6

Green and Efficient RAN Architectures

Chapter Editor: S. Ruiz, Section Editors: M. Garcia-Lozano, D. Gonzalez, M. Lema, J. Papaj, W. Joseph, M. Deruyck, N. Cardona, C. Garcia, F. J. Velez, L. Correia, L. Studer, P. Grazioso and S. Chatzinotas

6.1 Introduction

The explosive increase of capacity demand has resulted in more dense deployments involving a growing number of BSs. To make future mobile networks sustainable from an economic viewpoint, they should provide more while simultaneously cost less. Installing smaller cells in areas where infrastructure can provide the required backhaul connectivity is the current natural evolution of radio access network (RAN) infrastructure. But heterogeneous networks (HetNets) and small cells require new techniques for configuration, management, and optimisation oriented to self-organising networks (SONs), as well as new RRM algorithms, so that full capacity can be used in the most efficient way, considering not only services but also users' profiles. Another key concept to improve RAN efficiency is referred to as Cloud-RAN, a centralised processing, collaborative radio based in resource sharing and real time cloud computing to adapt to non-uniform traffic. The cloud radio access network (C-RAN) concept is expanded to include also the evolved packet core (EPC) functionalities, in what is known as network Virtualisation. Considering that wireless networks will serve not only a huge number of mobile phones and computers, but also a dense deployment of devices and sensors, effort has to be done to improve the overall energy and spectrum efficiency, combined with opportunistic access to certain frequency bands and cognitive radio (CR) devices.

Chapter 6 deals with all these topics and summarises advances done by European cooperation in science and technology (COST) IC1004 researchers. Sections 6.2 and 6.3 are devoted to explain advanced resource management ecosystem for both, cellular and wireless networks considering resource scheduling, interference, and power and mobility management. Section 6.4

Cooperative Radio Communications for Green Smart Environments, 195–270.

describes recent energy efficiency enhancements by analysing power consumption, cell switch off, and deployment strategies. Section 6.5 is focusing on spectrum management optimisation and cognitive radio networks with tools and algorithms specifically for TVWS. Section 6.6 defines and models the virtualised and cloud architecture by proposing new algorithms under realistic (RL) scenarios. Finally, Section 6.7 is addressed to re-configurable radio for heterogeneous networking by including recent progress in resource allocation, diversity, and adaptive antenna techniques to improve capacity and energy efficiency.

6.2 Resource Management Ecosystem for Cellular Networks

6.2.1 Radio Resource Scheduling

In the context of long-term evolution (LTE) networks, the packet scheduler can be defined as the selection of the user to be served in every 1 ms transmission time interval (TTI) and frequency resource block, that is to say, the resource allocation is made in both time domain and frequency domain. A good design of a scheduling algorithm should balance the trade-off between the maximisation of the system throughput and fairness among users.

The two basic types of scheduling that have been defined in LTE are dynamic and semi-persistent. In the first case, the scheduler reacts to user equipment (UE) requests opportunistically. Resources are dynamically adapted to users according to their buffer status and channel conditions, but at the cost of an important layer 1 (L1)/layer 2 (L2) load, carried by the physical downlink control channel (PDCCH). The semi-persistent option is more adequate for services making use of small but recurring packets, such as voice over IP (VoIP). In this case, only the first assignment needs to be signalled and the same allocation is kept in subsequent transmissions, with the corresponding savings in the L1/L2 signalling. The policy is *semi*-persistent in the sense that the allocation can be changed at some point, for example to provide certain link adaptation. On the other hand, the lack of flexibility implies a less efficient use of radio resources.

LTE networks are continuously being improved through the different 3GPP releases, which means new challenges and open issues to be considered when designing scheduling policies. Examples are the problem of L1/L2 signalling, self-optimisation of scheduling policies, specific uplink (UL) issues, and operation through different aggregated carriers. All these topics have been a matter of research in the context of COST action IC1004 and are covered in the following lines.

The impact of control channel limitations has been widely evaluated for VoIP but missed for non-real time services. González et al. [GGRL11] carefully modelled practical limitations of PDCCH and studied the trade-offs associated to control channel usage and the provision of quality of service (QoS) for this type of services. Thus, covering an aspect omitted in previous literature. The PDCCH is critical in LTE networks for a correct functioning of scheduling algorithms, since, among other information, it carries DL and UL scheduling assignments. The authors conclude that the selection of scheduling and QoS parameters is clearly affected by the limited amount of control channel resources in LTE. Results suggest that in order to let the system operate efficiently from the radio resource allocation perspective, it is important to carefully characterise the system performance. Not only considering traffic features, but also the scheduling policy and availability of resources for the PDCCH. This is particularly true for cases in which users mostly target low bit rates, in this case the system capacity would be seriously penalised.

Scheduling policies have been widely studied in the last years. However, the complexity of the LTE scheduler, typically divided in a time domain policy plus a frequency domain one, leaves a lot of margin to improve classic schemes. One of the novel ideas proposed by Comsa [CAZ+13] et al. is a new scheduling approach able to find the optimum scheduling rule at each TTI, rather than using one single discipline across the whole transmission as typically done. This idea is investigated along several research works that are interrelated and constitute a very detailed description of this novel strategy.

In the work of Comsa et al. [SMS+12b], the authors propose to assess the corresponding scheduling utility function in a TTI basis, rather than a global and unique evaluation. Thus, the problem to be solved is the selection of proper rules to achieve local optimisation that would yield a higher global utility. Since evaluating all possible rules in a 1-ms time scale would lead to an unaffordable complexity, a Q-learning approach is proposed and investigated. The algorithm uses reinforcement based on rewards from previous transmission sessions, so that prior experiences are turned into permanent. However, since convergence might require excessive number of evaluations, the authors extend the approach to work as a multi-agent (parallel) system.

The description of several scheduling rules is done in Comsa et al. [SMS+12c]. Also trying to capture the well known trade-off between system throughput and fairness, a candidate objective function being an aggregation of the Jain index and normalised system throughput is proposed. Initial performance results in synthetic scenarios (pedestrian and vehicular) indicate a successful operation. The evaluation is done in terms of new quality metrics named global utility, tradeoff utility, and supreme utility.

The scheduler should dynamically self-tune and respond to the network state. Hence, given a certain array of reported channel quality indicator (CQI) values and the distribution of normalised user throughput, the best scheduling rule is decided. However, one of the problems of Q-learning is the exploration–exploitation compromise. Since the number of possible states is extremely large, it is unfeasible that all states had been explored in the past, thus the authors introduce the use of a neural network working as function approximator [SMS+12a]. Results indicate a superior performance of this approach when compared to standard schedulers. The complete architecture of the scheduling policy is depicted in Figure 6.1.

The authors propose extensions in the form of a self-optimisation strategy that adjusts the level of system fairness [CAZ+13]. In this case the trade-off throughput versus fairness is studied by evaluating the cumulative distribution function (CDF) of the normalised user throughput, which provides a more complete picture of the system performance. The key point is that a level of accepted fairness should be provided and the scheduler should be kept on that range of tradeoff. A possible fairness criteria is guaranteeing that $(100–x)\%$ of users achieve at least $x\%$ of the normalised user throughput. This way, under or over fair situations can be detected and avoided. This is graphically depicted on Figure 6.2. The fairness level ranges from pure

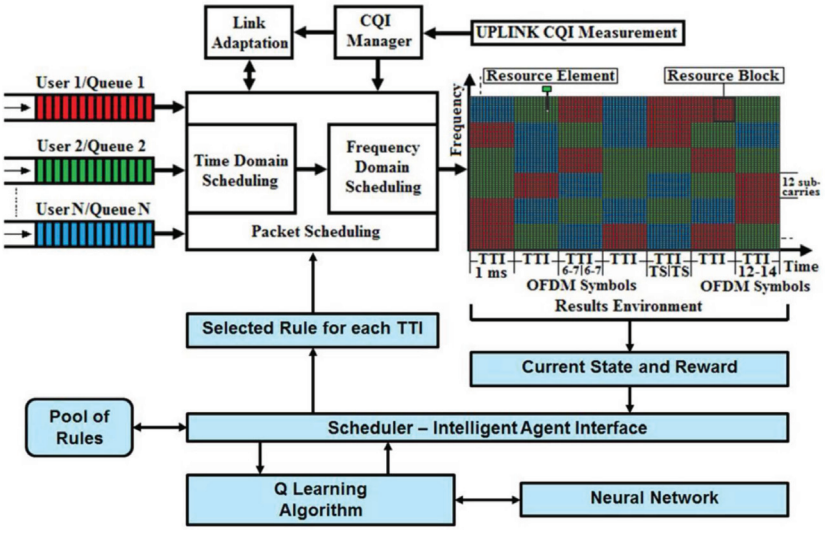


Figure 6.1 Long term evolution - advanced (LTE-A) packet scheduler framework from Comsa et al. [SMS+12a].

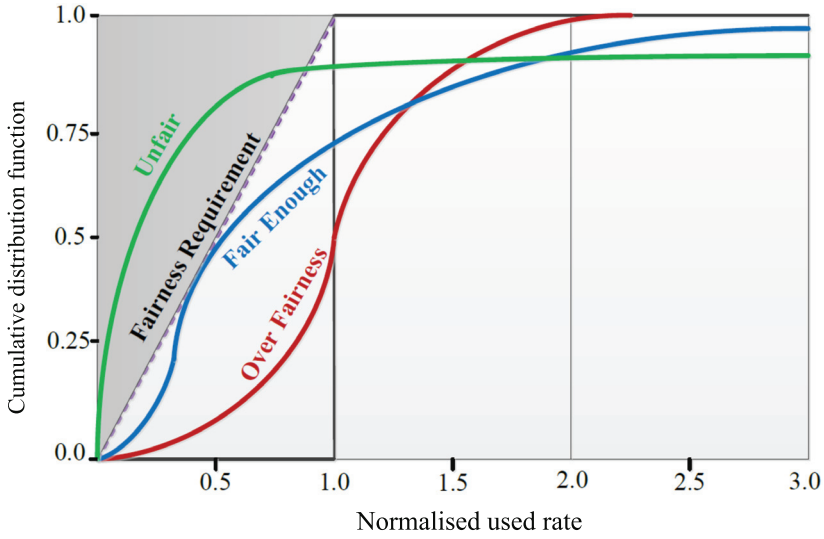


Figure 6.2 Fairness requirement from the CDF point of view [CAZ+13].

proportional fairness to a classic maximum carrier to interference ratio (CIR) selection, to be dynamically controlled by a parameter α. The evaluation of all possible α is infeasible at every TTI; however, this type of brute force search is not necessary since, as previously indicated, reinforcement learning is a component part of the approach. Four reinforcement learning techniques are proposed and described in this work. Results are further improved by Comsa et al. [CZA+14], where the previous idea is extended to a double parametrisation scheduling policy. This way, two parameters α and β are optimised to get the desired feasible state, situation with the required throughput versus fairness trade-off. Given the more restrictive domain of the parameters, a faster convergence time is obtained, which is an interesting feature for real time applicability. Improvements in terms of throughput and percentages of TTIs with a feasible state are achieved with respect to previous approaches. A second benefit is that lower fluctuations are obtained when the traffic load drastically changes, indeed the mechanism shows the ability of recovering a feasible state in less than 10 TTIs when severe changes in the traffic load and user channel conditions appear.

Allocation of resources in the UL poses new challenges. For example, in the first two releases of LTE the use of SC-FDMA imposes a contiguous allocation and equal power distribution. This implies that finding an optimal solution to the problem of ergodic sum-rate (SR) maximisation under proportional rate

constraints requires high-computational complexity and existent proposals are unsuited for practical applications. Cicalò and Tralli [CT14a] address this problem and proposes a novel sub-optimal heuristic solution after using Lagrangian relaxation of rate constraints. In particular, in order to reduce the search space, the authors make use of a linear estimate of the average number of sub-carriers that are allocated to each user when the optimal rate is achieved. After posing the dual problem, it is possible to exploit an adaptive implementation to build the heuristic algorithm, which either dynamically tracks the per-user channel condition and estimates the number of allocated subcarriers at each TTI.

The approach is compared against an optimal solution and a quasi-optimal from the existent literature. The SR gap with respect to the optimal solution is limited to the 10%, whereas the number of required iterations is two orders of magnitude lower. The complexity of the algorithm just increases linearly with the number of users and number of resources, thus being a much more interesting option for practical operation. The algorithm also preserves long-term fairness for homogeneous full-buffer traffic, with a Jain Index greater than 0.99 in all the investigated cases. However, in order to improve the real applicability of the scheduler, short-term fairness evaluation is also required. The study by Cicalò and Tralli [CT14b] shows that the proposal provides excellent results when evaluated over time windows larger than 200 ms. Hence, the scheduler results in an attractive solution for applications where QoS requirements depend on the rate averaged over intervals of that order of magnitude. Finally, under heterogeneous traffic conditions, the scheduler is able to achieve a higher ratio of satisfied users with granted bit rate services when compared to proportional fair-based schedulers. All this, with a significant reduction of the packet delay and without starving non-granted bit rate users.

Abrignani et al. [AGLV15] present an approach of UL scheduling for LTE networks deployed over dense HetNets. In particular for the co-channel layout, in which macros and small cells operate at the same carrier, thus with strong inter-cell interference. The problem is tackled from two different viewpoints, first a three-step-based algorithm and, second, a method that solves the scheduling optimisation at a time. These algorithms are posed as mixed integer linear programing aiming at the maximisation of throughput, optimisation of radio resource usage and minimisation of inter-cell interference. Special care is put on the evaluation of their computational feasibility and so their implementation in the standard, considering that the LTE scheduler has to be executed every 1 ms. After solving the programs, the authors observe that the compact case performs better and executes faster, though still not being compliant with the execution time requirement. The algorithms are then

compared against a heuristical greedy approach. Results indicate that just in the 90% of the cases, it reduces the performance with respect to the optimum by less than 10%. Besides, in this case the solution is met in less than 0.75 ms in more than 80% of the cases and in less than 1 ms in 100%.

The previous work was contextualised in the case of localised single carrier frequency division multiple access (SC-FDMA). With the introduction of LTE-A (release 10 and beyond) non-contiguous resource allocation is also possible in the UL, though with less flexibility than the downlink (DL). In particular, up to two separate clusters (sets of contiguous sub-carriers) can be allocated. The aim is to increase the spectral efficiency by exploiting users frequency diversity gain. On the other hand, this new transmission scheme brings an increase in the signals peak-to-average power ratio (PAPR) although still lower than pure orthogonal frequency division multiple access (OFDMA). Maximum power reduction (MPR) is introduced in 3GPP when multi-cluster transmission is used so that the general requirements on out of band emissions are met. On the other hand, this power de-rating implies and extra loss in the UL link budget that may result in throughput reductions.

Lema et al. [LGRO13] propose a novel scheduler that considers opportunistically the information on the MPR that is to be applied. The packet schedulers main task is to evaluate the gain or loss in throughput of the multi-cluster transmission over a conventional contiguous one. Based on the sounding reference signal (SRS) channel estimation, the eNB can predict the multi-cluster transmissions performance. In order to assess the performance of considering MPR wise scheduling decisions, it has been compared to other three benchmarks: Pure multi-cluster transmission, contiguous allocation, and a previous proposal of the literature in which multi-cluster transmission is pre-defined by a given path loss threshold. Also different cluster sises have been tested as the MPR to be applied strongly depends on this parameter. In all cases, the new approach has presented throughput gains. Enabling the MPR information in the scheduler adapts the transmission mode to each particular case, enhancing 30% cell edge throughput compared to a pure multi-cluster transmission. Average throughput is increased almost 18% when compared to classic contiguous allocation.

With the increase of supported bandwidths by modern cellular systems, spectrum aggregation (SA) has become mandatory. Indeed, it is a solution for the highly fragmented spectrum that operators typically have. However, this imposes new challenges over schedulers, that now must be able to handle inter carrier resource allocations. Acevedo Flores et al. [AFVC$^+$14] analyses the cost/revenue performance of a mobile communication system in an international mobile telecommunications (IMTs) advanced scenario with SA over

the 2 and 5 GHz bands. A system accounting for a general multi-band scheduler is compared against another in which users are allocated to one band without possibility of changing it after this first association, throughput gains of 28% are obtained. From the economic viewpoint, costs, and revenues are analysed on an annual basis for a 5-year project duration. Revenue and cost is evaluated, for different prices per MBytes (impacts revenue) and cell radius (impacts cost). For the worse study case, profits of 240% are obtained when the multi-band scheduler is introduced and 170% without it. Finally, an energy efficiency strategy is proposed and analysed. It opportunistically reallocates user to available bands with better propagation conditions. This yields a reduction in the transmission power, thus showing an additional benefit of the multi-band allocation scheme. It reduces down to 10% the required energy when compared to a solution with no possibility of link reallocation between bands.

Robalo and Velez [RV14] constitute an extension to the previous work, now in the context of LTE and for the 800 and 2.6 GHz bands. In this case, two multi-band schedulers are implemented and compared. The first one allocates users to a single carrier at a time, this is smartly done by using integer programming optimisation. The total number of users on each band is upper bounded by the maximum load handled, that is to say, the scheduler takes into account a per-band admission control constraint. The second scheduler is able to allocate users to one or both components and indeed this is the reason why the same integer programming approach is not feasible, due to prohibitive computation complexity. Thus, a per user scheduling metric is computed and the allocation is performed starting from the highest value. A complete evaluation is done from several viewpoints. The second case outperforms the first one and other benchmarks in terms of packet loss ratio and delay. A quality of experience (QoE) model for multimedia applications, computed as a function of packet loss ratio, delay and bit rate is also evaluated with maximum performance for the approach dealing with both bands jointly. Similar to the previous work, a cost/revenue analysis is performed, with similar profits for both approaches but far superior to the case without carrier aggregation.

One of the constraints in the previous work was considering admission control requisites in each band. More generally, the admission control in LTE has to consider the load situation in the cell when admitting a new UE in the system. A new request is only granted if the QoS for the new user is going to be satisfied without jeopardising the already connected ones. Lema et al. [LRFGL$^+$12] present an scheme that estimates the new user demands but in the context of UL. Thus, the approach is based on the SRS. The pre-allocation of the sounding bandwidth is crucial for assessing correctly the user demands.

The criteria followed to allocate the sounding bandwidth is to update those spectrum regions that have oldest channel information. With this, a complete channel information across the system bandwidth is obtained. The proposed process is compared to an admission algorithm that pre-sets a maximum data rate per eNB. Each UE is assumed to consume the guaranteed bit rate. Simulation results show that the resource consumption estimation through SRS increases the number of UEs admitted to the system. There is an overall improvement of the cell rate at the expense of increasing the outage probability.

In the previous works, the research methodology closely relies on system level simulations. Indeed, the definition of new simulation strategies able to correctly model carrier aggregation deployments is a key issue. [RVPP15] describes an open source and freeware extension to the LTE-Sim simulator, which implements carrier aggregation functionalities. The work explains the degree of parametrisation that LTE-Sim allows and describes the three available DL multi-band schedulers. First, an integer programing based one that aims at maximising the cell goodput, second, a basic multiband scheduling that allocates users sequentially to the available carriers, and third an approach introduced in the work itself. In this last case, unlike the previous proposals, the user can be allocated to several bands simultaneously thus providing an improved frequency diversity. The authors close their contribution by providing a comparison of scheduling policies simulated with LTE-Sim. Results quantify average cell packet loss ratio, delay, and application level throughput in a reuse 3 deployment and indicate a superior performance of multi-band scheduling.

