortion, impact of initialization, actual subsets of selected sensors, and computed power allocation policies is assessed by means of computer simulations. To that aim, an EH-agnostic sensor selection strategy, a lower bound on distortion, and an online version of the SS-EH and JSS-EH schemes are derived and used for benchmarking. The proposed sensor selection policy attains a lower bound on the reconstruction distortion of the optimal (joint) solution, as soon as the percentage of selected sensors is above 40%.

6.2.4 Measurements for specific applications

Operational activities of public entities, like border patrols, force development of communication solutions. The reinforcement of radio-communications or IT (Information Technology) systems increases the security of the officers on duty and allows to carry out all activities more effectively. Authors in [KRCJS17] provide results of research conducted to test, design, and build digital radio link for high-speed multimedia data transmission in maritime environment. The measurement campaign was carried out in the Vistula River Lagoon area near the Baltic Sea coastal waters in Poland. In addition, to verify the behavior of the developed digital radio link in real maritime environment, the selected measurement location is characterized by the possibility to ensure the LOS (Line of Sight) propagation conditions over a distance of several kilometers. It must be noticed, that entire radio link is realized in SDR (Software Defined Radio) technology and each change must be deeply verified. The key change was the hardware modification in USRP (Universal Software Radio Peripheral) devices. The positive results of the measurements verify and confirm the rightness of some minor changes in the software and hardware layer of the realized technological demonstrator.

6.3 Spectrum management and sharing

6.3.1 IoT/machine type communications

IoT has been widely used as a new technology that offers a revolutionary way to interact with the environment connecting objects and devices to improve the everyday lives of citizens around the world. ITU estimates that 25 billion devices will be connected to the Internet by 2020. Nowadays, new research is being carried out to define the technology and frequency bands for IoT. The requirements expected to be fulfilled are the support a massive number of devices, at low transmission rate

and reduced power consumption while maintaining long battery life. In addition, it is expected that there will be reduced complexity with extended coverage and limited latency and prices that minimize the cost of the equipment and maintenance [Eva15]. For the operation of such a large number of devices, [MPLC16] proposed the use the TV White Spaces (TVWS) of the Digital Terrestrial Television (DTT) band through spectrum sharing between DTT and IoT networks. The aim of the work presented in [MPLC16] was to evaluate whether the coexistence of both technologies in co-channel or adjacent-channel is possible by determining the margins of protection and EIRP that could be transmitted by base stations and IoT devices. Authors considered the Digital Video Broadcasting Terrestrial Access 2 (DVB-T2) standard as the most widely used in the world. The proposed scenario considered a DVB-T2 network offering fixed rooftop reception as a primary service and Narrow Band LTE (NB-LTE) network IoT as a secondary service. Five representative scenarios to deploy IoT were evaluated while both uplink and downlink were considered. The first two scenarios considered urban or suburban environments characterized by buildings of 2 or 3 floors with wide and narrow streets looking into smart parking and traffic congestion. The third and fourth scenarios considered smart agriculture looking into smart farming and animal tracking. Finally, the fifth scenario was about eHealth – Indoor, looking into patients' surveillance inside a hospital. Results demonstrated that the coexistence between DVB-T2 and NB-LTE-IoT is feasible in adjacent channel. Measurements set the maximum allowable EIRP for a NB-LTE Small Cell between 10 and 15 dBm, considering a 1 MHz band guard. Hence, NB-LTE could be efficiently used as a secondary service when having a 6 or 7 MHz white-space between two DTT channels. The maximum EIRP that could be transmitted by a NB-LTE device varies in a range between 9 and 14 dBm for the best scenario (Smart Parking), whereas this power is reduced to a range of 3 and 8 dBm considering the worst one (Traffic congestion), from a 2 MHz guard band. Considering high duty cycles and a 2 MHz guard band, the maximum EIRP could be reduced up to 4 dBm if the waiting time between transmissions is lower than the duration of the DVB-T2 frame (250 ms). For a 6 MHz guard band, this impact is almost negligible. Given that there is an inversely proportional dependence between throughput and battery life, for a specific waiting time between transmissions and long transmission time high throughputs and short battery life are achieved. Authors in [US17] describe how 5G networks should be able to support Machine Type Communication (MTC) traffic generated by the IoT as well as traffic generated from the traditional Massive Broadband (MBB). They also state that both types of traffic can be achieved by utilizing media access schemes based on Random Access (RA) while resource partitioning is necessary to avoid collisions between both traffic types. The paper is proposing a dynamic partitioning scheme that utilizes a control loop to manage the amount of reserved radio resources for MTC traffic. Authors employed Slotted ALOHA to exploit its characteristics for adjusting the amount of available Small Packet Blocks (SPB) for the one-stage access. This introduced discrete time-slots where IoT devices could only try to access the medium at the beginning of a time-slot. This reduced collisions and therefore increased the maximum throughput. The control loop was then used to measure the average collision

probability of the RA based MTC traffic and estimated the MTC traffic load. More specifically, the control loop observed the SPB collision probability and adjusted the amount of SPB such that a certain target SPB collision probability with a hysteresis of 2.5 percent was achieved. This estimate was then used to control the amount of reserved radio resources. The performance of the control loop was evaluated using constant as well as variable MTC traffic load. Results suggest that the proposed control loop can be applicable for low load situations with small load variations. As the medium access scheme was Slotted ALOHA, the target SPB collision probability of the control loop determined the operation point of the system. The paper concluded that there are three possible target Packet Block (PB) collision probability regions which result in different operation points:

$< 44\%$: more successful transmissions, bad resource utilization
$= 44\%$: less successful transmissions, best resource utilization
$> 44\%$: increasing SPB collisions, should be avoided.

Furthermore, the control loop achieves a better control quality for short optimization interval duration, but too short duration leads to many SPB reconfiguration and thus to decreased control quality and throughput.

Another contribution [MHD+17] compares non-intelligent and intelligent methods of channel assignment and their influence on optimal path selection in the process of message delivery from the source to the destination. The work focuses on finding the path between the source and the destination node in a Cognitive Radio Mobile Ad-Hoc Network (CR-MANET) environment, where the single channel and multichannel assigning were implemented with and without primary user (PU) activity. The paper describes how Cognitive Radio (CR) technology is implemented to Mobile ad-hoc network and how spectrum sensing is an important process before channel assignment. Authors present mobile nodes in CR-MANET as self-organized, decentralized, and autonomous routers, which can select different frequency channel between a couple of nodes

6.3.2 Coexistence and sharing

The advent of 5G will cause a massive increase in required spectrum, that will be necessary to accommodate new, bandwidth-hungry applications. For this reason, alongside bands already assigned to mobile services, new spectrum chunks have been identified for 5G at both European and global level. In this complex scenario, coexistence must be ensured both with other services in adjacent bands as well as with legacy mobile communications systems using the same bands in neighboring countries, given that the time schedule for 5G roll-out might differ significantly among different countries. The latter issue is particularly significant in Europe, owing to the high number of nations in a relatively small area, often sharing terrestrial and maritime borders with several neighbors.

The issue of coexistence has, naturally, attracted a lot of interest within the COST-IRACON community, which is reflected in the remarkable number of documents

devoted to the subject, that in some cases evolved in papers published in scientific journals.

An analysis of various 5G bands may be found in [CGMP18], which examines the three pioneer bands identified to offer 5G services in Europe by 2020 on a large scale, namely the 700 MHz, the 3.5 GHz, and the 26 GHz band, using for each band appropriate propagation models.

In regards to the 700 MHz band, in [CGMP18] the focus is on the study of the technical conditions that allow the coexistence between the victim LTE M2M base station in the 733-736 MHz and the interfering LTE Supplemental Downlink (SDL) base station in the 738-758 MHz. Results are derived in terms of separation distances, that guarantee the Interference to Noise Ratio (INR) requirement, as a function of SDL transmitted EIRP. Separation distances ranges between 0-50 km for [IR16], 0-2.5 km for [DC99], 0-18 km for [IR09], and 0-5 km for [IR13].

The 26 GHz analysis in [CGMP18] studied the impact of an International Mobile Telecommunication 2020 (IMT-2020) BS interferer on a victim incumbent Point to Point (P-P) link for the propagation model defined in ITU-R P.452-16 [IR15] in combination with ITU-R P.2108 [IR17]. It considered an IMT-2020 outdoor urban hotspot base station and three Fixed Service (FS) receivers at different heights (15 m, 30 m, and 60 m). The minimum separation distances that allows to meet the constraints on INR are about 27-44 km for the different FS antenna heights. However, the FS antennas at 26 GHz are strongly directive, so that even a small angular separation between the FS and the P-P radiation pattern will considerably reduce the protection distance.

It is well known that the Radio Spectrum Policy Group (RSPG) of the European Commission considers the 3400-3800 MHz band to be the primary band suitable for the introduction of 5G-based services in Europe even before 2020, noting that this band is already harmonized for mobile networks, and consists of up to 400 MHz of continuous spectrum enabling wide channel bandwidth [EC16]. However, there are some issues related to this band, owing to the fact that it can be used also by other services, and in particular FS and Fixed Satellite Service (FSS). In [CGMP18] tests for the 3500 MHz band analyzed the coexistence between an interferer LTE eNodeB and a victim incumbent FS for both co-channel and adjacent channel interference scenarios and various cell types considering the transmitted parameters suggested by [ECC13]. In [CGMP18] results are provided for different propagation models and for different combinations of horizontal and vertical angular discrimination. Angular discrimination considerably reduces the separation distances. As an example, in a Macrocell scenario the separation distance reduces from about 120 km for 0° of horizontal angular discrimination to about 23 km for 30° of horizontal angular discrimination.

Concerning the FSS in the 3400-3800 MHz band, a remarkable example is the reception of broadcast television by means of Very Small Aperture Terminal (VSAT), widely used in Africa, Asia, and Latin America. The issue of coexistence between land mobile services (and particularly LTE Time Division Duplex (LTE-TDD) and VSAT-based television reception) is analyzed in [CGP+17], which presents the re-

sults of an extensive measurement campaign and laboratory tests. Proper separation distances for both co-channel and adjacent channel were identified in different real propagation conditions, LOS and NLOS. The characteristics of VSAT equipment, studied by laboratory measurements, show a relatively small variation in selectivity between different receivers.

In addition to the three pioneer bands identified to offer 5G services in Europe other frequency bands have been analyzed for a possible introduction of future mobile systems. In [SSTV18] the viability of 5G New Radio (5G NR) spectrum sharing in Ultra High Frequency (UHF), Super High Frequency (SHF), and mmWave in outdoor environment is investigated. Performance evaluation includes the study of the behavior of Physical Layer (PHY) throughput as a function of the coverage distance for 2.6, 3.5, 28, 38, 60, and 73 GHz. The preliminary analysis shows that the highest system capacity and the highest modulation and coding schemes are achievable for the shortest coverage distances at mmWave whereas the supported throughput for long coverage distances is more favorable for UHF and SHF bands.

Even the 2.3-2.4 GHz band attracts a lot of interest for 5G deployment. However, this band is used in various European countries for incumbent services such as FS, Programme Making and Special Events (PMSE), Governmental use and WiFi above 2.4 GHz. Protection shall be guaranteed to all the incumbent users. One promising technique towards this aim is Licensed Spectrum Access (LSA). To investigate its potential, Italy hosted a world-first LSA pilot, promoted by the Italian Public Administration and the Joint Research Centre of the European Commission under the technical coordination of Fondazione Ugo Bordoni. It counted on the participation of research and industrial partners from Italy, France, and Finland. The achievements of the pilot were reported in [GCP+16] and in [GCP+17]. Seven LTE-TDD radio base stations (BS) at 2.3-2.4 GHz have been installed both indoor (5 Femto BSs) and outdoor (2 Macro BSs) in the building of the Italian Ministry of Economic Development (MiSE). Fig. 6.1 shows the LSA Pilot end-to-end architecture. The various elements of the pilot were located around Europe. The indoor and outdoor LTE nodes located in Rome were connected to the Evolved Packet Core (EPC), which allowed the communication toward user equipment (i.e. commercial smartphones equipped with authenticated test SIMs). An Operation Administration & Management (OAM) communicated with the LSA controller (both elements were located in Finland) and was capable of managing the mobile network in order to cope with the requirements imposed by the sharing rules stored in the LSA repository, located in France. The Sharing Tool used to determine the sharing rules was in the administration domain.