The allocation of resources in the data and control plane has been traditionally done by the same eNB. However, a plane split might yield additional cost savings [Zha14]. This way, data would be scheduled by cells without signalling, which opens the door to dynamic capacity activation, with the corresponding energy saving. This is graphically depicted by Figure 6.3. Zhang [Zha14] provides an insight on this issue with a particular emphasis on modelling of control signalling traffic and mobility management.

6.2.2 Interference Management

Without a doubt, interference mitigation is one of the most important elements in cellular systems. This subsection presents interference mitigation strategies and proposals for cellular networks based on OFDMA, in the DL, and SC-FDMA, in the UL, such as LTE and LTE-A. Both access schemes provide intrinsic orthogonality to the users within a cell, which results into an almost null level of intracell interference. However, intercell interference (ICI) is

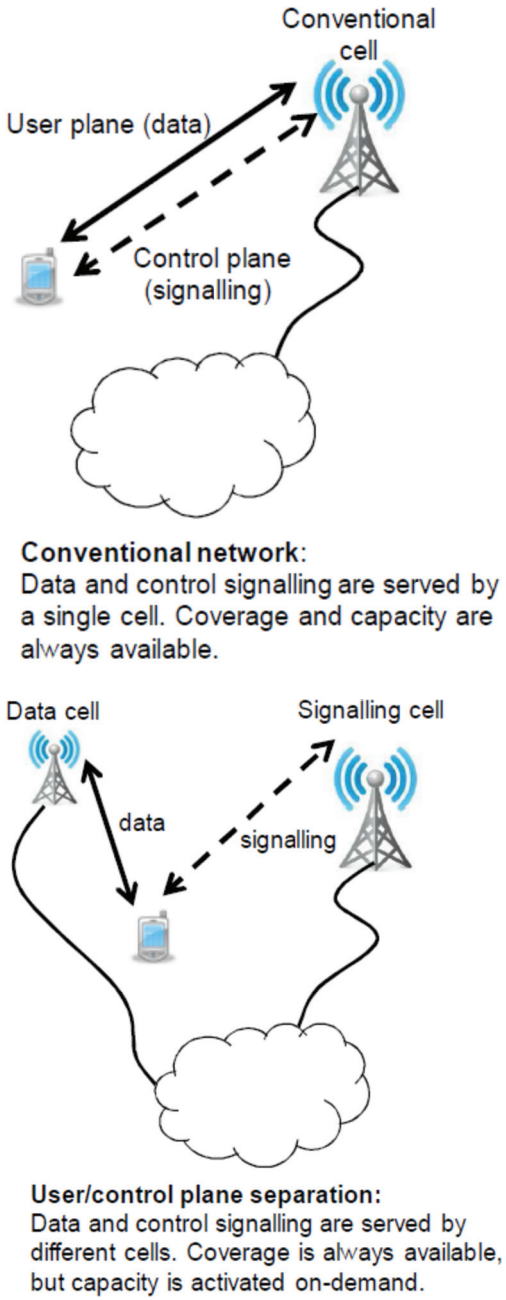


Figure 6.3 Brief concept for and control plane separation.

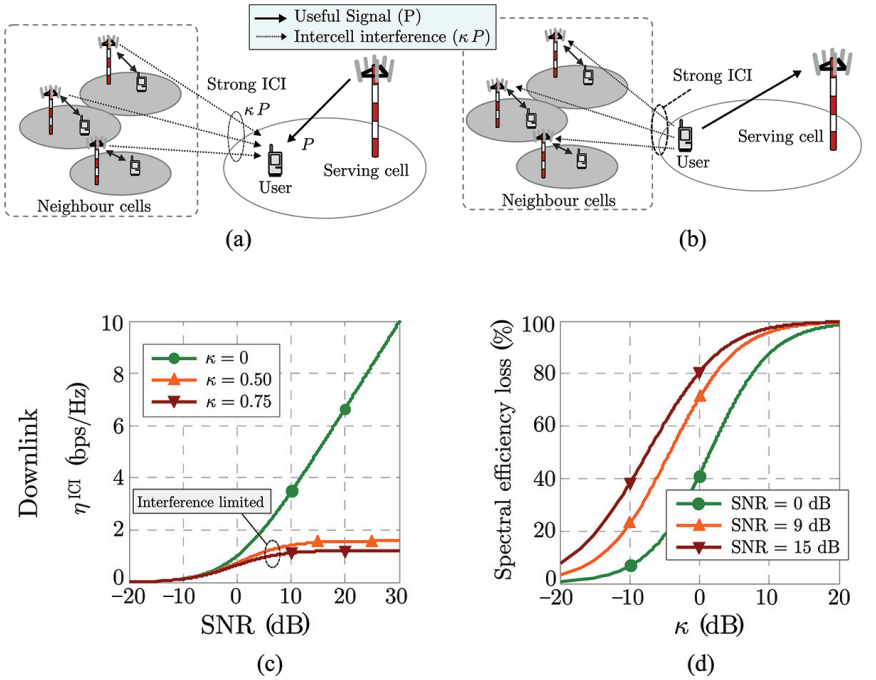


Figure 6.4 ICI in OFDMA-based cellular systems: (a) Intercell interference: downlink, (b) Intercell interference: uplink, (c) ICI: impact on capacity, and (d) ICI: losses.

created when the same channel is used in neighbour cells. Hence, the target of the strategies and techniques presented hereafter is to reduce or avoid ICI.

6.2.2.1 Essentials

In general, the rate at which the network is able to transmit depends on the signal-interference plus noise ratio (SNR). In order to introduce some fundamental notions related to ICI and its impact on the performance of cellular systems, the scenario shown in Figure 6.4 is considered. Figures 6.4(a) and (b) illustrate the case of DL and UL, respectively. As it can be seen, in the DL, ICI is generated by neighbour cells while in the UL, it is created by UE. The following analysis focuses on the DL, but it can easily be extended to the UL. According to the Shannon's formula, the spectral efficiency measured in bit/s/Hz of the *User* (Figure 6.4(a)) can be written as follows:

$$\eta^{\mathrm{ICI}} = \log_2\left(1 + \frac{\text{Useful signal received power}}{\sigma^2 + \text{ICI received power}}\right), \tag{6.1}$$

$$\eta^{\mathrm{ICI}} = \log_2\left(1 + \frac{P}{\sigma^2 + kP}\right). \tag{6.2}$$

σ^2 represents the noise power. In case that $kP \gg \sigma^2$, the scenario is interference-limited, and obviously, this is the case of interest from the interference mitigation standpoint. Note that the amount of ICI can always be expressed as kP, and hence, it can be understood that the goal of any interference mitigation strategy is to make k as small as possible. In order to illustrate this reasoning, the impact of ICI on the spectral efficiency is considered. Given that the signal-to-noise-ratio (SNR) and signal-to-interference ratio (SIR) correspond to $\frac{P}{\sigma^2}$ and k^{-1}, respectively, the spectral efficiency without ICI can be written as follows:

$$\eta^{\mathrm{NoICI}} = \log_2(1 + \mathrm{SNR}). \tag{6.3}$$

The spectral efficiency loss due to ICI will be given by:

$$\mathrm{Loss} = \frac{\eta^{\mathrm{NoICI}} - \eta^{\mathrm{ICI}}}{\eta^{\mathrm{NoICI}}} = \frac{\log_2\left(\frac{1+\mathrm{SNR}}{1+\left(\frac{1}{SNR}+k\right)^{-1}}\right)}{\log_2(1+\mathrm{SNR})}. \tag{6.4}$$

Figures 6.4(c) and (d) clarify the meaning of Equation (6.4) by showing the relationship among the involved variables. Figure 6.4(c) clearly shows that when ICI is high ($k \geq 0.50$), increasing P (to increase the useful signal received power) has a negligible effect on the spectral efficiency. Hence, ICI is the main capacity limiting factor. In addition, Figure 6.4(d) shows that in scenarios with relative low noise (SNR = 15 dB), the capacity is reduced between 40% and 80% when the SIR goes from –10 dB to 0 dB ($k^{-1} \in$ [–10, 0] dB), makes evident the severe impact of ICI. Indeed, the losses in the heavily interfered region (k $\geq$ 0 dB) are above 80%. In the light of the previous analysis, the need for effective interference mitigation is clearly justified. Additional practical aspects that further complicates the problem at hand, that of reducing or avoiding ICI include:

1. ICI is highly non-predictable (even in environments without mobility) due to the time-varying transmission patterns in neighbour cells. Moreover, frequency selective fading and practical limitations of channel state information (CSI) feedback schemes make harder to have accurate estimations of ICI. As it will be shown shortly, some of the proposals are aimed at overcoming this issue by employing predictive techniques, such as by Garcia et al. [GLG11], while other techniques are focused on

improving the *quality* of the CSI, see González et al. [GGRO12a] and Lema et al. [LGR14].

2. SINR levels are not uniformly distributed in the network coverage area. Indeed, users located near to cell edges receive not only weak signals from their serving base station (BS) but also high ICI. Thus, the QoS experienced by users strongly depends on the location in the network coverage area, and as a result, a *fairness* issue among users is created. An example of strategies focused on *protect* those *unlucky* users include [KMM11].
3. Finally, but no less important, the trend towards *densification* as a paradigm to increase the areal frequency reuse, and hence, as a mechanism to increase capacity, makes evident the pevalence of the techniques for effective frequency planning and interference mitigation. Indeed, as it is indicated in Gonzlez et al. [GGRO12b], the notion of frequency reuse is key in many of the proposals whose main target is to reduce the amount of ICI, mainly at cell edges, such as Peng et al. [PKHAM13] and Acedo-Hernández [AH13].

6.2.2.2 Classification of strategies for interference mitigation

Classifying the strategies for interference management is not an easy task because several criteria can be used. In addition, the boundaries in this context are blurred. However, a widely accepted classification is shown in Figure 6.5. In a nutshell, interference mitigation can be done by means of:

1. Coordination, where restrictions to the radio resources (mainly bandwidth and power) are applied at each cell (with more or less dynamism) to reduce ICI at cell edges,
2. Cancellation, where interfering signals can be either subtracted from the received signal,
3. Randomisation, where the interference is distributed uniformly across the system bandwidth through scrambling, interleaving, or frequency-hopping (spread spectrum).

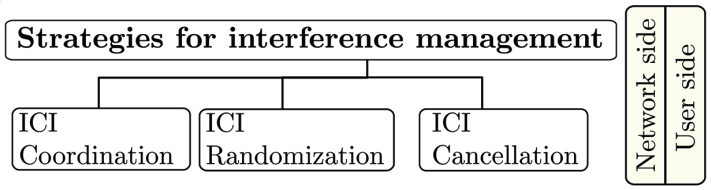


Figure 6.5 Strategies for interference management.

It is worth mentioning that, although most of the practical schemes are fully implemented in the newwork side, these strategies can also be implemented in the user side, and often in both. What technique and how is implemented mainly depends on the type of link (DL or UL). Other elements to take into account include: the type of environment, mobility conditions, and computational complexity, among others.

6.2.2.3 Static intercell interference coordination (ICIC): an approach to planning

As it was mentioned before, coordination implies apply a set *rules* in the network. This rules can be applied at different time scales, and hence, it is possible to talk about *static* and *dynamic* ICIC. Thus, in static ICIC the basic idea is to apply different frequency reuse factors to different groups of users based on a certain criterion (usually average SINR measurements are considered) over periods of times ranging from days to weeks. In static ICIC, there two fundamental schemes: soft frequency reuse (SFR) and FFR. The operational principle is depicted in Figure 6.6. In short, once users are classified (e.g., based on SINR_{avg} in the figure), the resource allocation is

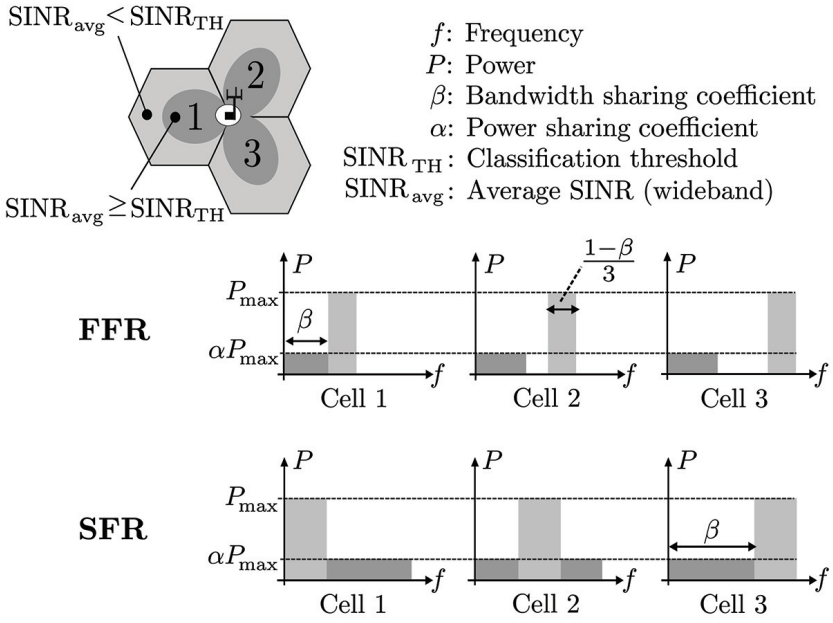


Figure 6.6 Operational principles in static ICIC.

done respecting the bandwidth and power allocation pattern that is defined by the paraneters α and β as it is shown. Further information can be found in Gonzalez et al. [GGRO12b].

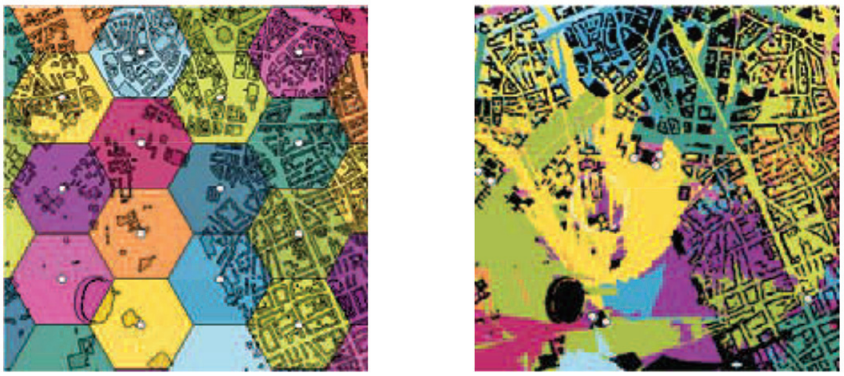
One fundamental problem with static ICIC is the calibration of the aforementioned parameters. While such selection is quite straightforward is perfectly hexagonal cells [GGRO12b], the problem is much harder in RL cellular networks featuring irregular layouts [GGRL13] (Figure 6.7). In Peng and Kürner [PK13a], the authors proposed a method for SFR planning, in the DL, that takes into account an interference metric that is inversely proportional to the distance between interferers, i.e., BS. Thus, for two BS i and j, the interference metric is computed as follows:

$$m_{ij} = \begin{cases} \frac{1}{d_{ij}^2}, & i \neq j, \\ 0 & i = j. \end{cases} \tag{6.5}$$

Therefore, the objective of the proposed iterative scheme, based on the Gauss-Sidel algorithm, is to minimise the overall interference (IS) expressed as a function of the previous metric, thus

$$IS = \sum_{c=1}^{C} \sum_{i,j \in M(c)} m_{ij}, \tag{6.6}$$

where C is the number of available channels and $M(c)$ is the set of cells in which the c^{th} channels is being used. In this manner, it can be seen that the



(a) (b)

Figure 6.7 Cellular layouts: (a) Idealised hexagonal cells, and (b) Realistic irregular cells.

proposed method realies on the assumption that ICI is proportional to the path loss in order to update the channel allocation. This scheme is a clear example of how static ICIC can be applied in a planning-like fashion to determine frequency–power profiles. Indeed, taking ICIC as design criteria in more general planning tasks is quite common. For instance, in Acedo-Hernández [AH13], the planning of the physical cell identifier (PCI) is studied from an interference management point of view and potential gains are quatified. PCI planning is important because it defines the location of the cell-specific reference signals (RS). RS (often called *pilots*) are always transmitted in the same orthogonal frequency division multiplexing (OFDM) symbol, but in the frequency domain, RSs are shifted determined by the value of the PCI. Since RS are used for CSI feedback (among other thinks), an improper PCI planning may give inaccurate SINR estimates, which leads to inefficient data transmission in the DL. The analysis presented in Acedo-Hernández [AH13] considers several plan models for the primary synchronization signal (PSS) which in conjuntion with the PCI determine the exact location in time and frequency of subcarriers where RSs are transmitted. Key performance indicators (KPIs) include figures obtained from the distribution of the SINR observed in RSs. To be more precise, the median and fifth percentile, as measures of connection quality, and statistics from the CQI and DL throughput. Figure 6.8 shows one representative, yet important, result presented in Acedo-Hernández [AH13]. As it can be seen, the PCI planning has a significant impact of the resulting CDF of the cell throughput. Figure 6.8(a) shows that, for low network load, average cell throughput decreases by up to 30% with respect to

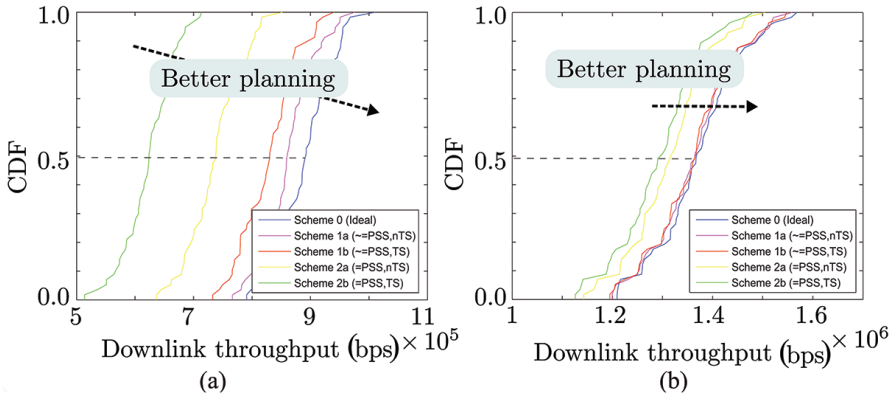


Figure 6.8 Impact of PCI planning on DL throughput (taken from [AH13]): (a) Low load scenario, and (b) High load scenario.

the best planning scheme. For high network load, shown in Figure 6.8(b), a similar behaviour is found, although gains are smaller. Further details about simulation setups and additional results can be found in Acedo-Hernández [AH13]. However, the study makes evident the impact of minimising the effect of ICI even in planning tasks. Obviously, as it was mentioned earlier, ICIC is also suitable to be applied in smaller time scales.