Tests were made about the coexistence between LTE-TDD base stations and two types of incumbent services, namely FS and PMSE. Coexistence is achieved thanks to sharing rules properly identified based on the incumbent services characteristics and tested in a real scenario. Compliance with the sharing rules for the FS has been verified on the field through measurements, both for the protection zone and restriction/exclusion zone approach.

Subsequently, tests on channel evacuation were realized considering a possible incumbent PMSE user, requesting a channel for its operations in a given location.

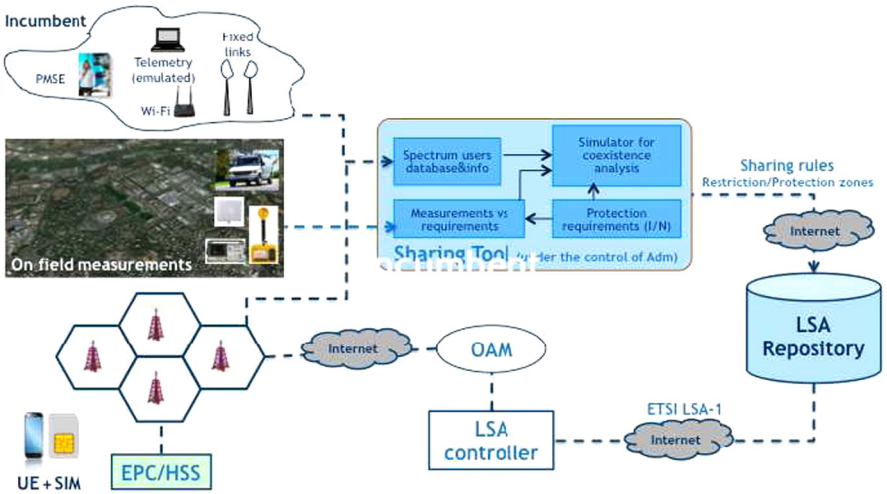


FIGURE 6.1

Italian LSA Pilot end-to-end architecture.

Results show that it was always below 40 seconds. Results from the pilot indicate that LSA is a straightforward approach that provides high predictability and certainty for both the incumbent and the licensee while preserving existing ecosystem and business models. Space limitations don't allow to provide a detailed account of tests results here, and the interested reader is referred to the quoted references for details.

Also related to the 2.3-2.4 GHz band was the experimental campaign carried out in Italy in response to the invitation of the European Commission to the Member States to perform specific coexistence analysis on developing conditions for the introduction of Wireless Broad Band (WBB) in this band. Both on-air measurements in an indoor environment and laboratory tests were performed aiming at investigating the possible interference towards WiFi due to out-of-band emissions from LTE equipment in the adjacent bands. Results of this activity were reported in [GCC+17]. The main trends in terms of interfering effects of the LTE signals are:

- The effect of LTE interference decreases as frequency separation between the interfering and victim signals increases.
- The interfering effect of LTE Down-Link (DL) onto WiFi victim receivers is stronger than LTE Up-Link (UL).
- Although different WiFi equipment may have different performance (e.g. in terms of throughput), similar effects of the LTE interference can be highlighted due to LTE signals (e.g. effect on different WiFi channels, effect of LTE DL vs LTE UL, etc.).

The 2.4 GHz band is also of interest for the future introduction of Low Power Wide Area Network (LPWAN) in the ISM band. The coexistence between LTE and

LPWANs has to be analyzed. In [PPMK19], experimental laboratory measurements were realized in order to test the robustness of the LTE-DL full data load to interference, caused by a Long Range (LoRa) signal, in the 2.4 GHz band. Experiments were performed in order to determine C/I ratios needed to provide LTE-DL service with good quality. Guard bands were varied between 1 MHz and 5 MHz. The obtained results show that 64QAM modulation has the lowest robustness against interfering LoRa signal. It was also observed that a LTE-DL signal with wider bandwidth is more sensitive on the interfering LoRa signal.

Another dynamic spectrum sharing approach in a quite different context is presented in [Lys17], which contains an overview of the achievements and ongoing work in the domains of TVWS, Dynamic Spectrum Access (DSA), compatibility and protection against interference, and smart antennas in South Africa. The rationale behind the work is the need to provide wireless Internet access, especially in rural areas characterized by a sparse population distributed over a large territory. This situation is common also to several other African countries.

The South African government set the target of universal broadband coverage, including broadband access for every school in the country. The identified solution was to exploit the TVWS in a cognitive way, by means of DSA. To this aim, the Council for Scientific and Industrial Research (CSIR) built a Geolocation Spectrum Database (GLSD). This approach was validated since 2013 in a number of large scale trials of TVWS technology, together providing Internet connectivity to over 20 schools with over 20,000 pupils and teachers. They were also followed by trials in other African countries, such as Ghana and Botswana. The trials included compatibility testing and demonstrated the viability of the approach based on DSA and using the GLSD to exploit TVWS.

Another key issue to be considered is the coexistence between mobile radio and other services such as Public Protection and Disaster Relief (PPDR). In Colombia, as well in other and other American countries, the 806 to 824 MHz and the 851 to 859 MHz bands are allocated for PPDR services (UL and DL respectively), with a combination of technologies like P25, TETRA, and LTE-PPDR, creating a scenario where technologies with different bandwidths and technical characteristics must coexist. At the same time, the lower PPDR band is adjacent to the 700 MHz LTE band, while the upper PPDR band is adjacent to the 850 MHz LTE band. The issue is further complicated by the fact that both narrowband and wideband PPDR systems are in use and it is necessary for the regulator to guarantee their coexistence. An experimental study, based on both laboratory testing and simulations, is reported in [NCM17].

The authors show the results related to three different scenarios, considering both interference between PPDR and LTE bands, and internal interference between PPDR systems using the same band. Key findings are that it is possible to reduce the guard band between LTE 700 MHz and PPDR systems by up to 2 MHz, while more critical cases are located in the upper part of the 850 MHz band where the LTE-DL is adjacent to the PPDR-DL.

6.3.3 Field monitoring

One practical problem to be addressed in planning and rolling out a new network lies in the field exposure limits that may prevent adding new radiating elements, e.g. for 5G, in sites where already other radio technologies are present.

This is further complicated by the fact that there is no harmonization between exposure limits throughout Europe (let alone worldwide). Therefore, operators with a wide international footprint, which is the case for various European operators, must face the additional challenge of adapting their roll-out strategies to regulations changing from country to country and, in some cases, also at regional level within the same country.

Finally, the advent of new technical features, such as Massive MIMO and Carrier Aggregation, will also require an adaptation of field measurement criteria and equipment, as well as of computation algorithms, currently in use.

Therefore, having reliable and accurate measurement and simulation procedures is necessary in order to correctly plan and roll-out new systems and networks. This issue was dealt with by various contributions presented at IRACON meetings ([GCV+16], [CGV+19], [CPB+19]).

The authors of [GCV+16] presented both Narrow-Band (NB) and Wide Band (WB) measurements taken during the operation of the above mentioned LSA pilot. Measurement were performed both indoors and outdoors and over the frequency band from 87.5 MHz to 3 GHz, in order to include FM radio and terrestrial TV broadcasting along with mobile and wireless systems. Both the NB and the WB measurements showed that the field levels were always well below the limit of 6 V/m even when the LTE-TDD femtocell network was deployed at 2.3 GHz. In particular, the highest measured values were slightly below 1.2 V/m and results were corroborated by simulations. As an example, Fig. 6.2 shows indoor measurement results for three bands in several locations within the premises of the Italian Ministry of Economic Development. Additional results can be found in the paper.

The introduction of Massive MIMO causes a high variability of exposure levels, owing to the usage of smart antennas. Therefore, the actual exposure levels are lower than those experienced with traditional antennas, and the methodologies to assess them (both measurements and simulations) must be adapted accordingly [CPB+19]. A method adopted in Italy defines a reduction factor named “Alpha24” that evaluates the average level of electric field during 24 hours of a base station under test. This methodology allows a more realistic evaluation of actual exposure levels and, therefore, allows the introduction of base stations using Massive MIMO even under very restrictive limits.

A related contribution [CGV+19] presents a methodology for carrying out large-scale measurements of radiofrequency electromagnetic field and low frequency electric and magnetic field. The system consists in performing measurements using a control unit equipped with the appropriate probes positioned on the roof of a car moving along the territory (dynamic measurements). The possibility to perform massive measurements in a short time could become extremely useful in the future, when the implementation of the new 5G services will lead to an increased use of the

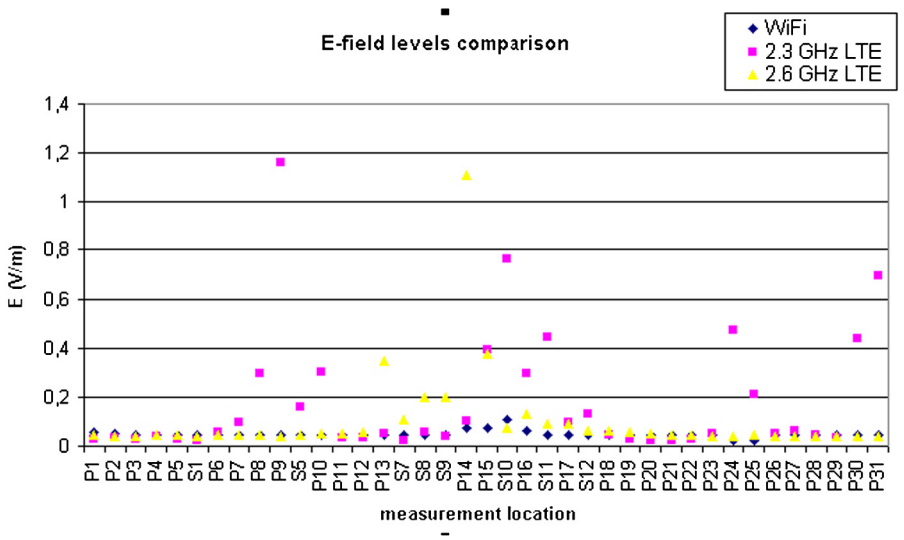


FIGURE 6.2

Measured E-field comparison.

electromagnetic spectrum, with consequent increase of electromagnetic field levels. Measurement activities carried out in different contexts and scenarios showed an excellent agreement between dynamically measured radio frequency electric field levels and the Electro-Magnetic Field (EMF) levels obtained by carrying out measurements with static probes. This agreement, makes it theoretically possible to perform large-scale radiofrequency field monitoring using the dynamic procedure, thus making the task much quicker and straightforward.

6.3.4 Virtualized networks

Wireless network virtualization is a key technology for 5G, both as an enabler for operational sustainability and business agility [AMG+16]. Wireless network virtualization requires the abstraction of physical resources such as spectrum, infrastructure, power, etc., into virtual resources. These resources can then be partitioned and allocated to different Virtual Network Operators (VNOs). Authors in [AMG+16] propose a virtualized network architecture for an infrastructure provider that shares the physical resources of a Massive MIMO cell among several VNOs using spatial multiplexing as well as orchestrate the interaction between the VNOs and the infrastructure provider through an auction-based mechanism that allocates spatial streams to the VNOs.

Furthermore, authors in [AAC+17] present a customizable resource virtualization algorithm for multi-user data scheduling in a LTE C-RAN deployment. The algorithm is based on the hypervisor specific dynamic assignment of air resources to the VNOs,

based on either joint scheduling or per-cell schemes. The objective is to improve the resource allocation mechanism for two scenarios linked with night-time and day-time.

Another contribution [CB18] presented the detection methods in the energy-efficient context in 5G networks. It was shown that the considered context-awareness information has enormous volume. Simulations performed showed that a power consumption reduction is possible if intelligent network organization is implemented. Authors suggest that the correlation between nodes used for finding clusters in the network allows to reduce the information exchanged in the network. However, mobility of nodes limits the clustering gain since the correlation in the network is limited in high-velocity scenario.

Finally, contribution [PKM18] presented the importance and the benefits of employing a geo-location spectrum database for deploying a TVWS network as a vital tool to limit harmful interference to primary TV spectrum users. This work outlines how a fully operational web-based TVWS system has been designed and developed for Cyprus using the web API framework and adhering to the methods defined by CSIR.

6.4 Radio resource management and scheduling

RRM refers to the efficient use of radio resources so that the system capacity is kept at maximum and the Quality of Service (QoS) of the users are met. Any mobile communication system should be able to offer services with different QoS requirements depending on each user subscription level. To fulfill different request service types such as voice, video streaming, web-browsing, etc., a mechanism to classify those types of bearers into different QoS Class Identifier (QCI) is needed. 3GPP specifies that each QCI is characterized by priority, packet delay budget, and acceptable packet loss [GCS17].

Due to the huge increase in traffic and services in mobile networks, network management has defined an alternative quality indicator known as QoE, which is defined as the overall satisfaction of the user of a service as it is subjectively perceived. QoE is widely measured using the Mean Opinion Score (MOS) scale, in which a particular value is assigned to the experience perceived according to the user's opinion raging from 1 (bad) to 5 (excellent) [GLGP13]. Therefore, when RRM approaches are developed, it is important to perform the allocations considering QoE.

When a cell becomes overloaded, the QoS is degraded, because incoming users or services suffer high blocking rate and call dropping probabilities due to the lack of available resources. Therefore, load balancing is required in order to transfer traffic from heavy loaded cells to the lightly loaded neighboring cells [GCS17]. It involves exchanging information periodically between evolved Node B (eNB)s and users. The load of the cells is compared, and, if needed, a Hand-over (HO) procedure is initiated. In frequency domain, load is typically defined as the percentage of used Resource

Blocks (RBs) over the total available in the system, as it is the smallest combination of data that can be transmitted or received to/from a terminal.

In addition, SON techniques have been developed to automate network management so that Load Balancing (LB) is achieved. The aim of load balancing techniques is to alleviate congestion problems caused by unevenly distributed traffic. This is achieved by sharing traffic among neighbor cells through the modification of HO margins. Different network/cell parameters can be modified in order to implement LB, with HO margins being one of most used options [RIBMB11]. This balance is expected to decrease the overall blocking ratio, thus increasing the total carried traffic in the network [ARG17].

When performing RRM, the user location knowledge is important from many points of view, not least in understanding and optimizing network performance. Mobile operators plan their network coverage dividing the area into cells and sectors. Cell edge users consume on average much more radio resources than users near the cell center, and even though the cell edge problem is well known and different techniques are addressed to mitigate this, very little work has been found regarding the actual user distribution within the cell, while some work has been found discussing different effects of it. To find the distance and the direction from the base station to the phone, the location of the serving cell is retrieved from the cellular network. The Cell ID from the smartphone and the cell list from the network are then correlated in order to calculate the distance and direction (see Fig. 6.3). Data collected from a live network clearly indicates that the angular direction follows a symmetrical distribution with the mean and median close to the center of the cell sector. The von Mises angular distribution is shown to be a good model for the angular distribution of the users [LG17].

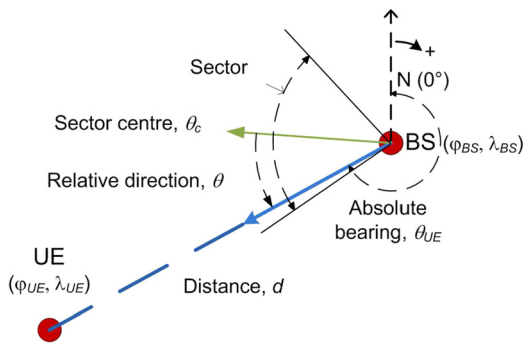


FIGURE 6.3

Definition of directional data within a sector [LG17]. The symbols φ and λ refer to the latitude and longitude.

In order to take the maximum profit of radio spectrum, higher frequency bands are exploited for low coverage areas, while lower frequency bands are used to provide great coverage areas but with a smaller capacity. This difference between coverage

and capacity of different frequency bands brings some challenges in network planning, as it needs to cope with the resulting interference [GCS17]. Hence, the use of frequency bands together with the power control plays an important role in RRM and scheduling. In the sections below, we present different approaches that target to utilize the radio resources effectively, scheduling the users according to a resource utilization plan.

6.4.1 Resource allocation in wireless mesh networks

This section focuses on the frequency and time resource allocation problem of the Orthogonal Frequency Division Multiple Access (OFDMA) Wireless Mesh Networks (WMN) system which differs from cellular network allocation due to the spacial architecture and because the OFDMA channel can be split into sub-bands of subcarriers. Using graph theory models and algorithms an optimization scheme was proposed to achieve the goal of minimum resource utilization over a network and allow frequency reuse [PR16], [RWM18]. The design of WMN considered is based on the following assumptions [RWM18], [RWP18]:

1. The resource allocation is initially performed, and routing will be done on a network which has resource allocations.
2. The resource allocation is performed by a centralized algorithm which has knowledge of all the connections in the network.
3. Connection between two nodes is possible if the signal of noise and interference ratio, SNIR is above a minimal level. A signal below this level does not contribute to the interference.

The following shared resources are considered:

- Time slots: The system operates with time slots division, having time synchronization among nodes.
- Frequency: The frequency division can be among different OFDMA bands, or among different sub-bands in an OFDMA band.
- Separation Rules

The non-collision rules assumed are:

- A node may not transmit and receive in the same time slot in the same frequency band
- All transmissions from one node to other nodes in the same time slot should have different frequency bands, or alternatively in a specific band should have different frequency sub-bands.
- All receptions at a node from other nodes in the same time slot should have different frequency bands, or alternatively in a specific band should have different frequency sub-bands. This includes intentional transmissions to this node as well as interference, such as transmissions from another nodes in the area of reception not intended for this node.

Each combination of time slots frequency band, and frequency sub-band is called a Resource Element (RE). The algorithms have the ability to obtain an assessment of the resources required globally in the system. An example of allocations produced with the simulation is given in Fig. 6.4. The links on the graph are marked with the allocated REs in both direction. We denote $|V|$ as the number of vertices, and $|E|$ as the number of edges. D_{max} is the maximum degree of the graph and $|C|$ (for "colors") is the number of resource elements allocated by the algorithm.

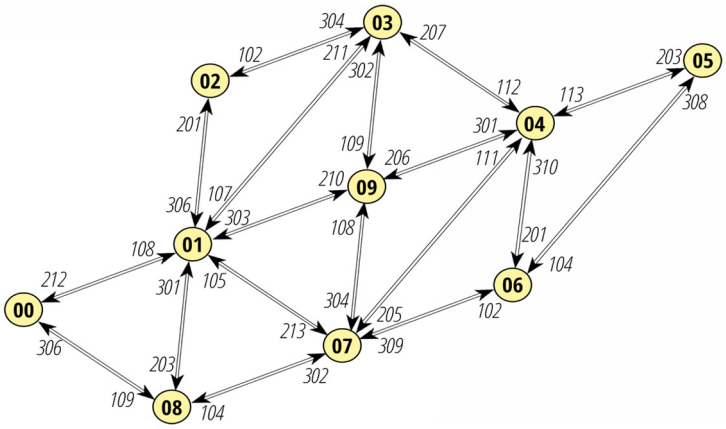


FIGURE 6.4

Example of resource elements allocation. $|V| = 8$, $|E| = 36$, $D_{max} = 6 \longrightarrow |C| = 24$.

6.4.2 RRM for D2D scenario

RRM also plays an important role in D2D communication. D2D communication in cellular systems is defined as direct communication between two mobile users without traversing the BS or core network [AWM14]. D2D in this case, typically re-uses UL resources since they are under-utilized. When allocating resources to D2D users, the interference and the user power levels play an important role. Most of the existing studies focus on a simple case of single user pair and hence are not practical. It is therefore important to develop approaches that can consider the intercell interference and enables the resource sharing between more than two users which promises great potential.

By adapting and expanding the interference structure of [ZCYJ15], an interference-graph based resource allocation algorithm for the LTE UL that respects the Single Carrier Frequency Division Multiple Access (SC-FDMA) constraint of continuous resource assignment to individual users, resource sharing between multiple users while controlling the interference is achieved [BK16]. Here, the interference situation in each cell is modeled as an interference graph G = {V, E} consisting of vertices V and edges E. Each communication pair forms a node, represented by a vertex in G and the edges between them are used to model the potential interference. Simulation

results have shown that such an approach increases the system capacity greatly. It should be noted that such capacity increase comes at the cost of a lot of information (pathloss and interference) required at the scheduler. Besides, the additional overhead to communicate the determined schedule depends on the dynamic of the scenario and may prohibit the straightforward application in an actual network.

6.4.3 RRM via frequency reuse

By reusing the existing frequency bands, either the same or different bands, the mobile users can be allocated to different bands in order to meet their QoS requirements. This is feasible as the coverage is frequency dependent as shown in Fig. 6.5. With Signal-to-Noise Ratio (SNR) thresholding and QoS requirement, an algorithm that minimizes the HOs is developed in [GCS17] calculating the required user data rate, expressing its SNR as [3GP10]

$$R_{b\,[\text{bits/s}]} = \frac{A}{B + e^{-C\rho_{[dB]}}} \tag{6.1}$$

where A, B, and C depend on the modulation type. By calculating the SNR for different frequency bands visible to the mobile users, a user is then allocated to an available BS. Users' load distribution to different bands provides high QoS with gain in throughput exceeding 8% [GCS17].

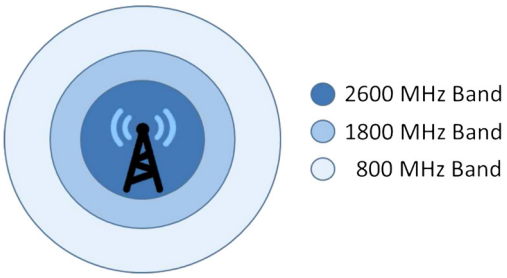


FIGURE 6.5

Coverage of different frequency bands.

Traditional load balancing techniques, which do maximize QoS statistics, may show certain limitations when optimization algorithms alternatively focus on QoE as the main figure of merit. In [ARG17], the cell load and QoE balancing techniques are investigated in an LTE network. A simulation tool is developed that offers performance statistics of every user connection, from which QoE values are estimated by using the so-called utility functions using MOS scale. These utility functions allow to quantitatively model the relationship between objective performance QoS indicators taken directly from the network and QoE, highlighting the QoS impact on the user subjective perception of the service. Moreover, HO thresholds are also modified

so that LB is better achieved, as shown in Fig. 6.6. In the tool, the QoE is measured through different equations defined for voice, video, and Web services. Results show that load balance algorithms have limitations and may even worsen the experience perceived by the users when comparing QoE performance indicators [ARG17]. Hence, the future systems shall also use QoE as the network performance.

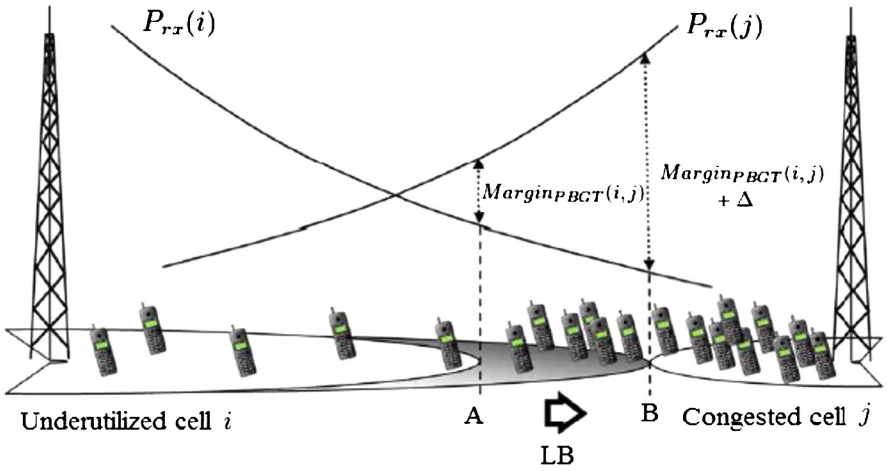


FIGURE 6.6

Handover margin modifications by Load Balancing [ARG17].

The effect of HO is further analyzed in [GLRT17], where a sensitivity analysis of throughput according to HO margins is presented. An alternative criterion for tuning HO margins is then introduced, focusing on end-user throughput. A mobile LB approach that modifies the service area of highly congested cells by decreasing HO margins, while simultaneously low congested cells increase their HO margins is introduced [GLRT17]. The assessment of this approach is carried out in a trial LTE network. Results show that the proposed indicator improves network performance in terms of end-user throughput from that obtained with classical mobile LB algorithms. The proposed algorithm changes cell service areas aiming to improve end-user throughput from that of classical mobile LB techniques. Unlike most research in this field [DTRM15], method assessment is not carried out in a simulator but in a trial LTE network [GLRT17].