6.2.2.4 Dynamic ICIC

Dynamic coordination implies adapt the bandwidth/power allocation patterns in time scales ranging from milliseconds to minutes. This can be done by means of generic resource allocation strategies, but it can also be done starting from a predefined pattern, such as FFR. The latter approach was employed in Krasniqi et al. [KMM11]. In this scheme, a model based on two users is used to determine the optimal FFR-based resource allocation. In particular, three different cases are considered: (i) both users are in the *inner* region (high SINR), (ii) both are in the *outer* region (low SINR), and (iii) there is one user in each region. Closed form expression are obtained by means of Lagrange multipliers. Therefore, the proposed can easily be extended to scenarios with multiple users to achieve dynamic resource allocation. Details of the formulation and assumptions can be found in Krasniqi et al. [KMM11]. Figure 6.9 shows the simulation results for the average SR taken over uniform users positions in the inner and outer regions versus maximum BS power. The lower curve represents the average of all SRs when no re-allocation of the outer bandwidth to the inner users is carried out for the static method used. A better performance in terms of average SR is achieved when we

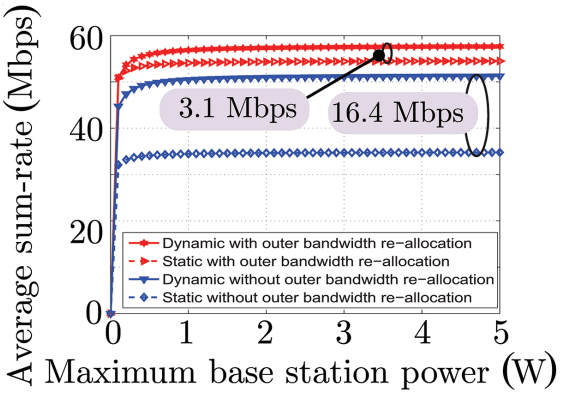


Figure 6.9 Maximum average SR (taken from Krasniqi et al. [KMM11]).

consider the dynamic method. Clearly, performance increase of approximately 16.4 Mbps is achieved when dynamic method is used. Considering also the outer bandwidth re-allocation to the inner users a performance increase of 3.1 Mbps is achieved when dynamic method is used. Hence, in the light of this results, the effectiveness of interference mitigation is demonstrated.

So far, strategies aiming at avoiding ICI have been reviewed. A completely different approach is just focus on cancelling its effect.

6.2.2.5 Topological analysis of ICI cancellation

In ICI cancellation, the basic concept is to regenerate the interfering signals and subsequently subtract them from the desired signal. In short, the approach is removing ICI rather than avoiding it. One good feature of interference cancellation is that the implementation at the receiver side can be considered independently of the interference mitigation scheme adopted at the transmitter, and hence, the coexistence with other techniques is not an issue. Some types of ICI cancellation include: interference rejection combining (IRC), successive interference cancellation (SIC), and interference alignment (IA).

One very interesing aspect of cancellation techniques is that they can modify significantly the reception area where transmitters can be *heard.* The work presented in Haddad [Had14] investigates this. Figure 6.10 shows the basic notions in this framwork. Transmitter s_x is supposed to be decoded by receiver r_x, and the region $H(s_x)$ is the region where transmitter s_x can be successfully decoded. Figure 6.10(a) shows zones $H(s_1)$ and $H(s_2)$ for s_1 and s_2, respectively, under the assumption of non-uniform power allocation, i.e., the transmit power of s_1 and s_2 are different. It can be proved that these two demands cannot be satisfied when both s_1 and s_2 transmit with the same power. The uniform power allocation is shown in Figure 6.10(b). Note that $r_1 \notin H(s_1)$, and hence, r_1 does not receive its information. In contrast, when

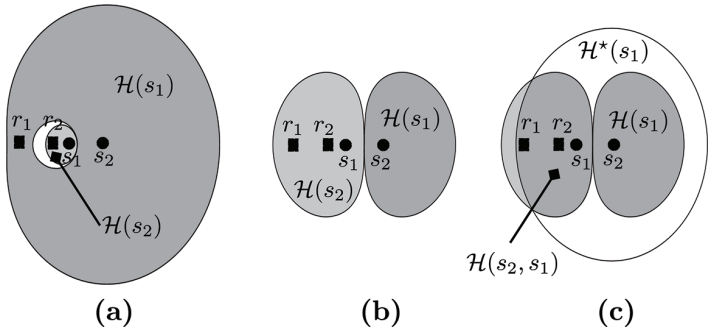


Figure 6.10 Topology with and without SIC (taken from Haddad [Hadl4]).

SIC is applied in r_1, it can first decode s_2, and then, it can cancel it to decode s_1. Thus, with SIC the two demands can be satisifed even with uniform powers as it is shown in Figure 6.10(c). Note that $H(s_2, s_1)$ is the intersection of two convex shapes, $H(s_2)$ and $H(s_1)$, where the latter (shown as an empty circle) is the reception zone of s_1 if it had transmitted alone in the network. It is clear that the total reception area of s_1 with SIC is much larger than without it. The most important contribution presented in [Had14] is that the authors present a tighter bound for the number of reception zones in terms of a novel metric called *Compacteness Parameter,* among other interesting results.

6.2.2.6 Enhancements through prediction and improved CSI feedback

As it was mentioned before, one fundamental problem with ICI is that it is highly unpredictable in real-life contexts. In this sense, several efford have been made to provide solutions to cope with this issue. The work presented in Garcia et al. [GLG11] introduce a model to predict the power variations in neighbour cells. Basically, the proposed framework try to avoid some drawbacks of distributed schemes, such as game theory that require static scenarios and sequential or per-round allocation to converge to equilibrium that may not be optimum solutions. Other practical radio resource management (RRM) schemes are difficult to implement due to their complexity or simply because they assume a global and perfect knowledge of the system. Thus, in order to avoid such difficulties, the proposed solution makes an assumption on systems evolution. The distributed power allocation, that does not require information sharing among transmitters, requires three phases:

1. *Estimation.* The channel is computed as a polynomial regression of degree 4, based on the five last measurements.
2. *Prediction.* The target is to determine the *a priori* values of the channels and interference for future steps. Finite-horizon-based prediction is considered.
3. *Decision.* Each BS produces a partially predictable interference, performing a trade-off between its inertia and its required power variations.

Another approach that can be taken to improve the performance of interference management schemes, and more precisely static ICIC, is to improve the CSI feedback. CSI is required both in DL and UL in order to allow for *opportunistic* scheduling. The schemes presented in González et al. [GGRO12a] and Lema et al. [LGR14] introduce mechanisms to improve the accuracy of CSI. To feedback the DL, UE transmit CQI-based reports through the UL to provide

the network with SINR estimates at different subbands. In the UL, sounding signals are sent by the UE and configured by the network. The fundamental idea in González et al. [GGRO12a] and Lema et al. [LGR14] is to take advantage of the bandwidth allocation pattern that is used in static ICIC schemes, such as SFR and fractional frequency reuse (FFR), to focus the CSI feedback on the bandwidth portions that are relevant to each user. In this manner, CSI reports are updated more frequently, and hence, the modulation and coding scheme selection can be done in an effective manner. In case of Lema et al. [LGR14], the solution reduces the time delay between sounding measurements, while in González et al. [GGRO12a] also the accuracy of the wideband CQI is improved. In both cases, results show significant performance in terms of average user throughput and cell edge performance, thus achieving one of the main targets of interference management.

6.2.3 Power Management

Power control is a key degree of freedom that has always been of paramount interest in wireless communications. The intelligent selection or adjustment of the transmitter power may improve the general performance in terms of interference, connectivity, and energy consumption.

For network deployment planning and optimisation, it is necessary to determine the enhanced node-B (eNB)-transmitted power which depends on the average SINR, the desired cell radius and the frequency reuse being adopted [AFRJ14]. If the average SINR is maximised the BS transmitted power increases to very high levels, which compromises the eNB energy consumption. So, average SINR levels should be maintained lower than the maximum. Also, for a pre-set value of SINR as the cell radius increases in planning, the transmitted power demands also increase. In deployments with SA, with different frequency bands, if a constant average SINR and same cell radius across bands is desired, different power allocations must be considered.

In the UL, the 3GPP proposed an open loop power control (OPLC) which adjusts the power spectral density of the mobile terminal (MT) based on the estimation of the channel gain. The main advantage of this power control method is that the path loss can be partially compensated, which is known as *fractional* power control. Hence, UE with high path losses transmit at lower power (than classic full compensation) thus generating less interference.

The UE calculates the OLPC over the allocated bandwidth by estimating its own path loss. The algorithm mainly compensates for long term channel variations (i.e. path loss and shadowing), yet the performance degrades due to

errors in the path loss estimation. For this reason, the algorithm also accounts of a closed loop component that aims to compensate these errors by sending feedback corrections to the UE periodically.

The power transmitted by a given UE in the Physical UL Shared Channel (PUSCH) is defined as:

$$P = \min(P_{\max}, P_0 + 10 \log M + \alpha L + \delta_{\mathrm{TF}} + \delta_i), \tag{6.7}$$

where $P_{\max}$ is the maximum available power in the UE, P_0 is both cell and user specific parameter, M is the number of resources allocated, α is the path loss compensation factor, L is the estimated path loss, δ_{TF} is a parameter that depends on the transport format chosen, and δ_i is the UE-specific closed loop correction.

Concerning the OLPC, P_0 and α are the most important parameters to be adjusted. P_0 controls the SINR target at the receiver end. A rise in P_0 increases the transmitted power, hence, interference rises and degrading the system throughput. However, if P_0 is low the UE may not accomplish the bit error rate (BER) requirements. The path loss compensation α, brings the system into a trade-off. If a UE placed on the cell-edge corrects only a part of its path loss it generates less inter-cell interference. Also, as α approaches one, all UE compensate their path loss so all signals are received with the same strength, resulting in similar SINR values along the cell area. On the other hand, when α is reduced, MTs compensate only a fraction of its path loss. The resulting SINR distribution is less fair, and cell-edge users are received with less strength than cell centre ones.

Performance changes arise in terms of cell SINR distribution when modifying both variables as shown in Lema et al. [LGRO11] and Vallejo-Mora [VM13]. A rise in P_0 improves the connection quality if the number of power limited users is small. A low value of P_0 provides the system with a reduced value of inter-cell interference; however, transmission power can result insufficient to overcome path loss and shadowing.

The performance of the OLPC depends very much on the network deployment environment. One synthetic scenario and one urban deployment are tested and compared in Lema et al. [LGRO11] The first presents high sensitiveness to interferences while the RL one is clearly sensitive to the availability of the transmission power, given the high number of sources of attenuation. The possibility of obtaining the best performance despite the environment nature is due to the versatility presented by the algorithm.

On the one hand, the OLPC formula shows a fairly good performance with low computational effort; but, ont the other hand it has not been optimised for the network throughput as it does not consider the inter-cell interference. Work in Peng and Kürner [PK13a] presents an iterative algorithm that calculates each UE power that maximises data rate, which is estimated using the truncated Shannon formula. On each iteration, the transmission power is updated to optimise the network performance based on the last iteration resulting users transmission power. Results show a fast convergence rate, throughput is improved by 21.9% compared to the LTE formula with only two iterations.

An automatic tuning algorithm for the parameters of the UL power control in LTE is proposed in Fernández Segovia [FS13]. The study is based on an initial sensitivity analysis that characterises the impact of P_0 and sub-carriers utilisation limit, both with an impact on interference levels. Next, the proposed algorithm is composed of two parts. Initially, it iteratively evaluates decreasing values of P_0 and sub-carrier utilisation until cell edge throughput requirements are met. In the second stage, optimisation is oriented to maximise capacity, and so, the average cell throughput. An extension of the basic algorithm towards non-ideal scenarios requires certain simplifications, otherwise computational cost would be prohibitive. The authors propose building a simplified scenario from the irregular original one and two mechanisms are proposed and investigated. The underlying idea is to approximate a real scenario by many single and regular ones based on geographical relations with neighbour cells. This means that at the end of the optimisation several solutions are provided to each cell, as depicted by Figure 6.11. Four mechanisms are compared for the selection of a final solution. Results indicate a successful operation of the approach with similar network performance among the mechanisms and selection methods. Finally, even though the method is conceived for the planning stage, it shows a complexity cost which is one order of magnitude lower than an exhaustive simulation analysis.

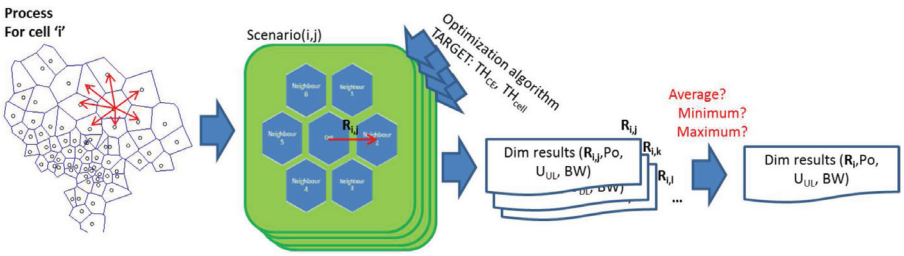


Figure 6.11 Statistical possible solution for cell *i*, from á [FS13].

6.2.4 Mobility Management

Handover is the mobility procedure that more directly affects the user experience, since it occurs during data transmission. More specifically, a sub-optimal setting of handover (HO) parameters could lead to a call drop or waste of network resources. Mobility robustness optimisation has gained attention in the research community, with several techniques for HO optimisation being proposed. Muñoz Luengo [MnL13] investigates the potential of adjusting handover margin (HOM) and time to trigger (TTT) to improve such robustness. An initial analysis performing a sweep of both parameters indicates that the network is more sensitive to variations of the HOM, with no additional benefits with the adjustment of the TTT. Hence, the authors propose the use of a fuzzy logic controller (FLC) that adaptively modifies just the HOM and being based on the values of HO ratio and call dropping ratio (CDR). Two fuzzy sets are suggested for the first input, on the other hand, three sets are defined in the second case. This allows having more granularity for the more critical performance indicator. The controller has been evaluated for different levels of traffic load and different user speeds. Results show an improvement in network performance, it is also shown that the network is more sensitive to margin variations for users with speed values above 10 km/h. Thus, the authors conclude suggesting the avoidance of changes in the margin greater than 1 dB.

6.3 Resource Management in Wireless Networks

6.3.1 Resource Management in Wireless Mesh Networks (WMNs)

A non-centralised, non-hierarchical, and distributed synchronisation process for WMN is introduced in [RPWD13, RPW14]. Each node has a time base that relative to an universal clock has an offset and a clock frequency difference. The purpose of the synchronisation is to correct the offset and the frequency so all the nodes will have almost the same time base. The synchronisation process is based on periodically transmission of messages with timing information by the all the nodes. The timing messages can be retransmitted to other nodes. Each node is correcting his time base according to the received messages. The synchronisation process has an acquisition phase when the clocks of the nodes have large variations and when the variations are reduced below a certain threshold the synchronisation process switches to the tracking phase. In the tracking phase, the variations among nodes are farther reduced and the changes in clock parameters are followed [RPW14].

Authors Ferreira and Correia [FC12] introduce a unified RRM strategy for multi-radio WMNs following self-organisation principles. This method built

on an abstraction-layer, radio-agnostic in the operation of multiple radios and transparent to higher layers. This methods integrates *channel assignment, rate adaptation (RA), power control*, and *flow-control, mechanisms,* achieving a high-performing WMN. This strategy is proposed for self-organised WMNs. The stratedy integrate multiple mechanisms:

- a RA mechanism aware of WMN traffic load specificities,
- an energy-efficient power control mechanism addressing the non - homogeneity of nodes rates,
- a load- and interference-aware channel assignment strategy that guarantees connectivity with any neighbour.

For a WMN of 13 multi-radio mesh access points (MAPs) using IEEE 802.11a, a fair aggregated throughput per MAP of 4.8 Mbps is guaranteed by the use of four channels and a total transmitted power of 34.5 dBm and RRM strategy exploits 100% of the network capacity, guaranteeing fairness and minimising the spectrum and power usage. Several system improvements multiple-input multiple-output (MIMO), higher modulations (available in 802.11n) and receiver sensitivities could be considered, resulting in higher system physical data-rates and larger ranges [FC12].

In Ferreira and Correia [FC13], a hierarchical-distributed strategy, combining RA, power control, and channel assignment mechanisms to efficiently guarantee max-min fair capacity to every node in WMN is introduced. In this model, a multi-radio WMN is composed of MAPs equipped with various radios, providing Internet connectivity to end-user terminals via a mesh point portal (MPP) gateway. Each MAPs aggregated traffic flows between an MPP gateway and the end-user(s) connected to that MAP, crossing one or multiple intermediary MAPs. The adopted fairness concept aims guarantee that all MAPs achieve a max-min fair aggregated throughput. If any MAP is favoured, increasing its load and associated throughput beyond maximal capacity that still guarantees fairness, there will be disfavoured MAP(s) that will have their throughput decreased. The RRM strategy is proposed too. This strategy include RA and transmission power control (TPC), a max-min fair flow-control mechanism and CA mechanism. Each multi-radio MAP has a radio agnostic abstraction-layer that implements the RRM strategy for optimisation of mesh radios resources, based on a mechanism for monitoring and sharing resources. It is hierarchical-distributed, triggered by each MPP, being sequentially run by each child. Paths are computed in advance by a routing algorithm. It follows max-min fairness principles in the allocation of capacity to MAPs. The combined RA and PC mechanism is proposed. It is sensitive to the WMN

fat-tree distribution of traffic flows. It uses the highest physical layer (PHY) bit rate possible only at MPPs (WMN bottleneck), and uses the remaining MAPs for minimum bit rates that satisfy their capacity needs. It is sensitive to the WMN fat-tree structure, using the highest possible bit-rates only at MPPs, and recurring, for the ramified links, to minimum bit rates that satisfy their capacity needs. This enables to reduce the transmitted power, an energy-efficient solution, and associated interference ranges, making channel reutilisation possible. A channel assignment mechanism is also proposed, aware of the load and interference of each MAP. It is concluded that with the proposed strategy max-min fairness (in the share of capacity) is guaranteed to every mesh node, interference is in existent, and resources such as power and spectrum are efficiently used [FC13].

In Cicalò and Tralli [CT13], a novel cross-layer optimisation framework to maximising the sum of the rates while minimising the distortion difference among multiple video users is introduced. The optimisation problem is vertically decomposed into two sub-problems and an efficient iterative local approximation (ILA) algorithm is proposed, which is based on the local approximation of the contour of the ergodic rate region in the OFDMA DL system. The ILA algorithm requires a limited information exchange between the application (APP) and the medium access control (MAC) layers, which independently run algorithms that handle parameters and constraints characteristic of a single layer. There is first formulated and discussed the feasibility of the optimisation problem showing that the optimal solution is achieved on the convex rate region boundary assuming subcarrier sharing. The problem has been then vertically decomposed into two sub-problems, each one characterised by parameters and optimisation constraints of a single layer. Numerical results have shown the fast convergence properties of the ILA algorithm and the significant video quality improvement of the proposed strategy with respect to optimisation strategies that only consider rate fairness.

6.3.2 Resource Management in Wireless Sensor Network (WSN)

In Sergiou and Vassiliou [SV12], the influence of source-based trees when they serve as topology control schemes in resource control congestion control algorithms and how they apply in a randomly deployed WSN are introduced and the simulation study is used to compare the performance of the two topology control schemes as well as the naive source-based tree when a resource control algorithm applies. Simulation results show that source-based trees could be more efficient when the data load is heavy, since they provide more routing paths from each node. On the other hand, sink-based trees provide

better results in terms of delay and energy consumption. The major conclusion that we extract from this study is that source-based trees, if carefully tuned, can provide an efficient topology control solution for specific applications [SV12].