RRM for DUDe scenario

RRM coordination can also happen between different BSs, where one BS is typically a macro BS while the other can be a pico BS. This is depicted in Fig. 6.7 and is termed as DL and UL Decoupling (DUDe) [SA18]. Some works claim that the gains of DUDe could reach up to 200%-300% in 5th percentile UL throughput when compared to a baseline case without Cell Range Expansion (CRE) + enhanced Inter Cell Interference Coordination (eICIC) [EBDI14].

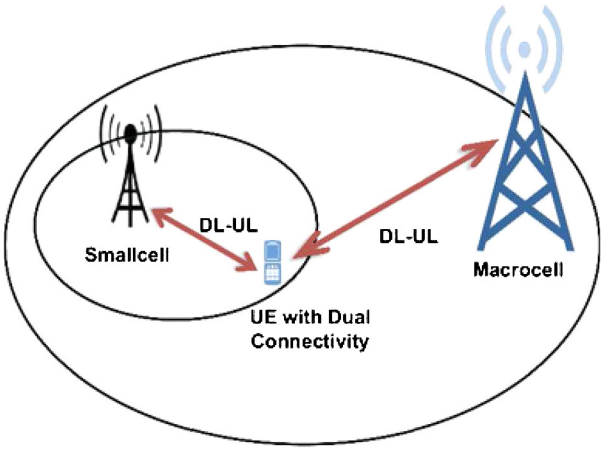


FIGURE 6.7

The diagram for the sample case of DUDe. Reprinted from [SA18].

In a DUDe scenario, a given user chooses the macro BS for its DL traffic, while for the UL the pico BS is chosen. Not only does DUDe split the UL and DL, but it uses also different cell selection criteria such as pathloss information or interference levels [WGLM+17]. When simulations for different user cases are run, it can be concluded that the CRE + eICIC slightly outperforms DUDe in improvement of UL user throughput when a dense user environment is deployed [WGLM+17].

6.4.4 PCA for higher capacity

Frequency reuse in cellular systems requires utilization of different methods for minimizing the interference [MH17]. One of these methods is to place a Power Control Algorithm (PCA) in charge of monitoring the link gains and managing the transmission power levels accordingly [Zan92]. There are two types of PCA: centralized and decentralized (distributed). Centralized PCAs have a single control center and calculate the best transmission power levels for an entire network. Decentralized PCAs have independent distributed control centers which are each in charge of a small region. Each of these control centers calculates the best transmission power levels for their region based on the live status of their local network links and sometimes their neighbors' as well. While decentralized PCAs generally provide less accurate results [Zan92], they are preferred today over centralized PCAs due to the fewer control links required and their real-time management speeds.

The decentralized PCA algorithm offered in the technical document [MH17], exploits the natural propagation of radio waves in a populated environment. It basically operates in a merge sort fashion; taking a large problem and splitting it into smaller overlapping ones, solving them with brute-force and then merging them back together. This Merging Brute Force Algorithm (MBFA) provides extremely accurate

minimum acceptable threshold. Simulations have shown that the MBFA, in certain scenarios, can provide surprisingly high accuracy even though it is an approximation of the optimum algorithm, the brute-force (a centralized PCA). Accuracy of 99.34% is achievable when the algorithm is configured correctly for each particular scenario.

6.4.5 Resource allocation and sharing for HetNets

RRM across different networks has been studied extensively as the future 5G-grade NB-IoT systems will be utilized for the sensory data collection. In this heterogeneous network, it is expected that 5G networks will carry heavy traffic like video streaming, while the NB-IoT systems will carry low-load data traffic. In [SG18], an analytic performance evaluation framework that captures both service and network centric metrics of interest as well as the peculiarities of resource sharing between LTE and NB-IoT transmissions is developed. Using this tool, three feasible strategies for RA and sharing between LTE and NB-IoT sub-systems, namely: static, dynamic and dynamic with reservations are developed. The RA for each of these cases is represented in Fig. 6.8. A unique property of the considered model is that the basic channels are assumed to be scheduled sequentially by taking into account the features of in-band NB-IoT technology [WLA+17].

The analysis shows that static resource sharing between LTE and NB-IoT is highly sensitive to the offered traffic load. On the other hand, dynamic resource sharing between LTE and NB-IoT drastically improves the overall system performance. However, it does not automatically provide any guaranteed performance to the over-the-top services and may as well, unpredictably degrade performance, especially when the offered traffic load across sub-systems is non-uniform.

The dynamic RA strategy with resource reservation allows to satisfy the feasible reliability requirements needed by big data analytics. The proposed dynamic strategy with reservation is therefore strongly recommended for the forthcoming IoT deployments over 5G mobile networks.

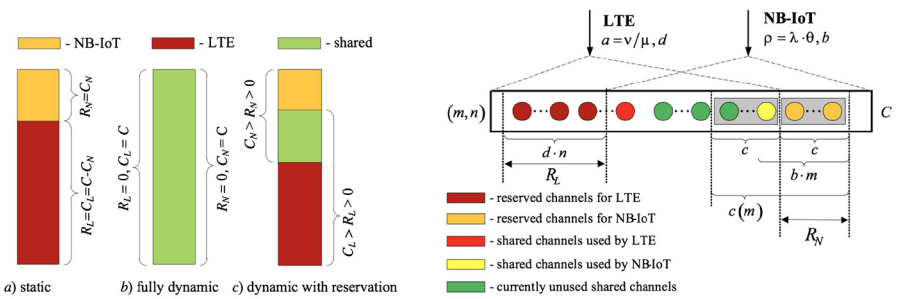


FIGURE 6.8

RA strategies and the process of sequential resource allocation in an LTE cell. Here, C represents the channels and R represent the minimum number of channels, with subscript L indicating LTE, and subscript N indicating NB-IoT.

6.4.6 A RRM tool

The previous sections presented different approaches and cases for RRM testing, where the researchers develop their own simulation platform, because the existing simulators are unlikely to find a one-fits-all solution. The Vienna LTE Simulator [MIŠ+11] is implemented in MATLAB® and it is rather focused on the lower layers, so it cannot appropriately model different service types. Besides, the simulation is rather time-consuming [DA17]. The NS3 [NS3] tool provides network level simulations but it also suffers from long simulation time [DA17].

To provide a better answer for such question, in Generic Wireless Network System Modeler (GWNSyM) a flexible platform that allows easy-configuration and easy-analysis of rather large and complex system deployments is developed. The tool has been designed with the main goal of being easily extended with either new functionalities or strategies. Unlike other tools that follow event driven simulation, the simulation with GWNSyM is based on snapshots, where the previous snapshot is used to feed the following one, which is required to analyze service evolution and dynamics (e.g. varying load) [DA17].

6.5 Heterogeneous networks and ultra dense networks

The ever-increasing demand for bandwidth hungry broadband services, and enhanced QoE, increased spectrum efficiency, and reduced energy consumption, have resulted in several challenges in designing next generation 5G wireless networks [HH15]. The use of network densification through the deployment of low power small cells, whether by a mobile network operator or an end user, is recognized as one of the key strategies towards achieving the 5G vision and targets. By densely deploying additional small cell nodes within the local area range and bringing the network closer to end users, the performance and capacity are significantly improved [SZHA15,FVV+14]. The energy levels decrease, while retaining seamless connectivity and mobility, resulting in improved QoE and user satisfaction.

6.5.1 Scenarios and capacity evaluation for small cell heterogeneous networks

Scenarios and Architectures for RRM and Optimization of Heterogeneous Networks have been discussed in [SVHR17], where the state of the art information related to scenarios and architectures from Energy efficient High-speed Cost Effective Cooperative Backhaul for LTE/LTE-A Small-cells (E-COOP) technologies is discussed. One of the proposed architectures exploits infrastructure based on small cell deployment using RRU technology which is connected to the core network using backhaul technology based on fiber optic links. Fig. 6.9 presents the general architecture for Cloud RAN.

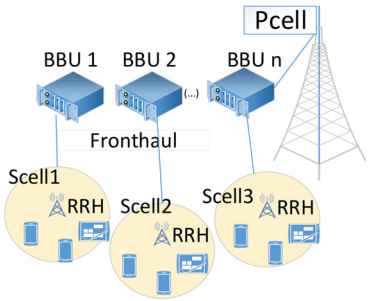


FIGURE 6.9

Cloud RAN architecture.

In Fig. 6.9, the Macro base station (PCell), operating in a low frequency band, e.g., 800 MHz, provides full coverage and Small Cells (SCells), operating in a upper frequency (e.g., 2.6, 3.5, or 5 GHz), provides local high data rate.

6.5.2 System level evaluation of dynamic base station clustering for coordinated multi-point

The authors from [Sch17a], [Sch17b] introduce promising techniques to achieve Coordinated Multi-Point (CoMP), which are partly standardized for LTE networks. The authors focus on Joint Transmission (JT) in the DL direction of the C-RAN, where the same data is transmitted from multiple eNBs, organized into CoMP clusters, so that the signals interfere constructively at the receivers. Under these clustering groups, cooperation is only allowed between eNBs belonging to the same cluster.

The realistic scenario under study accounts for user mobility patterns, Web-traffic modeling, and dynamic reconfiguration of clusters. Although LTE has been considered by the authors, as long as the concept of OFDM and the segmentation of time and frequency resources into resource blocks is utilized, the considered approaches are also applicable to the upcoming 5G.

Authors adopted the dynamic clustering algorithm (Algorithm 1) from [BBB14]. The simulation results of a realistic network including a vehicular mobility model indicate that the algorithm significantly improves the total performance in terms of sum data rate but also the individual data rates perceived by the UEs, even in the presence of mobility or highly dense scenarios.

6.5.3 Comparison of the system capacity between the UHF/SHF bands and millimeter wavebands

The viability of spectrum sharing in UHF, SHF, and mmWave in outdoor environments have been discussed in [BSV18]. In the mmWave the linear cellular topology is considered, while in the UHF/SHF bands cells with hexagonal shape are assumed. Performance evaluation includes the study of the behavior of PHY and supported

throughput for 2.6, 3.5, 28, 38, 60, and 73 GHz. While the two-slope model is considered for the 2.6 and 3.5 GHz frequency bands, the modified Friis propagation model, with shadow fading and different values for the standard deviation, is considered in the millimeter wavebands.

The variation of the supported throughput with the coverage and reuse distances is analyzed in [TV18], for different values of the Channel Quality Indicator (CQI) and reference Carrier to Noise-plus-Interference Ratio (CNIR) requirements recommended by 3GPP, and given International Telecommunication Union-Radiocommunication Sector (ITU-R) propagation models. The PHY throughput is computed through the implicit function analytical formulation that was already applied in [SVP17], [TV18]. One example of the variation of the PHY throughput that corresponds to the curves of CNIR for the UHF/SHF and mmWave bands is presented in Fig. 6.10.

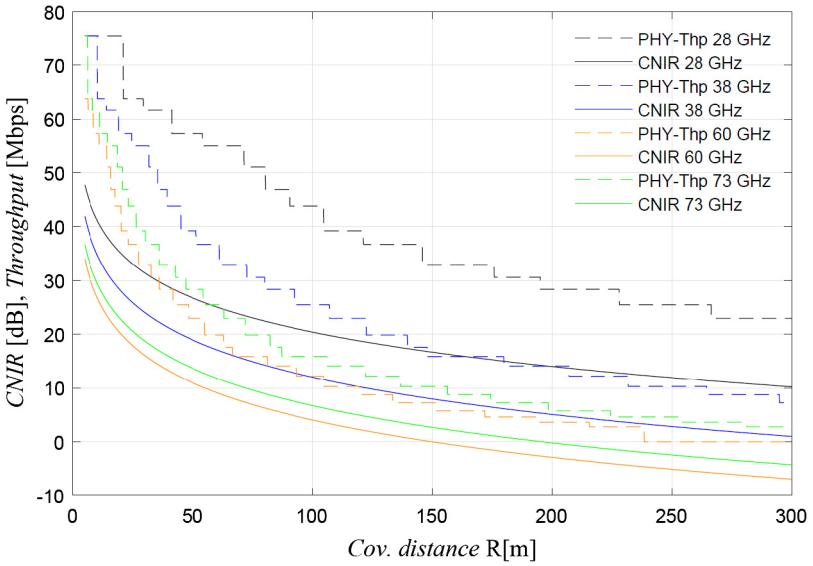


FIGURE 6.10

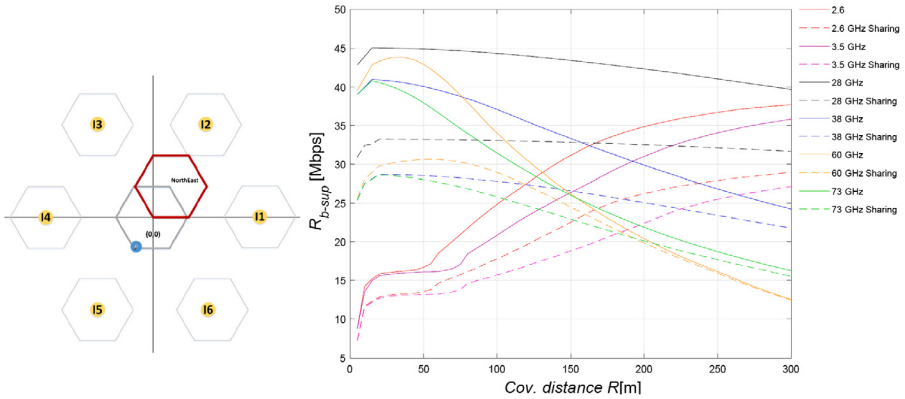
Variation of the CNIR/SINR and PHY throughput with d for 28 GHz, 38 GHz, 60 GHz, 73 GHz, for $R = 300$ m.