Problems with topology design and control of the large-scale system such as WSN is solved in Porcius et al. [PFMJ12]. There is designed a novel procedure for topology design and control and implement it in the TopoSWiM simulation tool. The procedure combines a mathematical approach based on graph theory with a physical model of the operating environment and of the radio propagation channel. The mathematical approach is based on a new *clustering algorithm for gateway positioning (CAGP)*. The procedure combines a mathematical approach based on graph theory with a physical model of the operating environment and of radio propagation channel. The mathematical approach is based on a new CAGP, which determines the minimum number of gateways that are needed and their positions in order to provide coverage and external connections for the nodes with predefined positions while maximising network accessibility. The procedure takes into account also the physical model of the environment where the network is to be deployed and the appropriate radio propagation channel model. The newly proposed CAGP algorithm is first compared to the benchmark k-means and k-means++ algorithms, in terms of execution time and number of isolated nodes or accessibility.

New block mechanism (BACK) for the aggregation of several acknowledgment (ACK) responses into one special packet for WSNs with contention-based random medium access control MAC protocol with multiple nodes sharing the same medium is introduced in Barroca et al. [BVF$^+$11b]. The sensor block acknowledgement (SBAK-MAC) protocol that improves channel efficiency by aggregating several acknowledgement (ACK) responses into one special frame is introduced in Barroca et al. [BVF$^+$11b] and Barroca and Velez. [BV13]. SBACK-MAC protocol protocol enables reduce the end-to-end delay and energy consumption due to the protocol overhead in S-MAC.

The block acknowledgment (BACK) mechanism improves channel efficiency by aggregating several ACK control packet responses into one special packet, the Block ACK Response. Hence, an ACK control packet will not be received in response to every data packet sent. The Block ACK Response will be responsible to confirm the data packets successfully delivered to the destination. This packet has the same length as a data packet [BVF$^+$11b].

The BACK mechanism starts with the exchange of two special packets: RTS ADDBA Request and clear-to-send (CTS) ADDBA Response, as shown

in Figure 6.12(a). ADDBA stands for Add Block Acknowledgement. Then, the data packets are transmitted from the transmitter to the receiver (10 messages, each one fragmented into 10 small data packets). Afterwards, the transmitter will inquire the receiver about the total number of data packets that successfully reach the destination by using the Block ACK Request primitive. In response, the receiver will send a special data packet called Block ACK Response identifying the packets that require retransmission. The structure of the packet is presented in Figures 6.12(b) and (c). The BACK mechanism finishes with the exchange of two special control packets: the RTS DELBA Request and CTS DELBA Response. DELBA stands for Delete Block Acknowledgment. The structure of the DELBA packets is the same presented in Figure 6.12(a). Figure 6.13 presents the message sequence chart for the BACK mechanism.

Barroca et al. [BV13, BBV14b, BBVC14] describe the novel mechanisms to reduce overhead in IEEE 802.15.4 based on the presence of BACK Request (concatenation mechanism), while the second one considers the absence of BACK Request (piggyback mechanism). The aggregation of ACKs aims at reducing the overhead by transmitting less ACK control packets and by decreasing the time periods the transceivers should switch between different states. In algorithm, there is introduced the mechanism

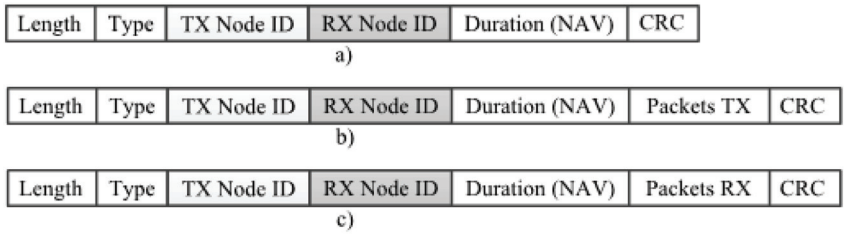


Figure 6.12 (a) RTS ADDBA Request and CTS ADDBA Response, (b) Block ACK Request, and (c) Block ACK Response packets format.

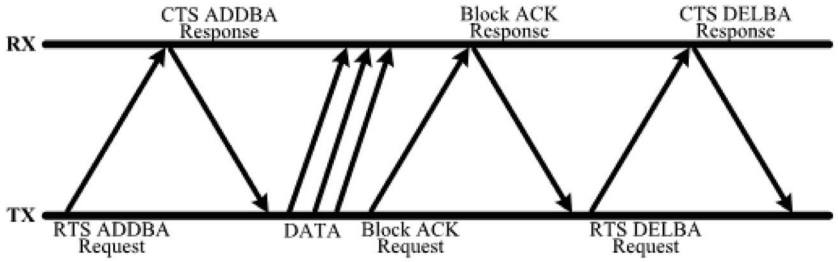


Figure 6.13 SBACK-MAC block ACK mechanism.

to increasing the throughput as well as decreasing the end-to-end delay, while providing a feedback mechanism for the receiver to inform the sender about how many transmitted (TX) packets were successfully received (RX). The mechanism also considers the use of the *request-to-send/clear-to-send (RTS/CTS)* mechanism, in order to avoid the hidden terminal problem.

A new mathematical model to characterise the interference in WSN based on the IEEE 802.15.4 standard is described in Martelli et al. [MBV12]. In this model, the consideration of the detailed operation of the carrier sense multiple access with collision avoidance (CSMA/CA) algorithm as described in IEEE 802.15.4 standard is integrated and included. There are assumptions that all nodes are distributed on a two-dimensional plane according to a homogeneous poisson point process (PPP), with a spatial density. They have to communicate with a central coordinator (CC), which is supposed to be located at the origin of the two-dimensional plane. In model a query-based traffic, the CC periodically sends a query to all nodes and nodes that receive it try to send their data to the CC, is considered. Authors mathematically derive the success probability for a generic node located at a distance r_{o} from the CC (reference transmitter TX_0) to correctly transmit its data and two factors can affect the success:

- connectivity, since only the nodes that receive the query from the CC can attempt their data transmission;
- interference, because all the nodes start the procedure to try to access the channel simultaneously.

Results collected during simulation show the impact of nodes density and sensing range on network performance evaluated in terms of success probability [MBV12].

Authors in Burati and Verdone [BV12] introduced a novel priority-based carrier-sense multiple access (P-CSMA) Protocol for Multi-Hop Linear Wireless Networks and it enables to increase throughput in comparison with traditional protocols like Slotted Aloha or 1- persistent Slotted CSMA. It partially enables solving the problem of interference among nodes simultaneously transmitting on the route. It's simple MAC protocol which exploits the linearity of the network topology. The protocol assigns to nodes different levels of priority, depending on their position in the route. Authors consider and describe the situations where all relays are (randomly) distributed over a straight line whose length (the source-destination distance) is fixed. The P-CSMA starts from the observation that, with traditional CSMA protocols, when a source nodes has many packets to send over an established route, they compete for accessing the radio channel with those that were previously

transmitted and are still being forwarded by some relays in the route [BV12]. All nodes in route have assign different level of priority in the access to the channel. It means that nodes closer to the destination have higher priority with respect to those closer to the source. Authors impose nodes to sense the channel for different intervals of time: the smaller is the sensing duration, the higher will be the priority in the access to the channel. Algorithm gives priority to those packets in the route which are closer to destination stands in the fact that they will not compete with those transmitted in the rear. In this way, P-CSMA mechanism speeds up the transmission of packets which are already in the route, making the transmission flow more efficient. The performance analysis in the sense of success probability and throughput is introduced too and from the results it is clear that the performance in terms of success probability and throughput are better than the ones achieved in the case of 1-persistent CSMA and Slotted ALOHA. Moreover, even under a distributed approach, the protocol proposed tend to behave similarly to a TDMA scheme were all transmissions are scheduled by a centralised controller. In Fabian and Kulakowski [FK13], the performance of geo-routing is validated in radio conditions that are RL for sensor networks is refereed. It is shown that the protocols that are known to guarantee the packet delivery for unit disk graph model and perfect location knowledge are far from that in real sensor networks. A new modification to the angular relaying (AR) was proposed and tested. It was shown that with two new types of routing messages, the packet delivery ratio can be significantly improved in unreliable networks. At the same time the algorithm modification did not increase its mean algorithm cost neither the message complexity, which directly influence the battery usage of nodes. Even though the modified algorithm does not guarantee delivery in the conditions mentioned, it is a first step in that direction, showing others a way to propose further improvements. A routing algorithm, composed of two algorithms: *greedy distance-based algorithm* and *AR,* was implemented in a new Java-based computer simulator. In the greedy mode the algorithm utilise CTS and RTS messages with a timer adopted from the AR algorithm. When a message arrives to a concave node AR is launched. The algorithm switches from the greedy mode into the AR mode when after sending an RTS message, no CTS message is received. In the AR mode, the data message contains information about the concave node and a flag indicating the mode is set to AR. Every node that receives the data message with the AR flag set, checks whether it is located closer to the destination than the concave node, in which case the algorithm switches back to the greedy mode.

6.3.3 Mobility-based Authentication in Wireless Ad hoc Network

A novel approach for authentication of network nodes based on their mobility patterns observed from the radio-wave propagation is introduced in Skoblikov [Sko13]. It relies on the fact that propagation of radio waves is immutable by an adversary. During the authentication procedure the of nodes location in space as well as the mobility pattern are verified in the context of current communication situation using a number of plausibility tests. For the new mobility-based authentication scheme, this information is the knowledge about the current location and mobility pattern of communicating party. For the new mobility-based authentication scheme this information is the knowledge about the current location and mobility pattern of communicating party. If in the given example scenario *Bob* utilises a two-step authentication protocol, he will first verify that *Alice* is approaching the junction from southern direction. Only after this is proven, the password would be requested from *Alice* in order to distinct it from the other cars driving on the same road. Significant advantage of the mobility-based authentication scheme is its flexibility. It can, but does not have to be combined with some higher-layer authentication protocols. Furthermore, the number of plausibility tests and their strictness can be adjusted depending on many parameters, such as desired level of security, accuracy of the wireless channel parameter estimation (WPCE), computation capabilities of Alice and Bob. The core of the mobility-based authentication scheme are plausibility tests. But since these heavily depend on the PHY-layer parameters of the communication system, the next section address the key features of the IEEE 802.11p and the classification of the wireless channel in car-to-car scenario.

6.3.4 Hybrid Mobile *ad hoc* Network (MANET)–Delay Tolerant Network (DTN) Networks and Security Issues

In MANET, the security mechanisms are based on the assumption that there is/are a connection between source and target nodes (end-to-end connections). opportunistic networks (OppNet) and DTN, disconnected MANET are more general than MANETs, because dissemination communication is the rule rather than conversational communication [PDC13]. Security is an important issue in disconnected MANET, DTN, and OppNet. OppNets are formed by individual nodes that can be disconnected for some time intervals, and that opportunistically exploit any contact with other nodes to forward messages. Each nodes computes the best paths based on its knowledge. The messages are routed and transmitted by *store-carry-forward* model. The main philosophy

of OppNet is to provide ability to exchange messages between source and destination nodes.

In OppNet, the security solutions need to reflect security for all nodes, all services and application that participate on routing and transmitting process. There is sporadic connectivity of nodes and we need to provide secure delivery of the messages from source node to destination node. In order to provide the effective communication between nodes, there is necessary to consider different aspects (disconnections, mobility, partitions, and norms instead of the exceptions [PDC12]).

The modification of the the reactive routing protocol *dynamic source routing protocol (DSR)* designed for MANET, which enables transportation of the packets between disconnected terminals is introduced in Papaj et al. [PDC13]. The main idea of the routing algorithms is to enable the routing mechanism not only if we have connected paths to the mobile nodes but when unable to find routes, if there are disconnected routes. If there are paths between the source and destination nodes, the standard routing algorithm DSR is use for selection of the paths. In the case that the DSR routing protocol cannot find the paths to the destination or when the disconnection of the paths are occurring then the new protocol is activated. Algorithm also uses the statistical information about all connections between its neighbours nodes and also provides useful information about how many times was the mobile node in contact with other mobile nodes. Based on these data the routing protocol can select the optimal forwarding mobile node. The modified algorithm has been implemented and tested in OPNET modeler.

In Matis et al. [MD14], the hybrid routing protocol is designed and described. This new routing protocol expands functionality of DSR routing protocol which implemented features of the DTN forwarding mechanisms, which enables delivering of messages in MANET networks also in the case when the height speed mobility of nodes causes that MANET network is fragmented to networks islands with zero connectivity between them. Because of the mobility a lot of new opportunistic connections between nodes are created. The proposed algorithm also enables to increase the performance of the network in the sense of the message delivery in disconnected environment [MD14]. The performance analysis of the proposed routing protocols was tested in MAT-LAB.

Hybrid MANET–DTN networks are new type of mobile networks, that provides the new way how the different application could be provided for end users. The main idea of hybrid MANET–DTN networks enable use the network not only for personal usage but for emerging applications and

services. The trust-based candidate node selection algorithm for hybrid MANET–DTN is introduced [PDP14]. This mechanism enables the selection of the secure mobile node used for transportation of the data across the disconnected environment. Selected candidate nodes provide the secure mobile node selected to transport of the data between disconnected island of the MTs [PDP14]. The trust values are computed from collected routing and data parameters. Each mobile node have its own values about neighbours mobile nodes and then the routing protocol can select optimal node for transportation of the data. The proposed algorithm is implemented into OPNET modeler simulator tool.

6.3.5 Resource Managent in LTE Network

Authors in [Lue13] investigate the potential of adjusting HOM and time-to-trigger (TTT) for intra-frequency HO optimisation. There is described HO mechanisms and a sensitivity analysis of HOM and TTT is carried out for different system load levels and user speeds. Next the performance analysis considering both HO parameters for different situations, highlighting the variation of the user speed is discussed. There is also introduced the new incremental structure of the Fuzzy Logic Controller (FLC), that adaptively modifies HOMs for HO optimisation. The design of the FLC includes the following tasks:

- define the fuzzy sets, and
- membership functions of the input signals and define the linguistic rule base which determines the behaviour of the FLC.

The first step involved in a FLC execution is the fuzzifier, which transforms the continuous inputs into fuzzy sets. Each fuzzy set has a linguistic term associated, such as high or low. In particular, two fuzzy sets have been defined for the input HO Ratio (HOR): high and low. In the case of the CDR, since it is a more critical performance indicator, three fuzzy sets have been defined to have more granularity: high, medium, and low. Note that the number of input membership functions has been selected large enough to classify performance indicators as precisely as an experienced operator would do, while keeping the number of input states small to reduce the set of control rules. The translation between numerical values and fuzzy values is performed by using the so-called membership functions. Its main advantage is to allow addressing numerical problems from the human reasoning perspective, making the translation of the network operator experience to the system control easier.

6.4 Energy Efficiency Enhancements

6.4.1 Energy Efficiency in Cellular Networks

In Europe, the telecommunication market accounts for 8% of the total energy consumption and for 4% of the CO_2 emission. The problem of increasing energy consumption and CO_2 emissions in different industrial sectors led the European Commission to identify the so-called 2020 objectives which foresee to reach the 20% of renewable energy production, to improve energy efficiency by 20% and to decrease the CO_2 emissions by by 20% by the end of the year 2020 [BCE+11]. In both fixed and mobile telecommunication sectors, the access network is responsible for a large part of the energy consumption. As an example, in a typical mobile radio network, up to about 80% of total power consumption occurs at the BS. It is worth noting that the power consumption in cellular networks is steadily increasing due to the growing demand of broadband wireless internet access through the usage of new MTs such as smartphones, tablets, and other high-end terminal devices, as well as laptops with cellular connectivity.

For future networks, an energy efficiency improvement is expected. Litjens et al. [LTZB13] assess the energy efficiency of mobile networks in 2020 and compares it to a 2010 baseline by taking into account the trends of mobile traffic increase, the corresponding network deployments (including the BS density and use of small cells), and technological improvements w.r.t. mobile network equipment (reflected in, e.g., power consumption models and sleep modes). An energy efficiency improvement factor of 793 has been observed in 2020 over 2010, thanks to the traffic increase (leading to more bits transmitted per cell), hardware evolutions (lower power consumption of BSs and backhauling), network sharing (leading to higher resource utilisation especially in coverage-limited cases), micro-sleep mode of macrocells in 2020 dense urban scenario (energy saving at low traffic), and MIMO.

6.4.1.1 Power consumption of different telecommunication technologies

To determine the power consumption of the cellular network, it is important to be aware of the power consumed by the BSs active in the network. A power consumption model for a macrocell base station (MBS) is proposed by Baumgarten et al. [BJRK12]. A BS consists of different hardware elements, as shown in Figure 6.14, which can be grouped into two distinct categories: sector-specific hardware, which is exclusively used to operate one sector (blue box in Figure 6.14(a)) and shared hardware, which is either baseband

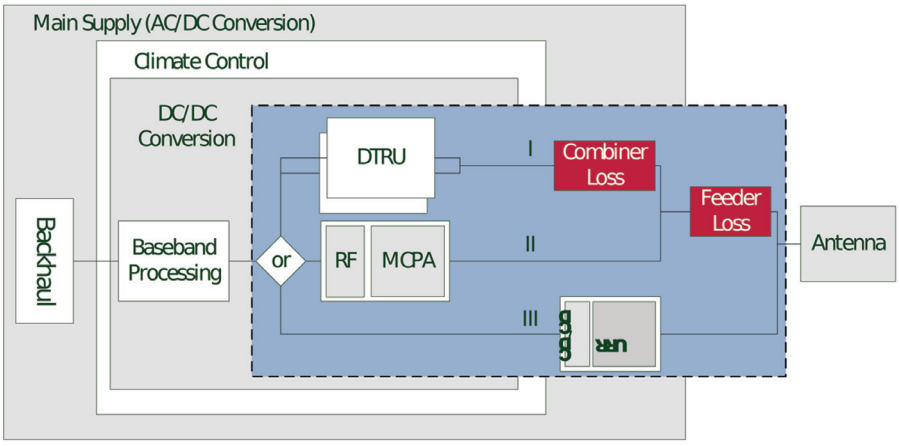


Figure 6.14 BS architecture as proposed by [BJRK12] (a) and [DJL+13] (b).

processing or site support. Based on this architecture, it is possible to set up a power consumption model for the BS. The power consumed by a BS consists of a part that is always used, even when the BS does not process any signals (idle state). The other part scales linearly with the load. Figure 6.15 shows the power consumed by a global system for mobile communications (GSMs) BS under different loads. A break-up between the power consumption of the different parts of the BS is also provided.

In Deruyck et al. [DJL+13], a BS similar power consumption model for a macrocell and femtocell BS is used for comparing the energy efficiency

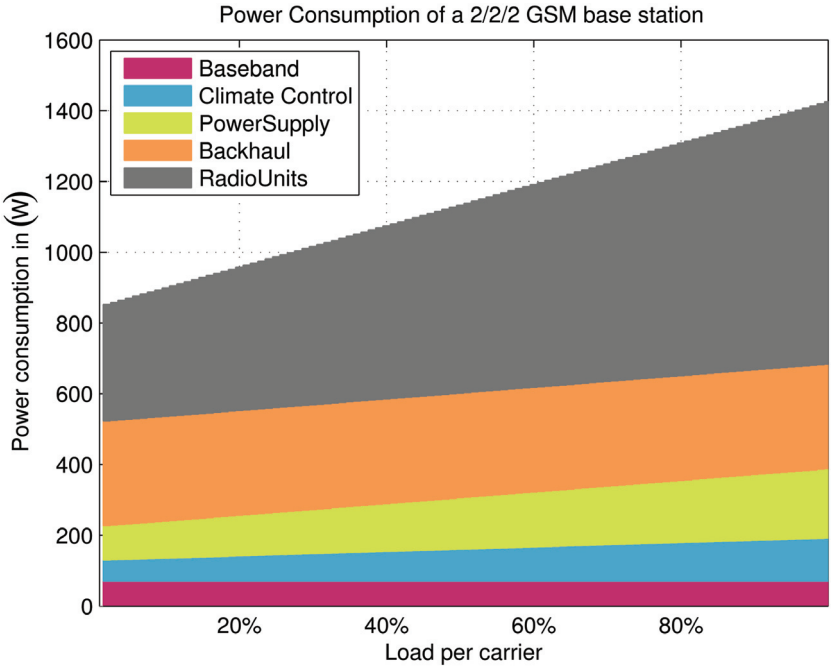


Figure 6.15 Power consumption of a GSM BS at different load levels [BJRK12].

between an LTE and LTE-A BS. The influence of three main functionalities added to LTE-A on the energy efficiency is investigated: carrier aggregation (to increase the bit rate), HetNets (whereby macro-cell and femtocell BS are mixed in one network), and extended support for MIMO (where multiple antennas are used for sending and receiving the signal). An appropriate energy efficiency metric is proposed, which takes into account the geometrical coverage, the number of served users, the capacity, and of course the power consumed by the network. In general, a higher bit rate results in a lower energy efficiency. However, due to carrier aggregation, LTE-A allows to obtain higher bit rates for even a higher energy efficiency. Heterogeneous LTE(-A) networks typically consist of macrocell and femtocell BS. For bit rates above 20 Mbps, the MBS is the most energy efficient. Below 20 Mbps, there is no unambiguous answer as it depends on the bit rate which technology is the most energy-efficient. For future networks, it is recommended to estimate accurately the required bit rate and coverage to decide which BS type or combination of these types should be used in the network. Finally, MIMO can also increase the energy efficiency, even up to 5 times by using spatial diversity and 8×8

MIMO. Future networks should thus support MIMO. It is recommended that future networks implement LTE-A as it will improve the energy efficiency compared to LTE.

Another way to improve the amount of radiated power (and thus the power consumed by the BS) is beamforming. Correia Gonçalves and Correia [GC11] developed two simulators, one for universal mobile telecommunications system (UMTS) and another for LTE, to evaluate in a statistical way the potential impact that adaptive antenna arrays have to reduce the radiated power, compared with actual BS static sector antennas. A logarithmic model was derived to represent the power improvement that is achieved by adaptive antennas in UMTS as a function of the number of users and the number of elements at the antenna array. When cells are near its top capacity, which was simulated with 70 users within the same cell, 90% of the power that is radiated by static sector antennas can be saved if adaptive antennas are used and, for antenna arrays with eight elements, more than 93% of radiated power saving. LTE power improvement, when compared with actual static sector antennas, does not change with the variation on the cell radius, for the same number of interferers in neighbour cells.

6.4.1.2 Cell switch off

A lot of power is consumed when it is actually not necessary. For example, during the night, all BS of the network are still active although there might be no activity taking place into their cells. Letting those BS sleep and turn them on when there are really necessary can save a significant amount of power. In literature, several cell switch off (CSO) schemes have been proposed [GYGR14]. Most of them address the problem of selectively switching off BS by means of heuristic algorithms. These schemes are preferred because CSO is a combinatorial NP-Complete problem, and hence, finding optimal solutions is not possible in polynomial time. However, González et al. [GYGR14] propose a multi-objective framework that takes the traffic behaviour into account in the optimal cell switch on/off decision making which is entangled with the corresponding resource allocation task. The exploitation of this statistical information in a number of ways, including through the introduction of a weighted network capacity metric that prioritises cells which are expected to have traffic concentration, results in on/off decisions that achieve substantial energy savings, especially in dense deployments where traffic is highly unbalanced, without compromising the QoS. The proposed framework distinguishes itself from the CSO papers in the literature in two ways: (i) The number of cell switch on/off transitions as well as handoffs are

minimised. (ii) The computationally-heavy part of the algorithm is executed offline, which makes the real-time implementation feasible. The proposed multiobjective framework achieves significant gains with respect to the considered benchmarks in terms of power consumption (>50%), number of HOs (>85%), and the cell switch on/off transitions (>87%).

Litjens et al. [LTZB13] observe that dynamic switching off picocell BS at low traffic for energy saving has limited impact on the overall energy efficiency due to the low power consumption and low total number of picocell BS. However, a higher potential saving is expected for scenarios with more deployed picocell BS to serve hot zone traffic.

6.4.1.3 Cellular deployment strategy

Barbiroli et al. [BCE+11] show that the power consumption of a mobile radio network is strictly related to the cellular deployment strategy. For a Manhattan environment, whereby the city is composed of a regular grid of square building blocks separated by wide streets, a macrocellular and microcellular coverage architecture has been considered and results show that MBSs are power efficient for outdoor coverage and indoor, high floor locations, but become extremely inefficient when a coverage in indoor locations at the first floor is needed. Coversely microcellular layout is efficient for outdoor coverage and indoor location at first floor, but inefficient for coverage at high floors. Thus the maximum efficiency (i.e., the minimum transmitted power density per km^2) is achieved by combining the micro- and macrocellular architecture: using MBSs for indoor coverage at high floors and for outdoor coverage, acting as umbrella cells, and using microcell BS as gap fillers for outdoor coverage and for indoor coverage at lower floors.

The idea that the power consumption of a mobile radio network is related to the cellular deployment strategy is also considered by Deruyck et al. [DTJM12] in which the authors propose a coverage-based deployment tool. The goal of this tool is to design a network that covers a predefined geometrical area and bit rate with a minimal power consumption. Possible BS locations are provided and the algorithm selects those locations that result in the highest coverage with the lowest power consumption. 3D data about the environment is taken into account when predicting the coverage of the network. For a suburban area in Ghent, Belgium, an wireless fidelity (WiFi) 802.11n network and an LTE-A femtocell network are developed. The LTE-A femtocell network is about 2.5 times more energy-efficient than the WiFi network due to the fact that the LTE-A network uses about half of the BS than the WiFi network.

6.4.1.4 Joint optimisation of power consumption and human exposure

When we talk about green wireless networks, we should not limit ourselves to energy efficiency; also exposure for human beings is an important issue. People are becoming more concerned about possible health effects. In a Belgian (Flemish) study, about 5% of the respondents does not own a mobile phone. 6% of non-owners say that they are afraid of the health effects [MSM12]. Electromagnetic field exposure awareness has significantly increased in the last few years. International organisations such as international commission on non-ionizing radiation protection (ICNIRP) provide safety guidelines and national authorities define laws and norms to limit exposure of the electromagnetic fields caused by wireless networks. In Deruyck et al. [DJT$^+$13], a capacity-based deployment tool is proposed which optimises the network towards power consumption or towards global exposure. The tool is capacity-based which means that it will respond to the instantaneous bit rate request of the users active in a considered area [DJTM14]. Appropriate distribution models for the number of active users, their locations, and their required bit rate had to be chosen. Preliminary results showed that when optimising towards power consumption, a network with a low number of BS with a high output power are obtained (leading to a low power consumption but high global exposure), while when optimising towards global exposure, the opposite situation is obtained: a high number of BS with a low output power (resulting in a high-power consumption but a low global exposure). As a compromise, one can optimise towards power consumption while satisfying a predefined exposure limit. This network shows a good compromise for the power consumption and the global exposure compared to the other two optimisations. However, further optimisation should be aspired. In Deruyck et al. [DTP$^+$15], the capacity-based deployment described above is extended to optimise the network towards both power consumption and exposure. By choosing an appropriate fitness function, the required trade-off between power consumption and global exposure is obtained. Furthermore, the requirement to connect as much as possible to an already active BS is dropped as this might result in a good optimisation towards power consumption but is not necessary the right choice when optimising towards global exposure. The network obtained with the proposed algorithm has a 3.4% lower power consumption and a 37% lower global exposure compared to the network optimised towards power consumption while satisfying a certain exposure limit. A better trade-off towards both parameters is thus obtained.

6.4.2 Energy Efficiency in WSNs

6.4.2.1 Energy harvesting

Power consumption is also a major bottleneck in WSN due to the difficulty to provide a continuous or sporadic energy source in situ for the operation of the wireless nodes. A natural component of any comprehensive solution to this problem is to leverage energy-harvesting technologies, whereby the needed energy is collected from the environment by converting different forms of energy, such as solar, elastic or radio frequency, into electrical power. Castiglione et al. [CSEM11] address the problem of energy allocation over source digitisation and communication for a single energy-harvesting sensor. Optimal policies that minimise the average distortion under constraints on the stability of the data queue connecting source and channel encoders are derived. It is shown that such policies perform independent resource optimisations for the source and channel encoders based on the available knowledge of the observation and the channel states and statistics. The main drawback of these policies is that they require large energy storage system (e.g., battery) to counteract the variability of the harvesting process and a large data queue to mitigate temporal variations in source and channel qualities. Suboptimal policies that do not have such drawbacks are investigated as well, along with the optimal trade-off distortion versus backlog sise, which is addressed via dynamic programing tools. Using large deviation theory tools, Castiglione et al. [CSEM11] designed also policies that show a good trade-off between energy storage discharge probability, data buffer overflow probability and average distortion.

6.4.2.2 Power consumption in electronics circuitry

Another major issue to be addressed in WSNs is the power consumption of the electronic circuitry in transmitter and receiver and the related energy consumption. Dimic et al. [DZB11] address this problem by deriving an energy consumption model for low-power wireless transceivers as a function of transmitter power and packet length. Their results show that both variables can be set optimally, minimising energy per bit of data, for any channel attenuation and control data overhead. Up to 85% of energy can be saved compared to full power transmission of short packets. However, there are two limits to this approach: output power in existing transceivers is set with coarser resolution than necessary and typical packet length is shorter than optimal. Since optimal packet length increases with control data length, packet aggregation with

fragment repetition (AFR) is applied to reduce optimal packet length. This approach shows further improvement of energy efficiency.

6.4.2.3 Routing protocol

The energy per bit received in WSNs can also be improved by using the mobility-based routing protocol proposed by Ferro and Velez [FV12]. Such a protocol forwards messages from the nodes to the sink in an effective manner. It calculates for each path a cost based on the probability of errors during transmission, and selects the path with the lowest cost. Constant information exchange about paths and their cost allow to keep the path information up-to-date, and avoid the use of dead links generated by node's mobility. The performance is compared with flooding, which is an easy-to-implement mechanism for forwarding packets in multi-hop networks. A node who receives a packet transmits it to all the neighbours. By 'flooding', the network with multiple copies of the same packet, eventually it will get to the destination. Simulation results of random networks showed that the developed protocol can deliver up to 77% more packets than Flooding, even with a lower energy cost as shown in Figure 6.16(a). With the protocol, one requires 34% less energy per bit received as shown in Figure 6.16(b). By introducing sleep cycles in the protocol, this energy saving rises to 47%.

6.5 Spectrum Management and Cognitive Networks

6.5.1 Cognitive Radio

According to Haykin [Hay05], spectrum utilisation can be improved significantly by making it possible for a secondary user (SU; who is not being serviced) to access a spectrum hole unoccupied by the primary user (PU) at the right location and the time in question. CR is an intelligent wireless communication system that is aware of its surrounding environment, and uses the methodology of understanding-by-building to learn from the environment and adapt its internal states to statistical variations in the incoming RF stimuli by making corresponding changes in certain operating parameters (e.g., transmit-power, carrier-frequency, and modulation strategy) in real-time, with two primary objectives in mind: (i) highly reliable communications whenever and wherever needed, and (ii) efficient utilisation of the radio spectrum. Given the importance of CR for future telecommunication systems such as fourth generation (4G) and beyond mobile technologies, the number of research projects and initiatives in this field is increasingly growing. Within COST IC

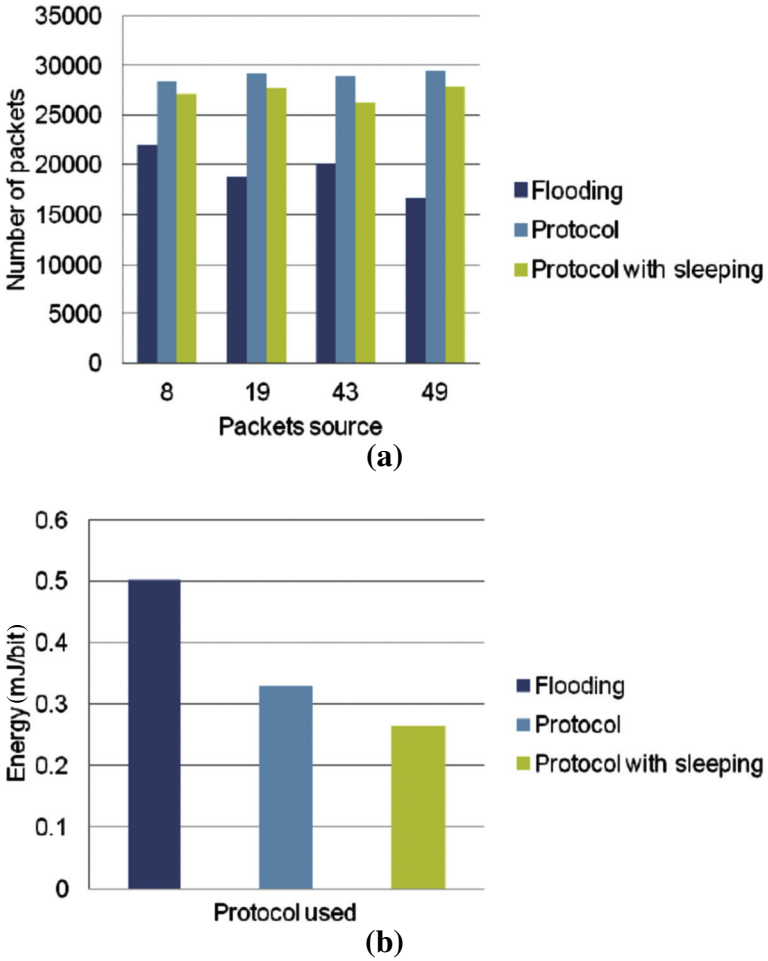


Figure 6.16 Influence of the routing protocol on the number of delivered packets (a) and the energy per received bit (b) [FV12].

1004, research has included aspects of how CR impacts on the performance of spectrum sharing and coexistence.

Cognitive radio terminals should be capable of acquiring spectrum usage information by means of some spectrum awareness techniques such as Spectrum Sensing (SS). Furthermore, in practical scenarios, the CR does not have *a priori* knowledge of the PU signal, channel and the noise variance. In this context, investigating blind SS techniques is an open research issue. To this end, Chatzinotas et al. [SKSO13] have studied several eigenvalue-based blind

SS techniques such as scaled largest value (SLE), standard condition number (SCN), John's detection and spherical test (ST)-based detection. The decision statistics of these techniques have been calculated based on the eigenvalue properties of the received signal's covariance matrix using random matrix theory (RMT). Several methods are considered in [CSO13a]. The sensing performances of these techniques have been compared in terms of probability of correct decision in Rayleigh and Rician fading channels for the presence of a single PU and multiple PUs scenarios. It has been noted that the SLE detector achieves the highest sensing performance for a range of scenarios.

The SS only approach ignores the interference tolerance capability of the PUs, whereas the possibility of having secondary transmission with full power is neglected in an underlay-based approach. To overcome these drawbacks, Chatzinotas et al. [SKSO14b] have proposed a hybrid cognitive transceiver which combines the SS approach with the power control-based underlay approach considering periodic sensing and simultaneous sensing/transmission schemes. In the proposed approach, the SU firstly estimates the PU SNR and then makes the sensing decision based on the estimated SNR. Subsequently, under the noise only hypothesis, the SU transmits with the full power and under the signal plus noise hypothesis, the SU transmits with the controlled power which is calculated based on the estimated PU SNR and the interference constraint of the PU [CSO13b]. Furthermore, sensing-throughput trade-off for the proposed hybrid approach has been investigated and the performance is compared with the conventional SS only approaches in terms of the achievable throughput.

In CR networks, the detection of active PUs with a single sensor is challenging due to several practical issues such as the hidden node problem, path loss, shadowing, multi-path fading, and receiver noise/interference uncertainty. In this context, cooperative SS has been considered a promising solution in order to enhance the overall spectral efficiency. Existing cooperative SS methods mostly focus on homogeneous cooperating nodes considering identical node capabilities, equal number of antennas, equal sampling rate and identical received SNR. However, in practice, nodes with different capabilities can be deployed at different stages and are very much likely to be heterogeneous in terms of the aforementioned features. In this context, Chatzinotas et al. [SKSO14a] study the performance of a decision statistics-based centralised cooperative SS using the joint probability distribution function (PDF) of the multiple decision statistics resulting from different processing capabilities at the sensor nodes. Further, a design guideline has been suggested for the network operators by investigating performance versus network sise trade-off while deploying a new set of upgraded sensors [SCO15].

Cooperative SS is a promising technique in CR networks by exploiting multi-user diversity to mitigate channel fading. Cooperative sensing is traditionally employed to improve the sensing accuracy, as shown in Figure 6.17 while the sensing efficiency has been largely ignored. However, both sensing accuracy and efficiency have very significant impacts on the overall system performance. In Zhang [Zha13a], the author first identifies the fundamental trade-off between sensing accuracy and efficiency in SS in CR networks. Then, several different cooperation mechanisms are presented, including sequential, full-parallel, semi-parallel, synchronous, and asynchronous cooperative sensing schemes. The proposed cooperation mechanisms and the sensing accuracy-efficiency trade-off in these schemes are elaborated and analysed with respect to a new performance metric achievable throughput, which simultaneously considers both transmission gain and sensing overhead. Illustrative results indicate that parallel and asynchronous cooperation strategies are able to achieve much higher performance, compared to existing and traditional cooperative SS in CR networks.

An opportunistic spectrum usage model facilitated with periodic SS and handoff in a two-hop selective relay network is considered in Wang [Wan13].