As there are limitations in the availability of dedicated spectrum for each Mobile Operator (MO), this work considers sharing without coordination, which means that each operator adopts the same frequency reuse strategy. Spectrum sharing access in heterogeneous networks assumes that two or more MOs have dedicated spectrum for macro cellular layer while SCells will share the access to spectrum in an opportunist manner.

In fact, scenarios where the frequency bands are shared are very relevant and an opportunity for new entrants, e.g., service providers, to own their light infrastructure.

These SCell networks will likely be deployed in a near future. The intermittent nature of traffic access will facilitate that usage of resources from different networks can temporarily access the infrastructure with the same resources. The authors from [BSV18] considered, as a case study, that one cell from MO #2 overlapped with cells from MO #1, as shown in Fig. 6.11(a), both using the shared band, as shown in Fig. 6.11(b).

In the sharing scenario from [BSV18], the supported throughput is highly reduced by the interference caused by the cell from a different MO. However, the ubiquitous coverage of the mobile communication system remains, as the PHY throughput does not reach zero inside the cell. To understand the practicability of considering spectrum sharing in UHF/SHF and mmWave, Fig. 6.11(b) puts together all the curves from the analysis of these frequency bands with and without sharing.



(a) Scenario with one UHF/SHF SCell interferer (Northeast)

(b) Supported throughput for a cell radius of 300 m

FIGURE 6.11

Comparison of the system capacity with/without sharing between the mmWave and UHF/SHF bands, with $K = 3$ and 20 MHz of bandwidth.

The supported throughput is higher for the 28 GHz frequency band compared to the remaining mmWave frequency bands (38, 60, and 73 GHz). One can also conclude that the system capacity decreases for increasing coverage distances. For shorter coverage distances up to circa 100 m, the supported throughput is higher at the millimeter wavebands, mainly due to the reduction that characterizes the application of the two-slope propagation model at the UHF/SHF bands.

It is noticeable that the supported throughput performance for the 28 GHz frequency band is the best compared to the other considered frequency bands, with values varying between the 45 Mbps and 39 Mbps for the scenario without sharing, and values varying between 30 Mbps and 33 Mbps for the scenario with shared spectrum. For all the other frequency bands, due to the behavior arising from the two-

slope model (Umi LOS) applied to 2.6 and 3.5 GHz the supported throughput at the mmWave is higher than for the UHF/SHF bands, for the shortest coverage radii.

As the coverage distance increases the upper MCSs are supported up to shorter distances inside the cells. These results allow for interpreting why at the mmWave bands the system capacity is clearly higher for the shortest coverage distances, ceasing to be better for longer distances (except for the 28 GHz band, whose supported throughput is always the highest one).

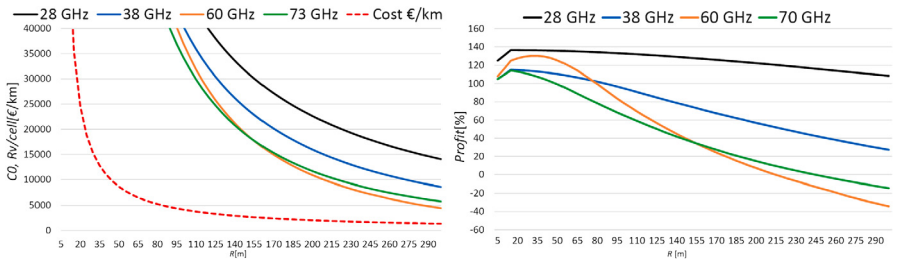
For all the other frequencies considered, we have learned that the highest system capacity and the highest MCSs are achievable for the shortest coverage distances at mmWave bands, which have a higher throughput up to 156 m, after this coverage radius, the 2.6 and 3.5 frequency bands have higher supported throughput values, reaching a maximum of 38 Mbps. In fact, due to the behavior arising from the two-slope propagation model (Umi LOS) applied to 2.6 and 3.5 GHz, for shorter R the mmWave have the highest values for the supported throughput. One concludes that the mmWave band a higher supported throughput is achievable for shorter radii ($R \leq$ 156 m) while at the UHF/SHF highest values of the supported throughput only occurs for longer radii ($R >$ 156 m). For cell radii up to 300 m, higher supported throughput always occurs for the 28 GHz frequency band.

6.5.4 Cost/revenue trade-off in small cell networks in the millimeter wavebands

Mobile cellular communications in the millimeter wavebands can support very high bit/data rates within small cells with short-range coverage. Based on the results for the equivalent supported throughput presented above, cost and revenues have been studied, [TV18]. One considers the cost/revenue model described in [VC03]. The revenues per cell, $R_v/cell$ [Euro], can be attained as a function of the throughput per Base Station, $thrBS$ [kbps], and the revenue of a channel with a data rate Rb [kbps], Rrb [Euro/min], and Tbh corresponding the equivalent duration of busy hours per day.

Revenues are considered on an annual basis. Six busy hours per day, 240 busy days per year [VCMV12], and the price of a 144 kbps "channel" per minute (corresponding to the price of approximately 1 MB) are also considered. In Fig. 6.12(a) presents the results for the overall cost per unit length per year, C_0, and the revenue per unit length per year, for only one carrier through the variation of R for $R_{max} = 300$. The revenues are noticeably higher than the costs of all frequencies. For 28 GHz band better results in terms of incomes are expressed. Fig. 6.12(b) presents the results for the Profit.

The results clearly indicate that short coverage distances should be used to optimize the system by having into consideration the maximization of the profit in percentage (mainly in 28 GHz), as the goal of operators and service providers is to enhance the profit.



(a) Network revenue/cost per unit length per year as a function of R

(b) Profit per unit length

FIGURE 6.12

Cost/revenue trade-off with $R_{max} = 300$ m.

6.5.5 Effects of hyper-dense small-cell network deployments on a realistic urban environment

In [RK16] a realistic urban scenario for a hyper-dense small-cell network deployment is presented, using full 3D building models and 3D ray-optical pathloss predictions. In addition to the macro-cell layer, different expansion stages of the small-cell network deployment have been evaluated and compared to a 3GPP-like environment with equal site density. For this purpose, the complementary cumulative distribution functions (CCDF) of RSRP and SINR in case of a macro- and small-cell deployment on the same carrier as well as a standalone deployment of the small-cell layer, have been discussed. It has been shown that in both cases, the 3GPP-like environment provides far better coverage than the considered realistic scenario. With respect to SINR, the macro- and small-cell deployment behave similarly for the two scenarios. For the standalone case, the 3GPP-like environment provides rather steep CCDF curves, which are good to optimize, but lack certain realistic aspects. Particularly the missing degree of complexity in the building information leads to too simple predictions. Furthermore, different expansion stages of the small-cell deployments in both environments are discussed with respect to the coverage and interference situations.

6.5.6 Advanced management and service provision for ultra dense networks

Given the increasing complexity and service requirements of cellular networks, their OAM tasks have become increasingly not addressable solely by human operators' analysis. New approaches are needed for a full-stack service provision that is both optimized and reliable. In this line, the SON paradigm [NGN08] aims to automate the OAM activities and the general behavior of the network in order to optimize and automate failure management.

Location-aware self-organizing networks

Until now, classic SON approaches are based solely on the analysis of network-based metrics [AGBFM15]. However, Ultra-Dense Network (UDN) environments, with the very variable nature of their demand and high base station density imply deep limitations to those approaches [FBAGM15]. To correct this, the increasing availability of context sources [FPSB18a] and, specially, indoor localization systems [AGFCB15] provide an additional source of information to support self-optimization [AGFDB16][AGFG+16] and self-healing [FGFL+16][FAGB+15][FBAG16] in those environments.

In this line, Fortes et al. [FAGB+17][AGFMG+15] proposed a location-aware detection mechanism for femtocell scenarios based on the comparison between the *expected* Received Signal Strength (RSS) values given the UE positions. Diagnosis of the root cause behind the issue is performed simultaneously including also the outcome of femtocell and router backhaul accessibility checks. The performance of the proposed algorithms was tested in a real four femtocell testbed. In this, different UE walktests were performed as shown in Fig. 6.13. These tests shown the capabilities of the proposed approach to support quick and reliable failure management in UDNs [FAGB+17][AGFMG+15].

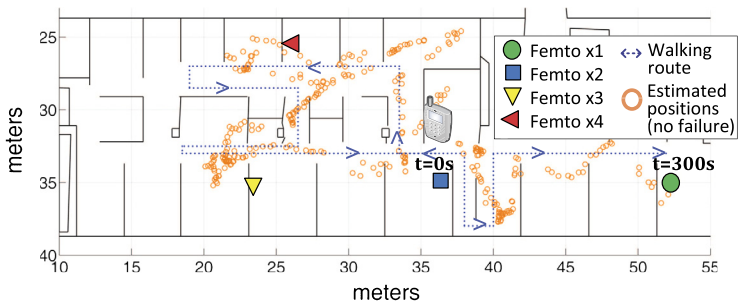


FIGURE 6.13

Testbed scenario and comparison between real and calculated positions.

Performance evaluation and packet scheduling in Home eNodeB (HeNB) deployments

Packet scheduling for UDNs is another key field of research, as studied in [PVP18b, PVP18a,PGB+11] following the simulation assumptions of [3GP09]. The study presented in [PVP18b] shows that the variation of transmitter power of the HeNBs did not influence the average SINR, only the variation of the area of the apartments influences it. Further details and study in an scenario of 24 apartments is presented in [3GP09] and another scenario with a different approach to cover the same area is presented in [PVP18a]. For the analysis performed in [PVP18b], the geometry of the scenario is presented in Fig. 6.14, where a frequency reuse of two is considered.

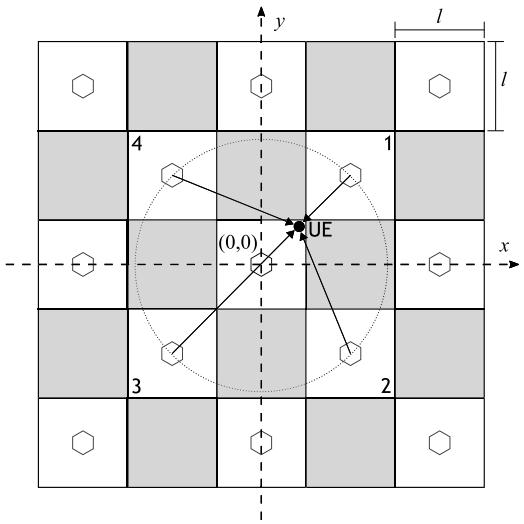
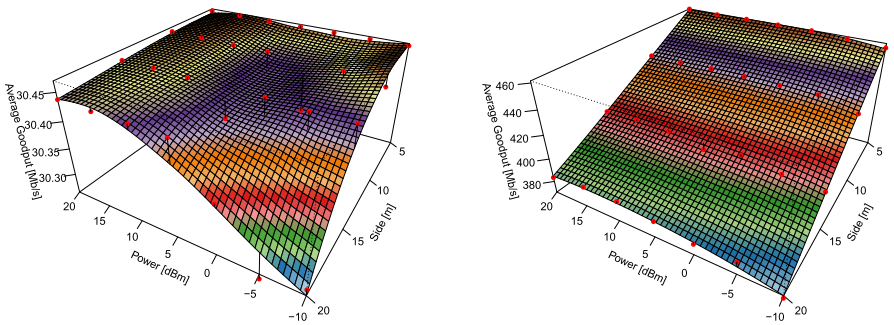


FIGURE 6.14

Simulation scenario with 25 apartments.