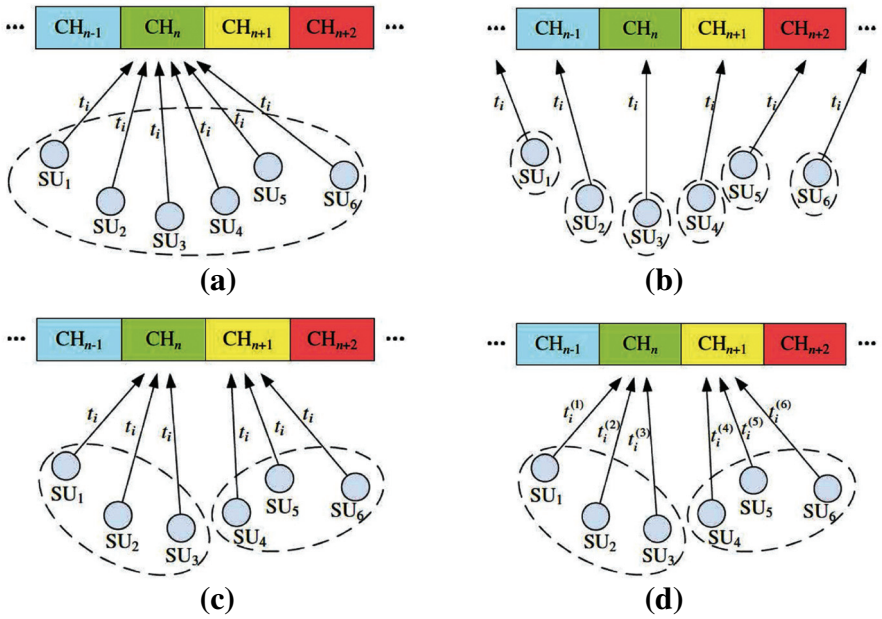


Figure 6.17 Illustration of cooperation mechanisms: (a) sequential cooperative sensing; (b) full-parallel cooperative sensing; (c) semi-parallel cooperative sensing; and (d) asynchronous cooperative sensing.

A novel sensing and fusion algorithm is proposed under a signalling bandwidth-constrained condition by introducing a quantised soft sensing and a test statistic restoration process. The reliability of secondary transmissions is studied, where expressions for the probability of collision, and the throughput of SU are derived.

Simulation results show the proposed sensing algorithm provides a better sensing performance with a lower signalling cost and a lower collision probability, compared to the conventional algorithm in selective relay networks in light of the missed detection probability and throughput. Finally, results show that the proposed algorithm offers a close performance to the throughput of a full soft algorithm in selective relay networks.

Wireless access networks based on dynamic spectrum access (DSA) in a densely deployed environment are promising solution to meet the future mobile traffic challenge with CRS capabilities. In such a network, not only the capacity improvement, but also energy efficiency (EE) are critical problems. It reported that in a multi-cell scenario the interference from neighbouring cells will degrade both EE and spectrum efficiency (SE). The challenges and solutions to achieve an energy efficient wireless communications are summarised in Li et al. [LXX+11], and Hasan et al. [HBB11]. As illustrated in Figure 6.18, in Tao Chen et al. [TCZ12] the energy efficient resource allocation in a DSA-based wireless access network is explored. In such a network APS are densely deployed to provide open access to MTs. Spectrum of the network is opportunistically shared from PUs. In multiple channels divided from available spectrum, an access point (AP) selects one working channel and all MTs attached to the AP use the same channel to communicate with the AP. The work in Chen et al. [CZHK09] is extended by taking into account EE in the channel allocation and MT association. The main contributions from Tao Chen et al. [TCZ12] include: designing a local information exchange method to estimate inter-cell interference, proposing a local energy efficient metric used for distributed energy efficient algorithms, and developing the distributed energy efficient algorithms for AP channel selection and MT association, respectively. The original problem is divided into two sub-problems, i.e., the channel selection problem of AP and the AP association problem of MT. The distributed AP channel selection algorithm and MT association algorithm are developed, which allow an AP to explore spectrum with less interference, and an MT to connect to an AP with the most bit/energy gain. The approach is flexible and provides scalability to large networks.

Smart grid is widely considered to be the next generation of power grid, where power generation, management, transmission, distribution, and

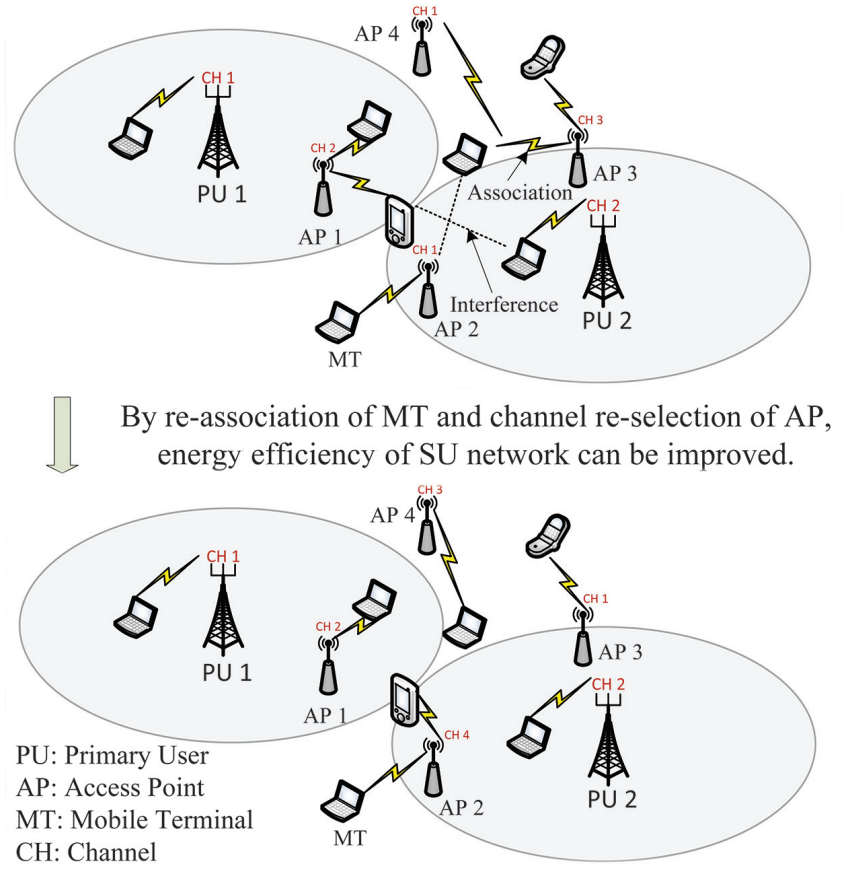


Figure 6.18 Reconfiguration of network by distributed MT association and AP channel selection algorithms to reduce interference and save network energy consumption.

utilisation are fully upgraded to improve agility, reliability, efficiency, security, economy and environmental friendliness. Demand Response Management (DRM) is recognised as a control unit of smart grid, with the attempt to balance the real-time load as well as to shift the peak-hour load. As shown in Figure 6.19, input of the system is provided by power plants, while feedback is demand of end users measured by smart metre. DRM acts as a control unit to balance and shape the realtime load. Output of the system is electricity delivered to each user through transmission and distribution. The forward path is *power flow*, and the bidirectional path is *information flow*, which provides two-way communications in smart grid.

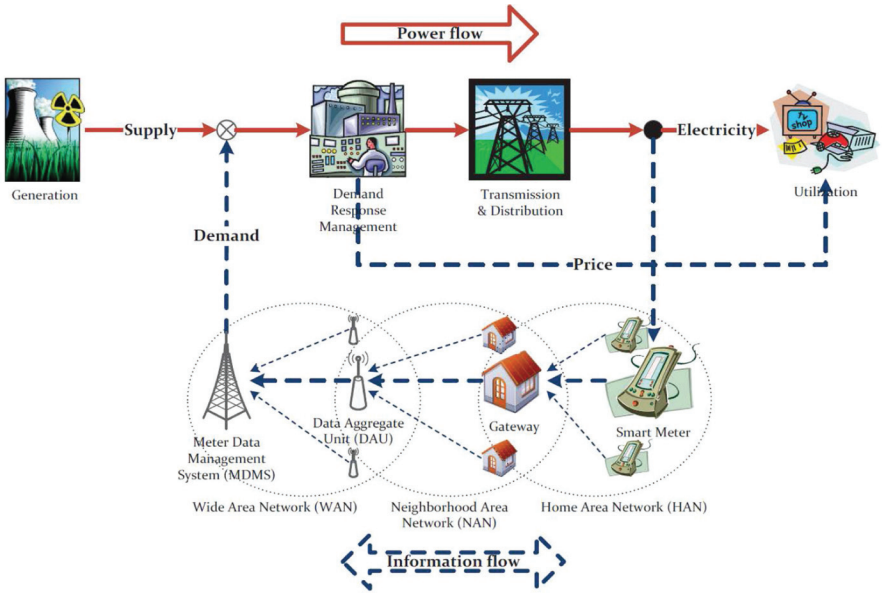


Figure 6.19 A closed-loop system scheme of smart grid.

Zhang [Zha13b] have introduced CR into smart grid to improve the communication quality. By means of SS and channel switching, smart metres can decide to transmit data on either an original unlicensed channel or an additional licensed channel, so as to reduce the communication outage. The impact of the communication outage on the control performance of DRM is also analysed, which reduces the profit of power provider and the social welfare of smart grid, although it may not always decrease the profit of power consumer. It is shown that the communication outage can be reduced. How the communication quality affects the control performance of DRM is also analysed. The work from Zhang [Zha13b] provides the guidelines of achieving better control performance with lower communication cost, paving the way towards green smart grid.

In the context of CR, spectrum occupancy prediction has been proposed as a means of reducing the sensing time and energy consumption by skipping the sensing duty for channels that are predicted to be occupied in future time instants [CAV13]. In a CR network, channel occupancy is not directly observable by the sensing nodes due to the wireless channel between the PU and SU. Therefore, SS can be modelled as a Hidden Markov Model (HMM) with a hidden process X_t and an observable process Y_t.

The hidden process *Xt* represents the PU activity and is modelled as a two-state Markov chain with state space $X = \{0, 1\}$, where '0' and '1' indicate an active and an idle PU, respectively. Similarly, Y_t is a random process with state space $Y = \{0, 1\}$, which represents the SS output, with '0' and '1' corresponding to an unoccupied and an occupied channel, respectively. Given the PU activity status x_t at time instant t, and y_t the corresponding sensing output, SS can be described by a two-state HMM with parametres $\lambda = (\pi, \mathbf{A}, \mathbf{B})$, where π is the initial state distribution: $\pi = [\pi_i]_{1\times 2}$, $\pi_i = P(x_1 = i)$, $i \in X$; $\mathbf{A}$ is the transition matrix that describes the probabilities of the PU activity status to change from active to idle: $\mathbf{A} = [a_{ij}]_{2\times 2}$, $a_{ij} = P(x_{t+1} = j | x_t = 1)$, $i, j \in Y$; and $\mathbf{B}$ is the emission matrix that describes the relationship between sensing output and the actual PU state: $\mathbf{B} = [a_{jk}]_{2\times 2}$, $b_{jk} = P(x_t = j | y_t = k)$, $j \in X, j \in Y$.

Having an observation sequence of past SS outputs, O, of length T, the first step to HMM-based spectrum occupancy prediction is the model training process. During this process the HMM-based spectrum occupancy predictor estimates its parameters, using the Baum–Welch algorithm, so that the probability of observing O is maximised. After training the model's parameters, the Viterbi algorithm is used to determine the channel occupancy sequence X_t that is most likely to have generated the observation sequence, O. Given the estimated model parameters λ^* and the decoded channel occupancy states, the channel occupancy at a future time instant, $T + d$, is estimated by comparing the conditional probabilities of an observation sequence, O_T, to be followed by an unoccupied, $P(O_{T+d}, 0|\lambda^*)$, and an occupied channel, $P(O_{T+d}, 1|\lambda^*)$, respectively.

A performance comparison between the HMM-based spectrum occupancy predictor and a 1st and 2nd order Markov predictors is shown in Figure 6.20. The prediction performance is evaluated in terms of the probabilities of true positive predictions (upper end of the figure) and false negative prediction (lower end of the figure) and the channel occupancy status transition rate. With reference to Figure 6.20 it is shown that the HMM-based predictor has a higher adaptability to the channel status transitions which in turn results in an up to 50% higher prediction performance.

In the underlay CR scenario, Sharma et al. [SCO13b] has studied Interference Alignment (IA) as an interference mitigation tool in order to allow the spectral coexistence of two satellite systems. Furthermore, frequency packing (FP) can be considered as an important technique for enhancing the spectrum efficiency in spectrum-limited satellite applications. In this context, this contribution focuses on examining the effect of FP on the performance of multi-carrier-based IA technique considering the spectral coexistence scenario

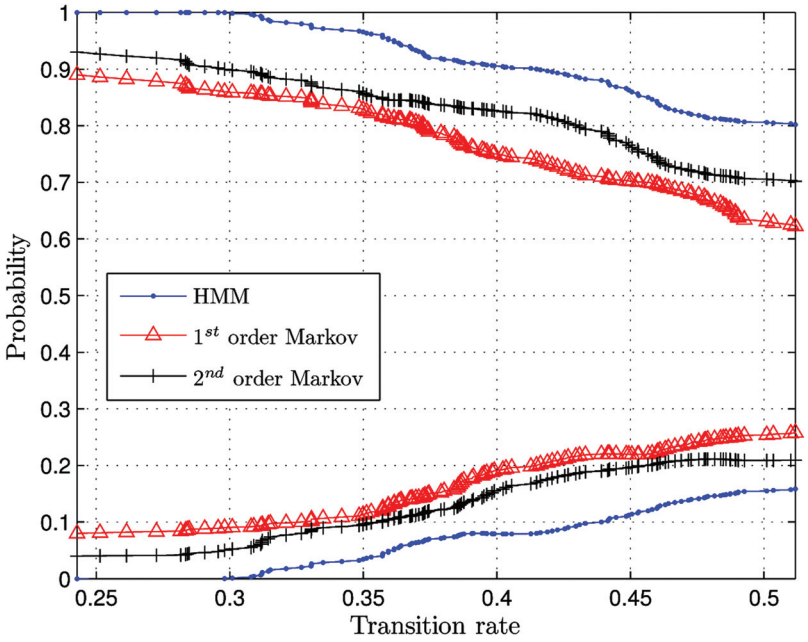


Figure 6.20 Prediction performance versus channel status transition rate.

of a multibeam satellite and a monobeam satellite with the monobeam satellite as primary and the multi-beam satellite as secondary. The effect of FP on the performance of three different IA techniques in the considered scenario has been evaluated in terms of system sum rate and primary rate protection ratio (PR). It has been shown that the system sum rate increases with the FP factor for all the techniques and the primary rate is perfectly protected with the coordinated IA technique even with dense FP [SCO14].

Author Diez [Die13] has studied different detection and hypothesis testing techniques that considerable influence on the performance of a cooperative CR network. Indoor trials with DVB T signals have been carried out at 690 MHz, using USRP-based devices. Measurement results proved that the wave-based form detector was more accurate than the cyclostationary feature detector. Besides, RF receivers with AGC and omni-directional antennas increased the signal detection probability. Taking into account the obtained results, a free channel detection algorithm has been defined based on installing some test points with a determined separation.

The cognitive device in X will be able to transmit a certain signal power with an omnidirectional characteristic if in the A, B, C, and D points the signal

transmitted by the cognitive device is not detected. If the signal is detected in any of the points of a particular arm of the cross, the radiation pattern of the antenna situated in X could be modified. Installing all of these fixed points could lead to a dense cognitive device network. This disadvantage can be offset by creating real time data bases with the information given by all the mobile cognitive devices on the area.

With the profusion of low cost wireless platforms, such as the Ettus family of Software Defined Radios (SDRs), experimentation in wireless research became a reality. Such platforms have allowed many ideas to become a reality, providing researchers a never before possible way to demonstrate the feasibility of their concepts. Nevertheless, such platforms are usually limited to a few at a time, since they become very hard to handle in larger quantities without a proper underlying infrastructure. This fact, unfortunately, limits the kind and scale of techniques that can be implemented and demonstrated. In Cardoso et al. [LSCG12], authors have introduced the CorteXlab testbed, a large scale testbed comprised of heterogeneous platform nodes able to deal with distributed PHY design and CR.

CorteXlab will provide researchers all over the world with an up-to-date high quality PHY layer testing tool. Being part of the Future Internet of Things (FITs) project, CorteXlab will be accessible through an easy to use web portal, with possible future connection to the other testbeds integrating FIT. Advanced PHY layer reconfigurability also allow to address issues common to this kind of networks, such as synchronisation and power consumption. Note that the optimal way to address this kind of scenarios remains unknown in the academia, and is in the forefront of the research efforts. Ongoing work includes the development of a reference PHY design, based on PHY layer relay network techniques, the deployment of the middleware (heavily based on SensLab) that allows the management and remote activation of the nodes and the choice of the node platform itself.

6.5.2 Digital Terrestrial Television (DTT) Coexistence & TV White Spaces (TVWS)

With the arrival of digital television technologies and new compression systems, the concept of TVWS was introduced to define such regions of space–time–frequency in which a particular secondary use is possible [TMS09]. This is, such parts of the licensed spectrum (mainly for broadcasting service) that are not in use in a specific area and that can be used for secondary devices (known as TVWS devices) to provide other communications services in such region.

These new TVWS devices, also named white space devices (WSD) and considered as cognitive devices [MM99], should determine which frequencies are free and then only transmit according to an appropriate range of parameters such as power levels and out-of-band emissions. There are several strategies to determine when a frequency is unoccupied: **SS** (also called detection) where devices monitor frequencies for any radio transmissions and if they do not detect any, assume that the channel is free and can be used; or **geolocation database** where devices determine their location and query a geolocation database which returns the frequencies they can use at their current location [Ofc09]. Although WSD based on SS presents many advantages as its low-cost infrastructure, its major drawback is that sensing to very low signal levels is costly and possibly not achievable, giving rise to a not accurate detection and causing interferences to DTT transmissions [Ofc10].

In Barbiroli et al. [BCGP12], authors propose three different approaches to determine whether a frequency is free, based on geolocation databases: based on a threshold approach applied to a coverage map; based on the location probability; and a combination of geolocation approach and field strength measurements (sensing approach). Results showed that a combined approach could provide better protection to the incumbent services.

After the determination of a free TVWS, a WSD should ensure that its transmitting power is below the maximum equivalent isotropic radiated power (EIRP) allowable for this free frequency in order not to interfere to DTT transmissions. Authors Petrini and Karimi [PK13b] developed and improved method for the computation of the DTT location probability and the calculation for the maximum permitted EIRP with an improved accuracy and less computational complexity. The DTT location probability is introduced as a metric for quantifying the quality of the DTT coverage which will be degraded in presence of WSD. These algorithms are improved in Petrini et al. [PMB13], based on carrier-to-interference ratio (C/I) calculation.

A review about some trials performed in Cambridge about the coexistence of DTT with WSD is presented in Cataldi et al. [CL12].

The use of crossed polarisations between DTT and WSD has also been addressed as a strategy for increasing the protection of both the DTT and WSD [Bro14] when the broadcast transmitter is vertically polarised and the WSD horizontally polarised. Furthermore, regardless of the polarisation of the PU, an increase in path loss from the WSD to the PU should be accounted if the WSD has horizontal polarisation [BT14a].

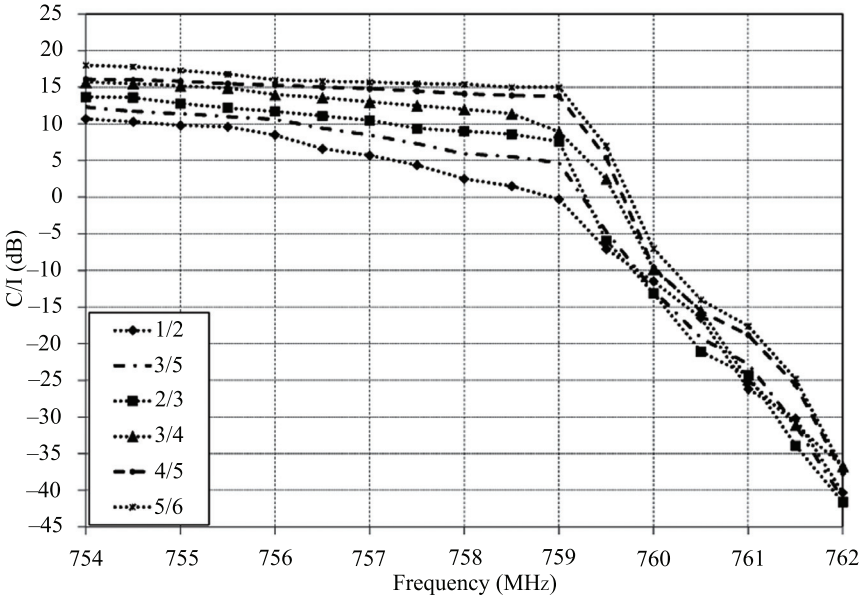
Many standardisation groups have taken care about the use of TVWS for wireless personal area network (WPAN), wireless local area network

(WLAN) and wireless regional area network (WRAN) producing 802.15.4m, 802.11af and 802.22 standards, respectively. Moreover, the 802.19.1 standard published in 2014 regulates the coexistence among multiple TV white space networks. In Fadda et al. [FMV+13] and Angueira et al. [AFM+13], authors evaluated through measurements in laboratory the PR required at a DVB receiver in case of being interfered by an IEEE 802.22 WRAN signal (digital video broactasting (DVB) wanted signal). Regarding the influence of the parameters of the digital video broactasting-terrestrial 2nd generation (DVB-T2) OFDM signal on the PR, measurements showed that the PR decreases as the DVB-T2 code rate does [AFM+13]. The modulation type also affects the tolerance to interference, being 7 dB higher for 64QAM compared to 256QAM for a constant pilot pattern (PP) as shown in Figure 6.21. The use of rotated constellations did not have any influence on the PR and hence, on the interference levels.

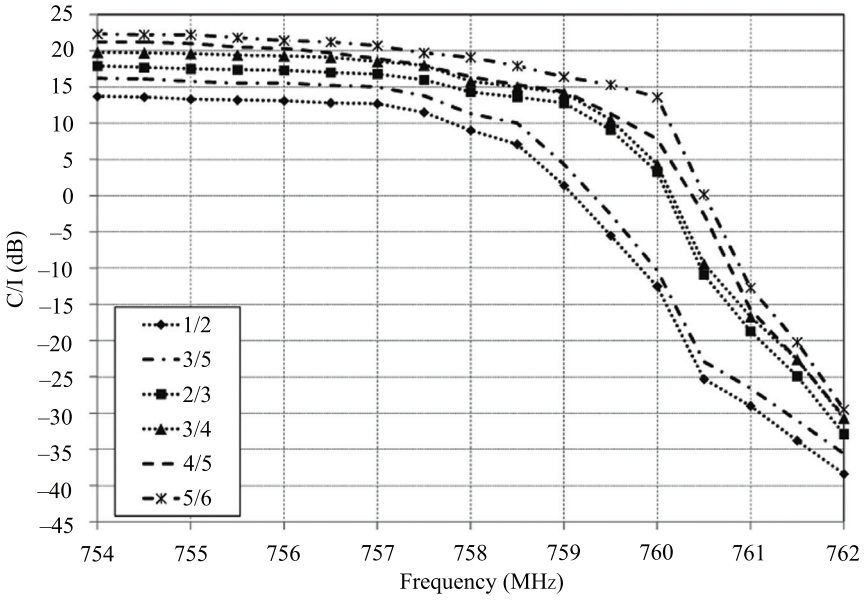
Following the analogue switch off, the ITU in the 2007 World Radiocommunication Conference (WRC-07) decided to allocate the upper part of the former analog TV band, that is, from 790 to 862 MHz, to IMT technologies (with 1 MHz of band guard between broadcasting and mobile services in ITU Region 1). This allocation was also called the digital dividend (DD) [Sam12]. This decision was supported by studies carried out by the European conference of postal and telecommunications administrations (CEPT) reported in CEPT [CEP09a] and [CEP09b], where some technical conditions for IMT emission in the DD band were given.

However, the restrictions established by CEPT do not assure avoiding interferences in the DD. Thus, for a digital video broactasting-terrestrial (DVB-T) reception interfered by a LTE DL signal in adjacency (1 MHz of band guard) Barbioli et al. [BCF+13] obtained that for broadband receivers (mainly installed in blocks of flats in urban areas), an area around 1.000–1.500 from the BS will suffer from receiver saturation due to the high level of interfering power, impeding the successful reception of any DVB-T channel. In case of a typical set top box (STB) receiver the probability of interference is lower than in the previous case but not negligible.

In addition, the ITU 2012 world radiocommunication conference (WRC-12) concluded with a decision to allocate additional ultra high frequency (UHF) spectrum to mobile services and invited to perform further coexistence studies and-report the results to the next WRC-15. The new mobile allocation, also known as second digital dividend (DD2), is to be made in Region 1 in the 700 MHz band (the actual range is to be decided in 2015 world radiocommunication conference (WRC-15)). The main difference compared



(a)



(b)

Figure 6.21 C/I requirements for all code rates for GI 1/16, PP4, and (a) 64QAM, and (b) 256QAM.

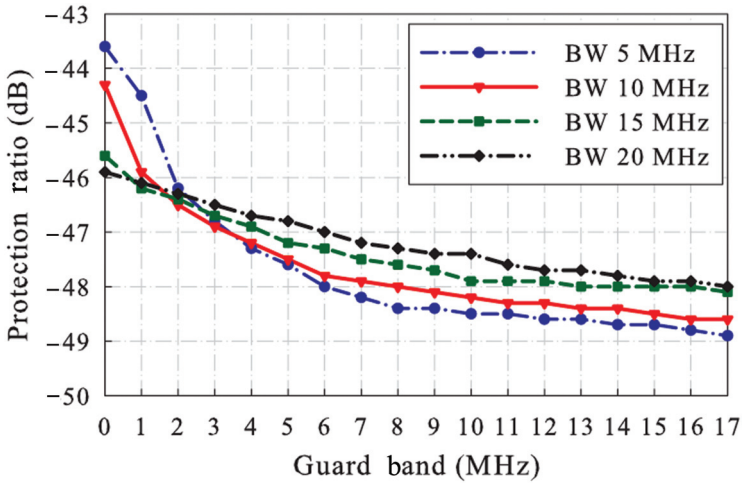
to the 800 MHz band lies in the fact that the UL is located in the lower part, instead of the DL.

In Fuentes et al. [FGGP+14b], authors present a thorough analysis of the the influence of the PHY parameters of DVB-T2 and LTE on the PR required for a-DVB-T2 wanted signal interfered by a LTE signal for both DL and UL cases and in the DD band. UL results to be the most interfering link as far as required PR are worse (higher) than those which are required for DL as observed in Figure 6.22. Furthermore, LTE signals with lower bandwidths are less interfering if LTE operates in adjacency (guard band higher than 9 MHz). In the same study, authors conclude that for LTE-DL operating in adjacency to DVB-T2, the higher the traffic loading, the higher the interference level. The contrary effect was observed for LTE-UL.

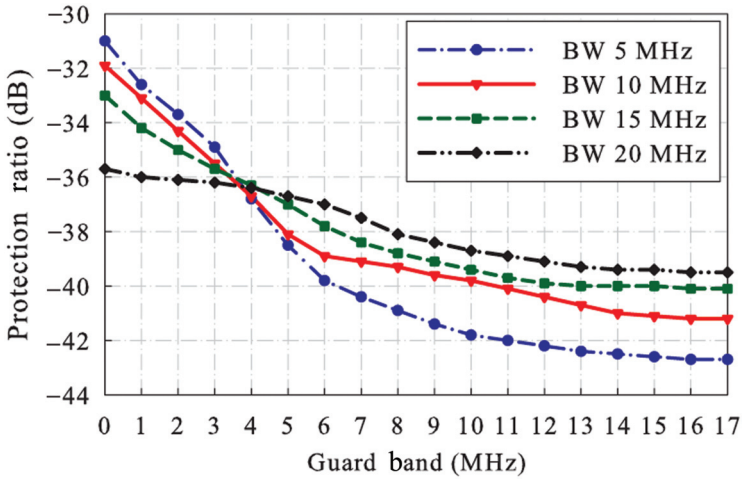
The PR obtained in Fuentes et al. [FGGP+14b], were used in Fuentes et al. [FGGP+14a] for planning studies in order to evaluate the coexistence of DVB-T2 and LTE in RL scenarios for DD and DD2. Thus, for the DD (DL in adjacency with 1 MHz guard band) case and fixed outdoor reception, with the receiver in the DVB-T2 coverage threshold (minimum DTT field strength), an average protection distance to LTE BS of 330 for urban and 580 m for rural environments must be leaved. In case of the DD2 (UL in adjacency with 9 MHz guard band) and DVB-T2 and fixed outdoor reception, a filter is not necessary if the LTE UE transmitted power is below 11 dBm. If portable indoor reception is considered, typical values of LTE UE transmitted power, assured a protection distance of 0.8 m in rural environments and 0.25 m in urban environments without the use of a filter in the DVB-T2 receiver.

The concept of micro TV white space (μTVWS) was introduced in Martinez-Pinzon et al. [MFGC15] to define the scenario in which the DTT signal is broadcasted to rooftop reception, and hence, obstructed for an indoor environment. Thus, these broadcast frequencies could be re-used in indoor for IMT technologies in the UHF band. In Martinez-Pinzon et al. [MFGC15], the authors perform some measurements in laboratory to evaluate the feasibility of the introduction of indoor LTE-A femtocells using DTT frequencies. The main conclusion of this study case is that, under the appropriate restrictions in EIRP, a LTE-A femtocell could operate in frequencies adjacent or even partially overlapping a DVB-T2 channel. The results fix the power limit for LTE in ranges from +20 dBm to –9 dBm depending on the proximity of the LTE-A central frequency to the DTT carrier. The co-channel case was disregarded.

Apart from the compatibility and coexistence between DTT and broadband and mobile communications, nowadays in Europe coexists the first and second



(a)



(b)

Figure 6.22 PR for fixed DVB-T2 reception, as a function of the separation between the carrier frequencies of DVB-T2 and LTE channels, for several values of the LTE channel bandwidth. (a) LTE-DL, and (b) LTE-UL.

generation of the European digital television standards, DVB-T and DVB-T2, with the mobile television standard, digital video broactasting-handheld (DVB-H). Thus, in Mozola [Moz12, Moz14], the author investigated the coexistence of these terrestrial and mobile broadcasting standards when a DVB-T2 signal interfered DVB-H reception.

6.5.3 Spectrum Management

Management of radio spectrum refers to the regulating process of RF which is necessary to efficiently use of such a limited spectrum resource. Because of increasing demand for more users and new services such as 4G mobile technology and beyond, spectrum management has become significant issue. During the last years, for efficiently use, sharing, coexistence, and aggregation of RF spectrum, several techniques have been introduced and different research initiatives have been conducted. In this section, recent researches results in the context of sharing, monitoring and aggregation of RF spectrum are presented.

When the primary is a rotating radar, opportunistic primary–secondary spectrum sharing can still be considered, as proposed in Saruthirathanaworakun et al. [SPC11] and as shown in Figure 6.23. A secondary device is allowed to transmit when its resulting interference will not exceed the radar's tolerable level, perhaps because the radar's directional antenna is currently pointing elsewhere, in contrast to current approaches that prohibit secondary transmissions if radar signals are detected at any time. The case where a secondary system provides point-to-multipoint communications utilising OFDMA technology in non-contiguous cells is considered. This scenario might occur with a broadband hotspot service, or a cellular system that uses spectrum shared with radar to supplement its dedicated spectrum, as described in Saruthirathanaworakun et al. [SPC11]. The secondary system considered provides point-to-multipoint communications in non-contiguous cells around the radar. Unlike existing models of sharing with radar, our model allows secondary devices to adjust to variations in radar antenna gain as the radar rotates, thereby making extensive secondary transmissions possible, although with some interruptions. Thus, sharing spectrum with rotating radar is a promising option to alleviate spectrum scarcity. Additional technical and governance mechanisms are needed to address interference from malfunctioning devices [Peh11].

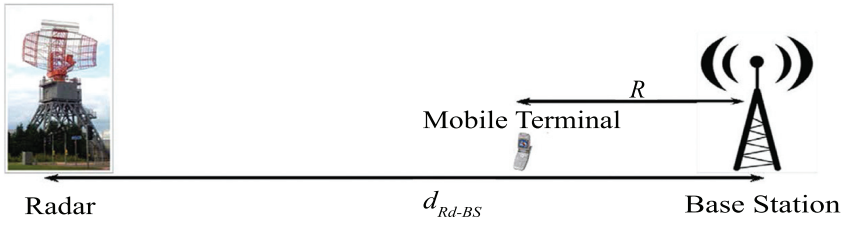


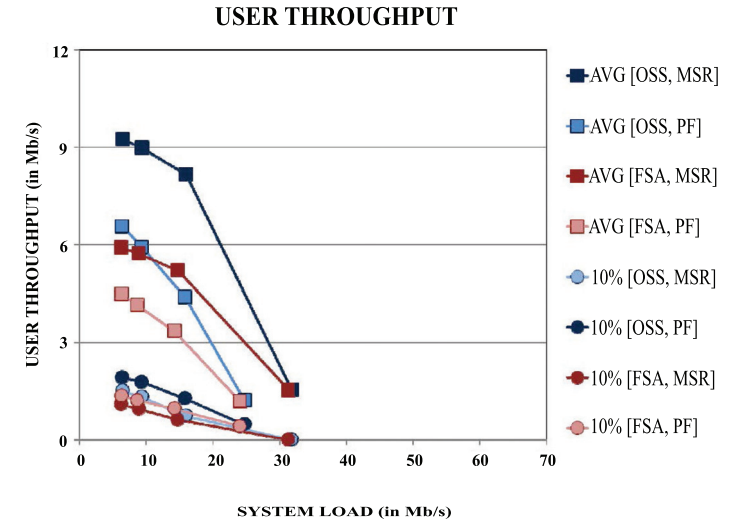
Figure 6.23 Sharing scenario with rotating radar.

It was also found that the secondary will utilise spectrum more efficiently in the downstream than in the upstream. Hence, spectrum sharing with radar would be more appropriate for applications that require more capacity in the downstream. The fluctuations in perceived data rate make sharing spectrum with radar attractive for applications that can tolerate interruptions in transmissions, such as video on demand, peer-to-peer file sharing, and automatic metre reading, or applications that transfer large enough files so the fluctuations are not noticeable, such as song transfers. In contrast, spectrum shared with radar would be unattractive for interactive exchanges of small pieces of data, e.g., packets or files, of which instantaneous data rate matters for the performance, such as VoIP.

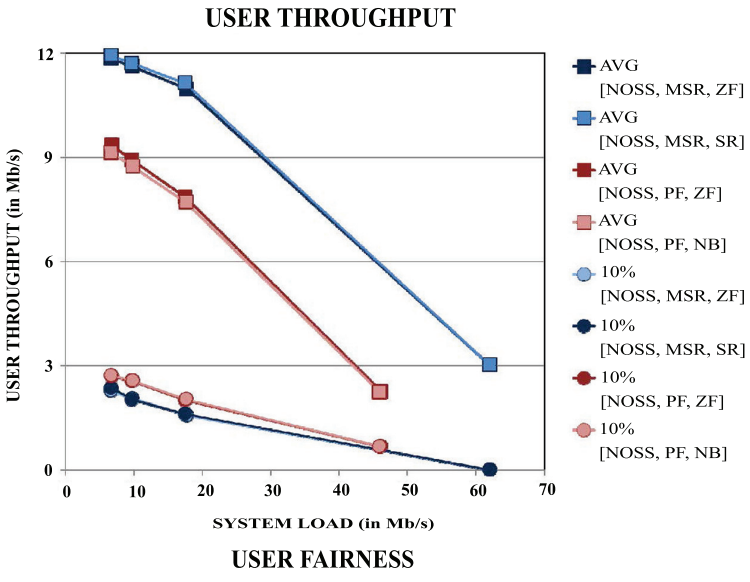
We assess the system-level performance of non-orthogonal spectrum sharing (NOSS) achieved via maximum sum rate (SR), Nash bargaining (NB), and zero-forcing (ZF) transmit beamforming techniques, (Figure 6.24). A lookup table-based PHY abstraction and RRM mechanisms (including packet scheduling) are proposed and incorporated in system-level simulations, jointly with other important aspects of network operation. In the simulated scenarios, the results show similar system-level performance of SR (or NB) as ZF in the context of spectrum sharing, when combined with maximum SR (MSR; or proportional fair PF) packet scheduler. Further sensitivity analysis also shows similar behaviour of all three beamforming techniques with regard to the impact on system-level performance of neighbour-cell activity level and feedback error. A more important observation from our results is that, under ideal conditions, the performance enhancement of NOSS over orthogonal spectrum sharing (OSS) and fixed spectrum assignment (FSA) is significant.

Regulators face new challenges for radio spectrum monitoring in the near future, not only because the intensification of the radio spectrum usage associated to the growth of mobile technologies like LTR, LTE-A, or the new fifth generation (5G), but also because of the new cognitive systems that will be in use in the near future. These challenges (some of them mentioned in the ITU-R Recommendation SM.2039 about spectrum monitoring evolution) leads to the development of new spectrum monitoring systems, smaller and cheaper than the current system in use by regulators.

In Arteaga et al. [NAVA12] and Navarro [NAV$^+$14], a spectrum-monitoring system based on Open SDR systems is described. First, a basic spectrum analyser is implemented and after that, the procedures for accomplishing the ITU-R SM.1392 Recommendation is implemented. This system, named SIMONES implements a SS algorithm based on energy detection, to



(a)



(b)

Figure 6.24 The performance comparison among NOSS, OSS, and FSA combined with different schedulers: (a) Average and 10%-tile throuhgput of OSS, and (b) Average and 10&-tile throughput of NOSS and FSA.

identify signals in a radio band, as well as bandwidth measurement, according to the $\beta/2$ method as recommended in ITU-R SM.443 and Chapter 4 of the Monitoring Handbook for fast fourier transform (FFT)-based systems. For frequency measurements, SIMONES has implemented the system in order to comply with ITU-R Recommendation SM.377, using the GPS locked reference oscillator for USRP. Because the system is based on FFT and Software Defined Radio, it is possible to measure variations in bandwidth and frequency for digital modulations. For occupancy measurement, the energy detection algorithm is adapted according to the ITU-R SM.1880 Recommendation, in order to obtain the occupancy parametre according to the recommendation. The spectrum monitoring system, SIMONES, has four functional components: the monitoring unit (SIMON); a set of drivers to interact with commercial monitoring software (TES Monitor suite); an independent web interface; and a user-based drive test unit.

In the Spectrum Monitoring Handbook 2011 version, a new section is introduced. In Section 6.8, the handbook describes the method for planning and optimisation of the monitoring stations. Since 2012, ITU Working Party 1C is developing a report on Spectrum monitoring network design for the future spectrum monitoring systems. In Navarro et al. [NA13], a modified method is proposed, which uses terrain-based propagation models as well as users density information to determine where the monitoring stations will attend bigger population (i.e., Spectrum Users). The population density criterion points to the prioritisation of monitoring in zones with more population density, which implies a bigger number of spectrum users covered by a monitoring station and therefore a bigger efficiency from the economical point of view, according to the recommendations of ITU-R Report ITU-R SM.2012.