The type of performance evaluation defined in the standard by [3GP09] was performed by applying the LTE-Sim simulator [PGB+11,PVP16]. Results presented in Fig. 6.15 were obtained for the Proportional Fair scheduler and follow the expected theoretical behavior. For lower values of the transmitter and higher apartment side, the average throughput tends to decay more rapidly. The results for different schedulers highlighted that the best performances are not obtained with the HeNBs transmitter power to its maximum values, allowing for optimized and greener systems.



(a) Video flows (b) Best effort flows

FIGURE 6.15

Variation of the average goodput, with PF scheduler and different values of transmitter power and room side. The fitting considers a polynomial surface with 95% confidence interval.

Optimization techniques and Knapsack optimization for MLB in small cells

Another main optimization challenge in UDN is the user association problem: the fact of determining the cell to be associated with each UE. The main modeling approaches for this are based on a "utility" cost function maximization which quantifies the satisfaction that a certain metric is met. Several techniques relying on combinatorial optimization were previously investigated [LWC+16,MAAV14,MBLM13]. Knapsack Optimization (KO) [MT90] is a new combinatorial optimization technique that is proposed for Mobility Load Balancing (MLB) for dense small cell deployments [NVM18,NM18].

An example use case based on a seven small cell scenario and 120 users unevenly distributed is used to compare the proposed MLB-KO in comparison with eICIC as defined in the 3GPP release 11 standard and with different CRE values and Almost Blank Subframes (ABS) of 20%. The Blocking Ratio (BR) is used to assess the effectiveness of the optimization as reported in Table 6.1. It is concluded that the MLB-KO technique outperforms CRE/ABS by at least 1.5 times for the 12 dB case increasing to 2.3 times for the 6 dB case and achieves more than four times improvement compared to the case with no MLB scheme.

Table 6.1 Comparison of BR for the different techniques.

Technique	Average BR (%)	Improvement factor of MLB-KO
No MLB	26.06	4.22 x
CRE = 6 dB	14.78	2.39 x
CRE = 9 dB	12.31	1.99 x
CRE = 12 dB	9.55	1.54 x
MLB-KO	6.17	-

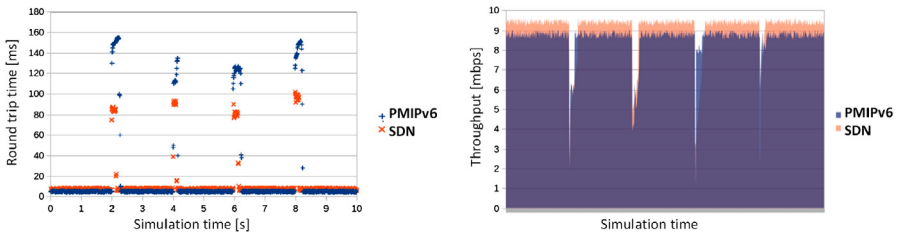
6.5.7 IP mobility and SDN

Another important area to be addressed in UDNs is the mobility at higher layers of the protocol stack as well as the challenges open by the adoption of SDN. As studied in [BBB17,BBB17] traditional IP networking cannot solve the Layer 3 (L3) mobility problem. However SDN approaches can solve it efficiently: these works proposed to enable L3 mobility translating the client address at L3 level. This is done by using SDN for dynamic translations between the real temporal Internet Protocol (IP) addresses and virtual permanent IP address, resulting in seamless connection on both server and client sides and preventing the interrupts and subsequent reconnections.

This allows continuous data flow, which is of uttermost importance for Smart Things. Such a functionality has modest equipment requirements, the changes of the existing network are at the core layer by implementing SDN functionality. The rest of the network is implemented using the traditional approach. It does not use tunneling, but original packets with the least changes at the central router. As clients packets

are not changed inside the local domain, optimal local routing and maximal link efficiency are preserved.

Compared to Proxy Mobile IPv6 (PMIPv6) via simulations in Mininet environment [Min], Fig. 6.16(a) shows the round trip time, while the maximal throughput is presented in Fig. 6.16(b). Switching time from a private IP address of one network to a private IP address of another network is reduced in the proposed solution, while the tunnel encapsulation causes visible decrease of throughput in the classical PMIPv6.



(a) Switch-over packet delay (b) Switch-over packet throughput

FIGURE 6.16

Simulation results.

6.5.8 Digital geographical data and radio propagation models enhancement for mmWave simulation

New simulators or simulation bricks could be used by equipment manufacturers or network providers in the mmWave infrastructure dimensioning, performance assessment, automatic network design and optimization. Previous works related to mesh backhaul performance assessment and design [MCL16], adjustment and validation of propagation models [CCAL16], [ACL17] were included on the connectivity prediction for mmWave urban mesh backhaul networks.

The backhauling prediction engine must fulfill three key requirements: first, precisely estimate the visibility condition; second, predict the mmWave propagation loss for both LOS and NLOS situations; and third, support the indirect antenna alignment, i.e. when the directive antennas may be oriented not towards a dominant but towards an indirect propagation path.

Four different propagation methods on a large set of backhaul link candidates in a North American city are briefly presented in Table 6.2. The presented methods rely on distinct geographical data: Open Street Map (OSM) buildings in the first method; High-Resolution (HR) geographical data in the second one; completed with Lidar point cloud for the third and fourth methods. A direct-path propagation technique Direct-Path (DP) is employed in the three first methods. Multi-Path (MP) are enabled in the latter one.

The three methods consider geographical data models on the same one square kilometer area of the city, where most building heights are above 10 meters, and rows of trees are present along most of the streets.

Table 6.2 Tested propagation methods and LOS Percentage.

Tested Methods	DP/OSM	DP/HR	DP/Lidar	MP/Lidar
Geo data	Open building database from OSM	3D HR vectors	3D building vectors and Lidar point cloud	
Propagation model	DP only. Diffraction above buildings	DP only Diffraction above buildings. Interaction with vegetation vectors	DP only Diffraction above buildings. Interaction with vegetation point cloud	MP Interaction with vegetation point cloud
LOS percentage	59.7%	33.3%	11.7%	

The first model was extracted from an open building database, i.e. OSM. The HR database was produced by Siradel from stereoscopic aerial pictures. The last tested geographical database has been specifically elaborated for improving the accuracy of the propagation model inmmWave scenarios. Lidar point cloud was collected in all streets of the considered area, with a density and a measurement protocol that guarantee a precise 3D representation of most trees located in the street.

The propagation models under test were all obtained from different configurations of the VolcanoUrban tool [CL09]. The tool was adjusted and validated in several mmWave frequency bands, including 60 GHz. More particularly, the way to manage the vegetation obstacles has been improved [CCAL16] so both the transmission through the vegetation and the diffraction on top of the vegetation are considered.

A comparison between prediction models, Table 6.2, shows that large variations from one model to the other are mostly caused by the representation of trees. A large majority of links are found as LOS in absence of any the vegetation, while only 33.3% links remains unobstructed if main vegetation blocks and main rows of trees are added in the geographical data. Finally, the LOS percentage decreases as low as 11.7% when vegetation blockage is estimated from the rich tree representation offered by the Lidar point cloud.

Summary and conclusions

This section started by motivating the use of SCells in the context of HetNets and UDNs followed by the presentation of scenarios and general architecture for C-RAN in this framework. Combining dynamic clustering and scheduling for coordinated multi-point transmission enables to considerably enhance the total performance in terms of sum data rate but also individual data rates perceived by the UEs, not only in LTE but also in 5G networks. The comparison of system capacity between the UFH/SHF and mmWave was followed by a comprehensive study of the cost/revenue trade-off of SCell networks in the mmWave. A realistic urban scenario for a hyper-dense small-cell network deployment was considered whilst considering full

3D building models and 3D ray-optical pathloss predictions. In addition to the macro-cell layer, different expansion stages of the small-cell network deployment have been evaluated and compared to a 3GPP-like environment with equal site density, which present much better performance than the realistic one. Femtocell scenarios and advanced management and service provision for UDNs have been discussed and research on location-aware SONs has been addressed. Performance evaluation of HetNets in UDNs compared different packet schedulers, and considered Knapsack optimization and simulation methodologies. The mmWave have also been addressed by considering an approach for radio propagation modeling enhancement accounting for digital geographical data, e.g., vegetation obstacles.

6.6 C-RAN

6.6.1 Resource management in C-RAN

C-RAN as a key feature of 5G, offers a lower cost solution to the exploding data rate demands. C-RAN splits the Base-Band Units (BBUs) of BSs from their radio units, i.e. the Remote Radio Heads (RRHs), and integrates the decoupled BBUs into a central location, known as a BBU-pool. By leveraging the advantage of load fluctuation, it improves the utilization of network resources by sharing the resources of under-loaded BBUs with the over-loaded ones. An efficient resource management strategy is needed however, to distribute BBU-pool scarce resources among BBUs, especially in case of resource shortage, when not all of the BBUs demand can be served at the same time.

Brahman et al. in [BCF18a] and [BCF19a] propose an on-demand resource allocation scheme to optimize the utilization of computational resources in a BBU-pool. Resource allocation is formulated as a centralized optimization problem, based on the concept of bargaining in cooperative game theory. Having the real-time BBU demand and the weight of the active services in the BBU, a bargaining power is calculated and assigned to a BBU that prioritizes it in case of a resource shortage. Meanwhile, a minimum amount of resources is guaranteed to be allocated to prevent the BBU from crashes. Therefore, the computational resource distribution among N_B BBUs in a pool is optimized by solving the following problem:

$$\max_{\mathbf{C}^{\mathbf{Al}}_{b,t_k[N_B\times 1]}} \quad \prod_{b=1}^{N_B} \left(C^{Al}_{b,t_k\,[\mathrm{GOPS}]} - C^{R}_{C\,b,t_k\,[\mathrm{GOPS}]} \right)^{B_{b,t_k}} \tag{6.2a}$$

$$\text{subject to} \quad C^{Al}_{b,t_k\,[\mathrm{GOPS}]} \leq C^{R}_{b,t_k\,[\mathrm{GOPS}]}, \quad b = 1, 2, ..., N_B \tag{6.2b}$$

$$\sum_{b=1}^{N_B} C^{Al}_{b,t_k\,[\mathrm{GOPS}]} \leq C^{Av}_{BP\,t_k\,[\mathrm{GOPS}]}, \tag{6.2c}$$

where:

- C^{Al}_{b,t_k}: allocated computational capacity to BBU b at time instant t_k,
- C^{R}_{b,t_k}: required computational capacity of BBU b at time instant t_k,
- B_{b,t_k}: bargaining power of BBU b at time instant t_k,
- $C^{Av}_{BP\ t_k}$: available computational capacity in the BBU-pool.

The constraints defined in (6.2b) and (6.2c), ensure respectively that a BBU cannot be allocated with more resources than it requires and that the total allocated resources may not exceed the available ones in the BBU-pool. Simulation results for a scenario with seven BBUs in a BBU-pool and a heterogeneous service environment show a minimum 83% improvement of allocation efficiency compared to a fixed resource allocation policy based on peak requirements.

Later, in [BCF18b] and [BCF19b] the authors show that the above model for computational resource management is also applicable to a time varying network where the BBU load and computational resource demand are dynamic. Based on changes in the instantaneous demand of the BBU, the model is re-calculated and the BBU-pool computational resources are distributed among BBUs dynamically at each time instant. Simulation results for a 5-minute interval in the same scenario show that the assigned resources change in accordance with BBUs demand. Moreover, the results confirm that BBUs demand is fulfilled more than 98% without any processing delay by using only 43.5% of the BBU-pool capacity.

Gomez et al. in [GBLO19b] proposed a similar approach for computational resource management inside a BBU-pool, the cost function given in (6.2a) being replaced by the following linear function:

$$\max_{\mathbf{C^{Al}_{b,t_k}}[N_B\times 1]} \quad \sum_{i=1}^{N_B} B_{b,t_k} C^{Al}_{b,t_k}\ \text{[GOPS]} \tag{6.3}$$

Later, the authors in [GBLO19c] propose the use of three different Maximum Likelihood (ML) techniques, i.e. Medium Gaussian Support Vector Machine (SVM), Time Delay Neural Network (TDNN), and Deep Learning using Long Short-Term Memory (LSTM), to provision computational resources more closely adapted to the actual needs of the BBU-pools.

Fig. 6.17 shows the error distribution of three proposed techniques for a BBU-pool resource provisioning. As the figure illustrates, SVM offers the best accuracy, with lower Root Mean Square Error (RMSE), compared to the other methods. In order to decrease the underestimation probability, negative errors in Fig. 6.17, the authors propose filtering data prior to the LSTM training process. The experimental results confirm that the underestimation probability reduces to zero however, the amount of the idle resources increases.

6.6.2 C-RAN deployment

The deployment of a reliable C-RAN capable of meeting network operators interests, is another challenging issue, since the trade-offs among conflicting objectives, e.g. delay minimization and load balancing, should be addressed.

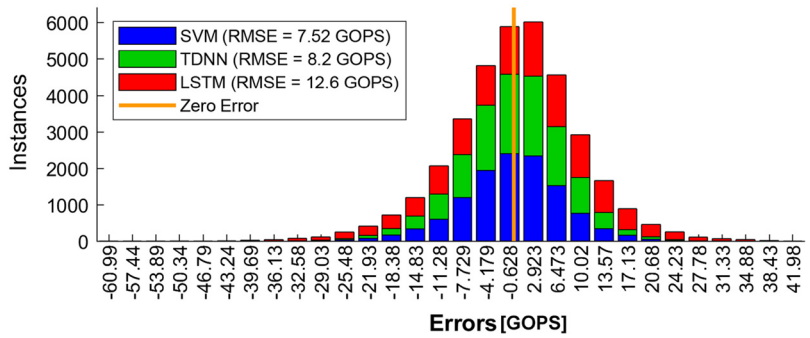


FIGURE 6.17

Histogram of the errors in GOPS for the three ML strategies.

Gomez et al. in [GBLO19a] take four different algorithms into account in the process of designing the C-RAN, each with a distinct criterion, i.e. delay minimization, BBU-pool load balancing based on the traffic, BBU-pool load balancing based on the number of the aggregated RRHs, and maximizing the multiplexing gain. Results on C-RAN deployment in a sector of Vienna city confirm that taking just the minimum delay algorithm into account minimizes round trip time, but causes lower performance and unbalanced load. On the other hand, the load balancing algorithm presents even traffic or even number of RRHs per BBU-pool. However, the round trip time is increased in both of them.

An appropriate C-RAN deployment improves both the load balancing of the network and the delay constraint satisfaction. However, the proliferation of smart IoT devices causes excessive load on the back-haul, in between massive devices and BBU-pool servers. Besides, C-RAN is an impractical solution for many delay-sensitive applications, because of the long distance between the device and the cloud center. An alternative approach is to offload some of the computing tasks from cloud servers to the network edge in the form of a new network topology known as fog computing.

Moratta et al. in [MAR+18] propose a model to determine the BBU, either in the fog node or in the cloud, that is eligible to process the workload of an RRH. The proposed model considers the round-trip delay as a composition of the propagation delay and of the processing delay. Since the propagation delay in a fiber optic link is a function of the distance between RRHs and its serving BBU, by acquiring the processing delay, the proposed model maximizes the distance of the RRH from its serving BBU so that the delay budget is not violated. Results on transferring a 6144 bits codeblock in Fig. 6.18 show that in order to be able to meet the delay constraints in case of low SNRs, either significant processing resources should be allocated to speed up signal processing, or the BBU must be located near the RRH.

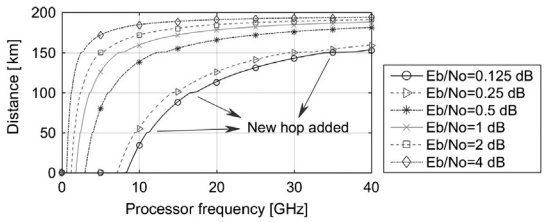
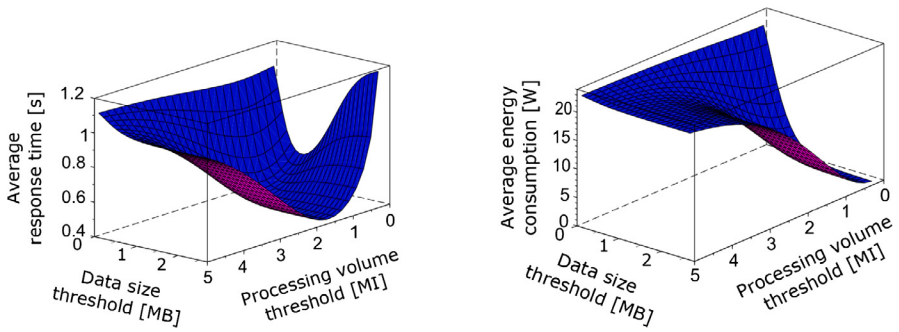


FIGURE 6.18

Maximum distance as a function of the BBU's available processing capabilities and SNR.

Besides the communication processing, the edge computing resources can also be utilized for application processing, the applications latency being reduced by offloading the computing on the edge devices, with more processing power, rather than the mobile devices with limited processing capability. Meanwhile, the trade-off between the energy efficiency and the response time, needs significant attention.

Sopin et al. in [SSC19] propose a model to offload processing tasks of a mobile device that are "heavy" in terms of the processing volume, that require more processing resources, and "light" in terms of the data size, that can be transferred fast, to a fog node. The threshold on the processing volume and the threshold on the data size filter "heavy" and "light" tasks. Fig. 6.19 depicts the impact of the offloading thresholds on the average response time in case of 20 mobile devices with random generated tasks. When the offloading probability decreases, smaller data size threshold and larger processing volume threshold, the load on mobile device becomes larger and causes an increase in average of both response time and energy consumption for locally processed tasks.



(a) Average response time (b) Average energy consumption

FIGURE 6.19

The impact of the offloading thresholds on the average response time and energy consumption.

6.7 SDN and NFV

Network Slicing, service customization and RAN Cloudification are important enablers of 5G networks, all of which are based on the capabilities of NFV and SDN [RCC18c]. Decoupling control plane from data plane in SDN, can considerably improve programmability and customization of virtual wireless networks. Meanwhile, virtualization enhances scalability and flexibility, leading to the improvement of resource utilization in SDN.

6.7.1 Virtual radio resource management model

This sub-section presents a high-level, service-oriented model for RAN slicing and RRM. The key objective is to establish a level of performance isolation between the VNOs, which are acting as network tenants, in order to ensure that their contracted Service Level Agreements (SLAs) will not be affected by the variation of different network parameters [SGSC19]. At the same time, another main goal is to share the aggregated capacity among various service slices on-demand, and in a fair manner [RCC17a]. To achieve these goals, a centralized virtualization platform called Virtual-RRM (VRRM) is modeled as an individual management entity based on the hierarchical network architecture which was presented in [RCC17b]. VRRM does not own the infrastructure and is in charge of satisfying the demands of different service slices based on the aggregated capacity provided by multiple Radio Access Technologies, owned by Infrastructure Provider.

The VRRM is analytically modeled as a convex optimization problem, according to the criterion of weighted proportional fairness. f_{VRRM}, as a measure of efficiency, maps a portion of network bandwidth that users utilize into a real number, quantifying the expected satisfaction of users given the allocated resources. This way, VRRM makes a bridge between the functionalities of MAC and higher layers, by optimizing the allocation of radio resources for different applications. There are also some constraints associated with the model, in order to apply the customized range of serving rates considering the individual VNO's policies and the contracted SLAs [RCC18d].

$$\underset{\boldsymbol{w}^{usr}}{\text{Max}}\, f_{VRRM}\left(\boldsymbol{w}^{usr}\right) = \underset{\boldsymbol{w}^{usr}}{\text{Max}} \sum_{k=1}^{N^{srv}} \lambda_k \log\left(\sum_{i=1}^{N_k^{usr}} w_{k,i}^{usr} \frac{R_k^{srv_{max}}{}_{[\text{Mbps}]}}{R_{[\text{Mbps}]}^{VRRM}}\right) \tag{6.4}$$

where:

- N_k^{usr}: number of users performing service k,
- N^{srv}: number of provided services,
- $R_k^{srv_{max}}$: maximum assignable data rate for service k,
- R^{VRRM}: total available capacity to VRRM,
- $w_{k,i}^{usr}$: assigned weight to user i, performing service k, ranging in [0, 1],
- λ_k: weight of service k, prioritizing data rate assignment.

6.7.2 Analysis of VRRM results

The performance of the VRRM model is presented and analyzed in this sub-section, based on the scenario which was proposed in [RCC18a], [RCC18b] for three types of SLAs: Guaranteed Bitrate (GB), Best effort with minimum Guaranteed (BG), and Best Effort (BE).

Fig. 6.20 shows the effect of the number of offered users on the total capacity share of the three VNOs. From this perspective, two parameters play decisive roles in decision making: the acceptable ranges of data rate variation for each VNO, which are indicated in the figure as one type of constraint, representing the internal policy of each VNO for service management. Another factor is the contracted capacity for each SLA type, which is fixed and does not change by the variation of traffic. The colored areas in Fig. 6.20 represent the case when all offered users in each VNO are being served. However, it is notable that when the number of offered users is very low, all users in VNO GB are served with the highest acceptable data rates, which are in fact lower than the values of SLA contract. In this special situation, the allocated capacity to VNO GB according to the VRRM algorithm, is set to the maximum acceptable data rate which can be offered to GB users.

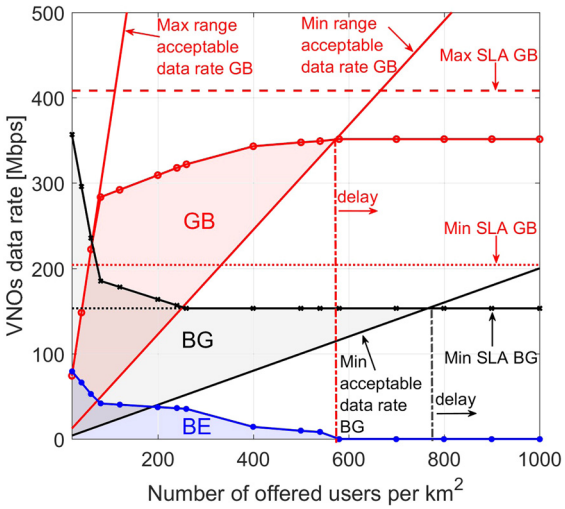


FIGURE 6.20

Capacity sharing of VRRM among the VNOs [RCC18b].

After the intersection points, when the minimum acceptable threshold of data rate surpasses the allocated capacity of each VNO, VRRM starts to delay the users from the suffered VNO based on the service priorities. As an example, when the delay in VNO GB starts, since this VNO has the highest priority among the three, all the users of VNO BE has already been delayed (as there is no minimum SLA guarantees for this VNO), and the allocated capacity of VNO BG has reached down to the minimum contracted SLA threshold.

6.8 UAVs and flying platforms

UAVs have gained particular interest in the literature thanks to their unique characteristic of moving almost freely in the sky. Equipping a UAV with radio devices can improve the path loss and link budget conditions. Since a UAV can travel when and where it is most needed, trajectories can be planned based on a specific task such as for example addressing user demand in a certain area or supporting the existing terrestrial infrastructure. This section discusses various opportunities of UAV-aided networks.

6.8.1 UAV trajectory design and radio resource management

When it comes to effectively exploit the ability of flying freely in the sky, UAVs are used for a cost-efficient deployment or real-time computed trajectories. Trajectory design is preliminary addressed by [MV17], considering one UAV as an UABS. The UABS has the purpose of supporting an underlying mobile radio network to provide service to users while accounting for the dynamic behavior of the network and user mobility. The trajectory is planned by clustering the users which remain unserved from the terrestrial network. For each cluster, a centroid point is defined such that the UABS can move from centroid to centroid by selecting the nearest to its current location. Even though, the trajectory design is straightforward, a significant improvement in the performance of the overall network can be noticed. A more advanced dynamic trajectory design is therefore introduced by [DMM+17]. Its novelty stands on the fact that it is event-driven, meaning that once a UABS arrives in a certain cluster, the algorithm is able to define its next target cluster on-the-fly. This allows the network to respond to the very challenging and quickly varying user traffic in future wireless networks. The trajectory design leverages again on the clusterization of [MV17] but chooses the next cluster centroid depending on a cost function that weights distance, user density, and direction fairness. Furthermore, an interference avoidance approach from the UABS to the terrestrial network is proposed, to allow the UABS to reuse the same frequency band in an attempt to efficiently reduce implementation costs. If efficiently implemented, RRM can both boost system performance and further decrease operational costs. [VM18] discusses the potential advantages of the integration between the terrestrial and aerial components of an UAV-aided mobile radio network, from the view point of RRM. RRM is a relevant aspect that should be addressed along with the real-time identification of the optimal route of the UAVs in a dynamic context. In fact, the UAV position, being a degree of freedom, can be seen as an additional dimension in the resource pool administered by L2 and L3 radio resource assignment algorithms. Simulation results show that a single UAV can bring an improvement in terms of network throughput of up to 1 percent. Using multiple UAVs, the performance increase can be much larger. The RRM becomes joint when the radio resources are scheduled in a coordinated manner between the terrestrial network and the UAVs, achieving a higher user throughput and an efficient network operation. The joint aerial-terrestrial resource management can be further improved by considering the decisions made in the RRM in the trajectory design as shown in [MV18]. Five

factors driving the UAV to the cluster through the cost function have been identified: energy, distance, throughput estimation, density of users, and Radio Resource Units (RRUs) availability.

UABS trajectory planning can also be formulated as an optimization problem. The goal of [MBCV18] is to maximize the number of served ground users, weighted according to their priorities, while considering constraints on the battery duration, UABS speed, and data rate. Users requesting for high throughput video services in urban environments are considered as served only when the entire video is downloaded. The problem is formulated as an ILP model, that computes the optimal path for the UABS, maximizing the number of served users, while accounting for the maximum number of radio resources available. A heuristic solution method is also presented to reduce the problem complexity and time of computation.

An important aspect for UABS network design is the impact of the Air-to-Ground (ATG) channel on which the UABS operate. [MBV18] determines how the speed and height of the UABS is influenced by the radio channel parameters. The effect of different antenna systems (with different antenna gains) is tested for UAV trajectories through simulations. The path planning accounts for user distance, density, spatial fairness, and additionally the UAV energy consumption. Moreover, the UABS is assumed to know a priori extracted Radio Environmental Map (REM) data, to better estimate the capacity it can provide to ground terminals. The results bring to the general conclusions that i) different antenna gains are more effective at different UABS heights: an increase of the UAV altitude requires a decrease in the radiation angle, ii) speed and radiation angle have shown no impact nor dependencies one to the other regarding throughput outcome. Furthermore, Ray Launching simulations for the city map of Bologna (Italy) are computed to create a realistic REM and compare these output data with statistical models [SM19]. For heights between 100 and 200 m, results show a small difference on the throughput gain in percentage for an antenna aperture angle of 120 degrees, but the gap is increasing as the aperture angle decreases.

Vehicular scenarios for 5G use cases are quite challenging for a cellular network not specifically planned for it. Moreover, when the penetration of cars equipped with wireless communication devices is far from 100%, the use of UAVs, carrying mobile base stations, becomes an interesting option. It is possible to integrate aerial and terrestrial network components operating on a single carrier frequency by using an efficient joint UAV trajectory design and radio resource assignment that accounts for vehicle mobility, vehicle density, and terrestrial network operation [MBBV19]. Network performance might increase thanks to the UABS of up to 10% in the proposed scenario.

6.8.2 UAV-aided network planning and performance

The paradigm of the previous four generations of cellular technology has led to high data rate demands from mobile users envisioned in the new 5th generation technology [SJM18]. To support this, ultra-dense networks have been proposed as

a promising 5G feature. However, ultra-dense networks face many challenges such as severe interference resulting in a limited capacity due to the dense deployment of small cells, site location and acquisition for the deployment of base stations, back hauling issues, energy consumption, etc. By showing that the coverage range of a UABS ranges from 50 to 200 m radius as well as the cell coverage variation with varying UABS fly height which can be as high as 45%, Sharma et al. conclude that the use of UABSs is a promising solution to overcome some of the aforementioned issues.

One of the limiting factors of UAV-aided networks is the inter-cell interference, especially at high altitudes. The study of [CVP18], which uses a realistic 3D simulator mode, conforms that UAVs at high altitudes suffer from significant interference, resulting in a worse coverage compared to ground users. By replacing the aerial base stations by mmWave cells, the ground coverage decreases to only 90%, but the UAVs just above rooftop level have a coverage probability of 100%. Unfortunately, UAVs at higher altitude still suffer from excessive interference. To reduce this interference, beamforming is a promising technology.

An interesting scenario for the applicability of UABSs is an emergency disaster scenario in which the existing terrestrial wireless infrastructure is saturated due to for example a natural disaster or a terror attack. By using UAVs, LTE femtocell base station can be brought to the disaster area to temporarily provide connectivity to the end user. Deruyck et al. have developed a deployment tool which allows not only to determine the required amount of UAVs, but also the most optimal locations of these UAVs to maximize the user coverage [DWMJ16]. The preliminary results of the study show that the envisioned scenario is a very promising application domain for the use of UABSs. Nevertheless, for the considered UAV specifications, a large amount of UABSs - approximately 1100 UABSs - is required to cover a city center of 6.85 km^2 like in Ghent, Belgium. The amount of required UABSs is highly influenced by the UAV specifications, the intervention duration, and the fly height or to limit the user coverage requirement. Since the latter is the less preferred choice, [DWP+17] investigate whether it is interesting to reduce the number of required UABSs by installing femtocell base stations in the vehicles of the emergency services and public transportation. Deruyck et al. show that this is indeed feasible but the effect is rather limited since only 5% of the users in the affected area can be reconnected through the vehicles. The main reason is that the vehicle's location is not known beforehand and can not be chosen as freely as it is the case for the UABSs. One issue related to using UAV-aided communication in emergency scenarios is a proper design of the backhaul network, i.e. from the drone to the core network. One way of doing this is to provide a direct link between the UABSs and the core network by using currently unoccupied frequency bands such as 3.5 GHz and 60 GHz [CDMJ19]. For the same emergency scenario mentioned above in Ghent, Belgium, this can be a good solution, especially when using the 3.5 GHz frequency band in combination with carrier aggregation. Without carrier aggregation, the contribution is limited due to the limited amount of available resource blocks in the 3.5 GHz band. For the 60 GHz band, the main limitation is the high path loss due to the buildings in the city environment.

Mfupe et al. discuss in [MK18] the different problems that should be addressed to use UABS communication as an alternative communication for Disaster Management services, but also to provide affordable broadband connectivity to the hard-to-reach rural communities.

As mentioned in the previous section, due to the flexibility of the UABS location, a swarm of flying platforms can also be considered as an integrated part of the future cellular network. This swarm can inject additional capacity and expand the coverage for hard-to-reach areas or exceptional scenarios such as sports events or concerts. Ahmadi et al. propose a novel layered architecture where Network Flying Platforms (NFPs) of various types - such as UAVs, unmanned balloons or High Altitude Platforms (HAPs)/Medium Altitude Platforms (MAPs)/Low Altitude Platforms (LAPs) - are flying in low/medium/high layers in a swarm of flying platforms and are assumed to be an integrated part of the future cellular network as shown in Fig. 6.21 [AKS17]. The position of the LAPs in the lower layer of the architecture is defined centrally and the NFP has the ability to re-organize the lower layer to achieve its target, which can be capturing as many users, maximizing the achievable rate, and/or fairness among users. The proposed airborne SON systems that reorganizes itself as explained above outperforms an NFP with a fixed placement in the lower layer.

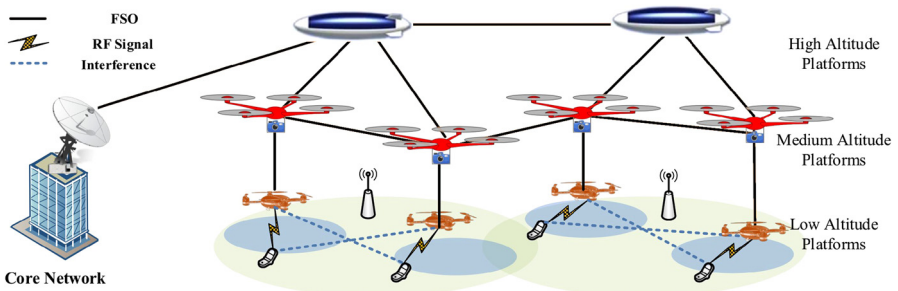


FIGURE 6.21

The hierarchical airborne self-organizing architecture where a variety of UABSs are flying at different altitudes assisting the existing wireless terrestrial infrastructure (Figure taken from [AKS17]).

When talking about UAV communication, one typically thinks about airborne scenarios, however, it can also be of major interest for maritime scenarios. In [FAZ19], three different UAV-based architectures are proposed to provide a solution for reliable, low latency cellular links for search and rescue operations over the sea by using a flying relay, a flying base station, and a flying remote radio head. Another possible maritime scenario is considered in [VNC+15] where the goal of swarms of surface autonomous vessels with multi-hop communication capabilities is to enable maritime tasks such as sea-border patrolling and environmental monitoring, while keeping the cost of each drone low enough to facilitate large-scale deployment. Experiments with their prototypes showed that LOS has critical importance at distance >100 m. The

maximum obtained distance for maximum transmit power is 1400 m. Their test considering the multi-hop technology has also been successful.

6.9 Emerging services and applications

The introduction of 5G and beyond technologies together with the IoT will result in the proliferation of a large number of communicating devices in different smart environments targeting different services and applications. In addition to traditional voice and data services, future inclusive wireless networks will be used to fulfill the needs of a large number of verticals in a variety of smart environments and sectors. Examples include smart homes, smart health, smart manufacturing, smart grids, smart vehicles, smart agriculture, and smart cities. New technical solutions are required to fulfill the diverse requirements of these applications. Within the COST-IRACON action, a number of new and emerging applications were investigated.

6.9.1 Smart grids

The introduction of smart grids results in increasing needs of real-time scalable and reliable monitoring, control and protection of electric power generation, transmission and distribution systems. Monitored state information, e.g. voltages, currents, and user demands, is delivered through a number of nodes to a central unit for performance analysis and different purposes including control and optimization of the grid performance. The technology used for data exchange adopts a Wide-Area Monitoring System (WAMS). Two typical nodes deployed in a WAMS, are Phasor Measurement Units (PMUs) and Smart Meters (SMs). A PMU delivers time-synchronized values of voltage and current phasors as well as other power system related quantities. A smart meter is an electronic device that records the consumption of electric energy and communicates the information back to the electricity supplier for monitoring and billing purposes. In [LJ16] and [AT18], the feasibility of using a Long Term Evolution (LTE) cellular network for real-time smart grid state estimation is investigated. The delay performance is assessed for different cell loads and scheduling schemes with the assumption that all the nodes are connected. It is concluded that with one Physical Resource Block (PRB), 2500 nodes can be accommodated within an LTE cell if the maximum allowable delay is relaxed from 1 s to 2 s. In [AT18] the impact of LTE random access procedure on the reliability of state information delivery was studied. It is concluded that the number of contending devices and cell coverage critically affect the performance in terms of the accuracy of state estimation in smart grids.

6.9.2 Vehicular applications

Vehicle to Vehicle and Vehicle to infrastructure (aka V2X) communications are drivers for road safety and efficiency. Smart and autonomous vehicles pose communi-

cation challenges in terms of delay, reliability, bandwidth, and energy consumption. It is necessary to investigate penetration losses into a vehicle to evaluate network performance and to provide reliable connectivity and large capacity for vehicular use cases such as smart cars and trains. In [Ber17], the LTE coverage was analyzed and the penetration loss was evaluated for a pickup truck in a live network. Depending on the orientation of the vehicle and the environment, the penetration loss can vary from 1.88 to 10.58 dB. A V2X system level simulator was presented in [ND18] including an optimized ray tracer for path loss predictions to study IEEE 802.11p and LTE-V mode 3 and 4 systems.

6.9.3 Public protection and disaster relief systems

Characteristics of current narrowband public protection and disaster relief communication systems include the following: reliability, speech and data transmission capability, point to point, group and broadcast calls, fast call setup, coverage, long battery life, flexibility, prioritization, and security aspects. In [PS17] and [PSP18], the migration towards broadband LTE/LTE-A networks was reviewed. Proximity-based services (ProSe), group communication system enablers (GCSE_LTE) and mission-critical push to talk (MCPTT) are essential features to be implemented. Network resource prioritization, Quality of Service (QoS) provision, and spectrum management are key for coexistence with commercial services. Carrier Aggregation (CA) is presented as a potential solution for an LTE-based public safety and disaster relief network. Issues to consider include negotiation between commercial and public safety service providers, coverage overlap, prioritization, and impact on battery life.