The criterion of the use of semi-deterministic propagation models points to improve the coverage prediction using DTM and considering terrain information and diffraction effects for prediction and monitoring stations location. In mountainous regions, obstacles like big mountains could be quite important and are not considered by models like ITU-R P.1546 or Hata. Therefore, the risk of a miss location of a monitoring station using such models as proposed in the ITU Spectrum Monitoring Handbook 2011, is very high.

In [CMM+11], authors proposed an Integrated iCRRM. The iCRRM performs classic common radio resource management (CRRM) functionalities jointly with SA, being able to switch users between non-contiguous frequency bands. The SA scheduling is obtained with an optimised General Multi-Band Scheduling algorithm with the aim of cell throughput maximisation. In particular, we investigate the dependence of the throughput on the cell

coverage distance for the allocation of users over the 2 and 5 GHz bands for a single operator scenario under a constant average SINR. For the performed evaluation, the same type of radio access technology (RAT) is considered for both frequency bands. The operator has the availability of a non-shared 2 GHz band and has access to part (or all) of a shared frequency band at 5 GHz, as shown in Figure 6.25. The performance gain, analysed in terms of data throughput, depends on the channel quality for each user in the considered bands which, in turn, is a function of the path loss, interference, noise, and the distance from the BS. An almost constant gain near 30% was obtained with the proposed optimal solution compared to a system where users are first allocated in one of the two bands and later not able to HO between the bands, as shown in Figure 6.26.

This work has been published in Cabral et al. [CMM+11].

6.6 Virtualised and Cloud-based Architectures

6.6.1 Architecture

Virtualisation and cloud computing are concepts widely used to enable the share of processing resources. This section presents several approaches on how these concepts have been adopted or extended in RANs, providing novel capabilities to these networks.

In Caeiro et al. [CCC12, 14a, 14b], the virtualisation of the wireless access is addressed. It is based on a network virtualisation environment that

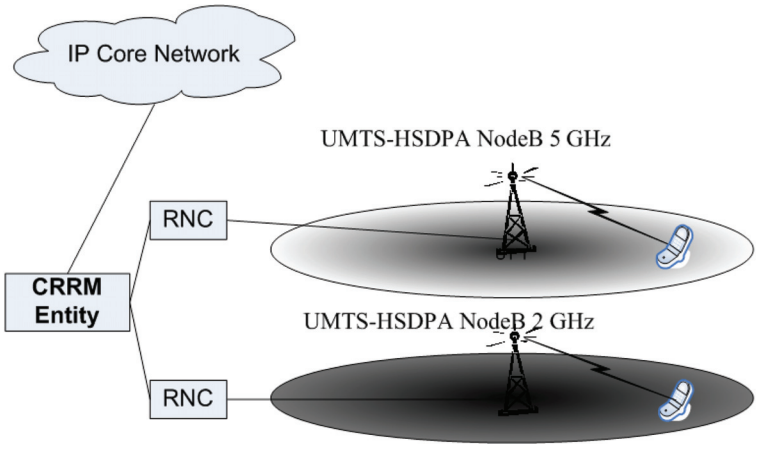


Figure 6.25 CRRM in the context of SA with two separated frequency bands.

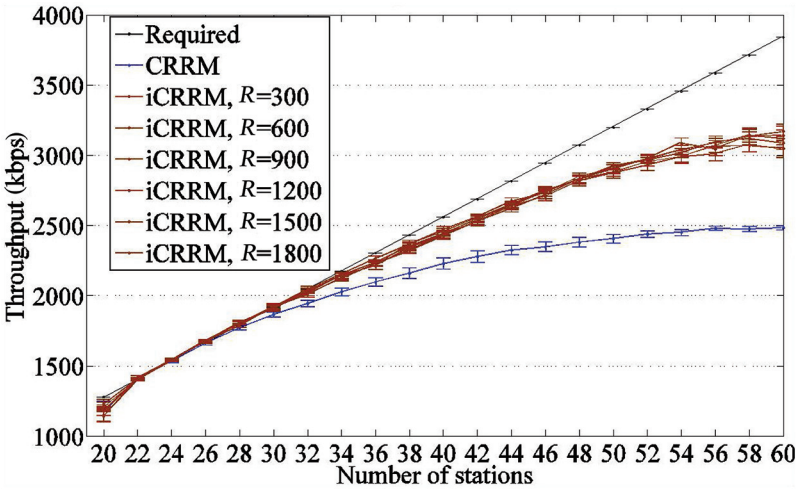


Figure 6.26 Average service throughput with the iCRRM with normalised power.

envisages the existence of multiple virtual networks (VNets) created by a VNet Enabler. A VNet is composed of one or multiple virtual base stations (VBSs), based on resources of heterogeneous RATs. In this environment, a virtual network operator (VNO) is a network operator that does not own any RAN infrastructure and needs wireless connectivity for its subscribers. For each VNO, a VNet is instantiated, settled on demand to satisfy its service requirements, in order to deliver services to their customers.

Similarly, the notion of virtual radio resources, as an aggregate of physical radio resources, and a model for their management is proposed in Khatibi and Correia. [KC14b]. It is a hierarchical management, as depicted in Figure 6.27, consisting of a virtual radio resource management (VRRM) on the top of RRM entities of heterogeneous RANs, designated as CRRM, while the classical local RRMs exist at the bottom. Given the VRRM, the RAN Provider, owner of the physical infrastructure, is capable of offering capacity to multiple VNOs according to certain service level agreements (SLAs).

On the other side, the concept of cloud has been adopted for RANs, where an architecture to offer cloud-based RAN radio access network as a service (RANaaS) is proposed in Ferreira et al. [FPH$^+$14]. It follows the concept of C-RAN, which splits the BS into a remote radio head (RRH) (an antenna and a small radio unit) and a software-based base band unit (BBU) (baseband, control and management functions). Software-based BBUs are deployed and managed on general purpose platforms, by applying virtualisation and cloud

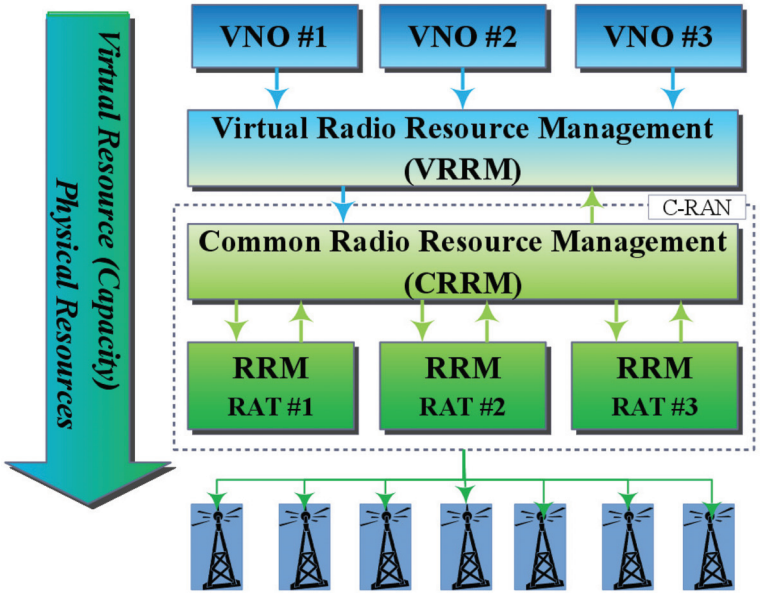


Figure 6.27 VRRM architecture (extracted from Khatibi and Correia. [KC14b]).

paradigms. RANaaS offers to operators an on-demand, scalable and shared RAN, efficiently adaptable to geographic and temporal load variations. It enables fast, dense, cost, and energy efficient deployments (1/3 roll out time, 15% CAPEX, 50% OPEX, 71% energy saving compared to traditional RAN). Also, new players may easily enter the market, allowing a VNO to provide connectivity to its users. One of the key research challenges is to match LTE processing requirements on general purpose processing platforms.

6.6.2 Models and Algorithms

Various models and algorithms are presented in this section to manage virtualised radio resources. An adaptive virtual network radio resource allocation (VRRA) algorithm is proposed in Caeiro et al. [CCC12], which optimises the utilisation of shared resources of heterogeneous RATs, in order to maintain the contracted capacity of VBSs. Initially, the algorithm pre-allocates radio resources units (RRUs) to each VBS. Then, sensitive to VBS changes in capacity, it dynamically uses compensation mechanisms to adequately allocate additional RRUs.

In Caeiro et al. [CCC14a,b], an on-demand VRRA algorithm is proposed, allocating resources elastically to end-users, depending on the current demand.

In order to compensate for the variability of the wireless medium and taking the diversity of available RATs into account, the algorithm continuously influences RRM mechanisms (admission control and MAC scheduling) to be aware of the VBSs state, to satisfy the SLAs.

A VRRM model is proposed in Khatibi and Correia. [KC14b] (Figure 6.27). It estimates the physical RAN capacity and allocates resources to support different service level agreement (SLA) contracts. It translates VNOs' requirements and SLAs into sets of polices for lower levels. VRRM optimises the usage of virtual radio resources, not dealing with physical ones. Nevertheless, reports and monitoring information received from CRRM enable it to improve the policies. VRRM is able to map the number of the available resources to the network capacity, optimising the network throughput and prioritising services with weights.

Obviously, offering the same data rate to a MT with low input SINR requires assigning comparatively more RRUs than to a MT of higher SINR. In Khatibi and Correia [KC15], the effect of channel quality in VRRM is studied. The capacity estimation technique is extended by considering three approaches according to pre-defined assumptions. In an optimistic (OP) approach, all RRUs are assigned to MTs of assumed high SINR. In a realistic (RL) approach, half of RRUs are assigned to MTs of assumed high SINR, while the other half of RRUs is assigned to MTs of low SINR. A pessimistic (PE) approach, all MTs are assumed to have low SINR. These assumptions will result in different boundaries for the possible data rates for each RRU.

An extension of the VRRM model considers traffic offloading support with WLANs [KC14a], where low data rate services are served by cellular RATs while high data rate ones by WLANs.

6.6.3 Scenarios and Results

A scenario with several physical BSs from multiple RATs (TDMA, CDMA, OFDMA, and OFDM) is considered. Three VNets are taken: A and C as Guaranteed Bitrate (GB), for services with stringent requirements; B as Best Effort (BE), for other services. Simulation results show that the adaptive VRRA algorithm [CCC12] satisfies minimum capacity requirements of VNOs. With the on-demand VRRA algorithm [CCC14b], the total capacity contracted for guaranteed VBSs should be limited, according to the average physical capacity. In Caeiro et al. [CCC14a], it is also concluded that by changing the quantity of created VBSs and the contracted data rate, the average data rate and efficiency may decrease if the number of GB VBSs (hence, the contracted

capacity) is higher than the number of BE VBSs. Overbooking the capacity contracted by BE VBSs achieves 30% higher efficiency and data rates.

In Khatibi and Correia [KC14b], a scenario in which coverage is provided through a set of RRHs, is assumed. RRHs support multiple RATs. Similarly, three VNOs exist: VNO GB, where the contract guarantees the allocated data rate to be between 50 and 100% of the required service data rate; VNO BE with minimum Guaranteed (BG), where at least 25% of the required service data rate is guaranteed; and VNO BE. It is concluded that, with VRRM, VNO GB receives the biggest portion (59%) of resources, 7% for VNO BE, and the rest of resources (34%) for VNO BG, better served than the minimum guaranteed. When Wi-Fi coverage is used for traffic offloading [KC14a], results show an increase up to 2.8 times in network capacity, enabling VRRM to properly serve all three VNOs: not only guaranteed services are served adequately, but best effort ones are allocated with a relatively high data rate.

The effect of channel quality on VRRM is evaluated in Khatibi and Correia [KC15]. Figure 6.28 presents the allocated data rate to VNO GB in conjunction with minimum and maximum guaranteed data rate. As long as one has the data rates within the acceptable region (i.e., shown by the solid colour), there is no violation to the SLA and guaranteed data rate. It can be seen that in OP approach up to 600 subscribers, the maximum guaranteed data rate is

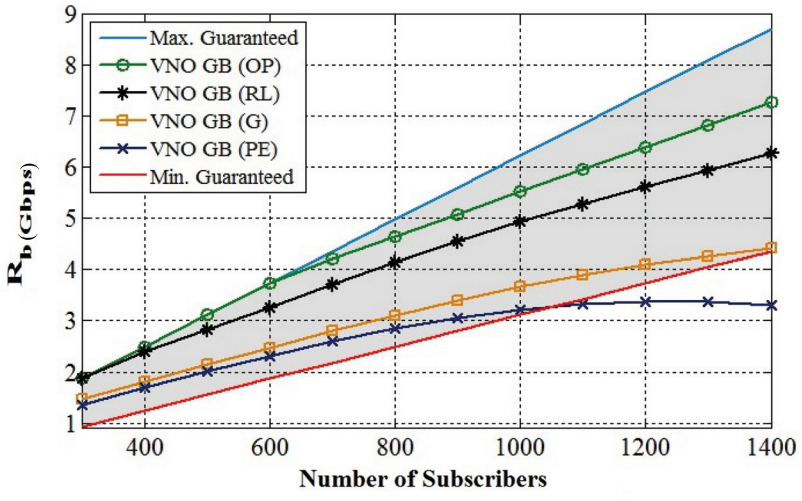


Figure 6.28 Allocated data rate to VNO GB for different approaches (extracted from Khatibi and Correia. [KC15]).

offered to this VNO. Considering the other approaches, as the number of the subscribers increases the allocated data rate moves toward the minimum level of the guaranteed data rate. Finally, in the PE approach, as the number of subscribers passes 1100, violations to minimum guaranteed data rate can be observed: this means that the network capacity is too low so that the model faced resource shortage and it has to violate the minimum guaranteed data rate of this VNO.

6.7 Reconfigurable Radio for Heterogeneous Networking

One of the main issues in performance evaluation of future mobile networks is the increasing mutual interaction among layers: thus, studying the performance of new systems requires an accurate cross-layer simulation, and possibly subsequent cross-layer optimisation, as the outcome of this interaction affects the QoE of users.

This issue attracted the interest of COST IC1004 researchers from the very beginning of the Action, when the authors of Werthmann et al. [WKBM11] developed a simulation tool to study these complex interactions.

They built an event-driven simulator that, as such, does not require a continuous time clock; the network simulator controls virtualised computers with unmodified operating system and unmodified applications. To model the interactions between the virtualised hosts and the network simulation, both need to share the same time source.

Since the network simulation performs all actions based on events, the simulated time follows from these events and is not related to the real-time. The virtualised computers need to adapt to this event-based time. Therefore the simulation has to control the time of these virtualised computers, and simulations may run slower or faster than real-time depending on the computational complexity of the PHY model. The tool is scalable, i.e., it can include a variable number of virtual machines.

To emulate the users' interactions with the virtual computers, the authors implemented an interface to the virtual computers keyboard. It is thus possible to schedule events in the simulation libraries calendar which send keystrokes to the virtual computers. Hereby it is possible, for example, to load different web pages in predefined intervals of time.

Thus, the authors built a framework for network performance evaluation, which can be used to evaluate applications performance in different situations as well as to evaluate the impact of PHY and MAC algorithms.

Cross-layer interaction is dealt with also in Litjens et al. [LGS+13], presented by representatives of the SEMAFOUR FP7 project. Their approach

builds upon and extends the concept of SON, developing a unified management framework that integrates the existing and future advanced SON functions across several types of RAT.

The first phase of stand-alone SON use cases and possible solutions focused on self-organisation mechanisms (such as mobility robustness optimisation, interference management, coverage/capacity optimisation, energy savings) that were isolated from one another. Typically, these solutions target individual RATs and cellular layers, rather than addressing the network as a whole.

The configuration and the optimisation of individual SON functions becomes highly complicated and the likeliness of conflicts increases significantly with an increasing number of different SON functions operating in parallel and in a non-coordinated manner, covering multiple layers/RATs and coming from different manufacturers.

The paper presents the self-management system for the unified operation of multi-RAT/layer networks, defined within the SEMAFOUR project. As some key partners of the projects were also institutions participating in COST IC1004 Action, mutual exchange of experiences, and information (within the confidentiality limits required by the project consortium agreement) provided benefits to both parties.

Current and future mobile networks are different from their predecessors in that they support a variety of multimedia applications with different (and often contrasting) requirements in terms of quality parameters such as throughput, packet loss, delay, and jitter. Thus, QoS definitions and targets that originated from previous generations of mobile networks (or even from fixed telephony) are not sufficient for current and forthcoming networks. The research community is striving to derive suitable QoE metrics from measurable QoS parameters, and COST IC1004 is no exception to this.

The authors Robalo and Velez. [RV13] presented their work on this theme. They define QoS requirements for various classes of services, including context-based information, i.e., information belonging to location dependent services. The available multimedia applications are subdivided in four broad classes, namely gaming, video, audio, and web-browsing.

The authors then map QoS onto QoE for various service classes, utilising (when available in literature) third party experiments to evaluate mean opinion score (MOS) as a function of QoS; for web-browsing and audio applications, since MOS results are scarce, the model considers the ITU-T G.1030 and G.107 recommendations, respectively, which are themselves base upon experimental MOS measurements.

Finally, the authors propose a unified model, that computes QoE as a weighed average of QoE evaluated for different applications. This allows building an effective QoE control mechanism onto measurable QoS parameters for multimedia networks, i.e., for improving packet scheduling in mesh and/or CR networks. In particular, when considering the additional delay and possible packet loss induced by changes in spectrum utilisation in CR networks, the model allows to verify their impact on the actual overall QoE.

The variety of services and applications is not the only factor adding to the complexity of current and future mobile networks: also cells appear in a variety of sises and types, while different RATs - and hence different types of AP - are available, according to the HetNet paradigm. Furthermore, also the backhaul network plays a key role in determining the overall performance of a given system: transport capacity in the fixed segment of the network may constitute a bottleneck, if it is not adequately planned to match the radio part requirements in different areas and parts of the overall network.

This issue is tackled in Olmos et al. [OFGZ13], whose authors propose a new algorithm for cell selection. Their main assumption is that cell selection strategies must be extended to consider backhaul load conditions as an additional input.

The authors consider three different cell selection algorithms:

- Best Server Cell Selection (BS_CS): the call is assigned to the best server, if it has enough available capacity. It is used as a reference for the other two algorithms;
- Radio Prioritised Cell Selection (RP_CS): if a call cannot be served by the best server, the algorithm redirects it to another BS in the Candidate Set, if possible (often used in legacy networks);
- Transport Prioritised Cell Selection (TP_CS), proposed by the authors: if transport load in the best server is high, the call is served by another node to avoid congestion at transport level. Weights are used to balance load among nodes according to their transport occupancy.

The authors built an analytical model, based on multi-dimensional Markov chains, to assess the performance of these cell selection algorithms.

Of course, having more than one cell in the candidate set implies a reduction in the overall network capacity and performance, as one MT occupies resources in more than one cell.

The systems composing the analytical model are undetermined, since there are more unknowns than linearly independent equations; then, among

all possible solutions, the one providing the minimum radio performance degradation should be chosen. In order to obtain such a solution, the authors propose an iterative procedure which is valid for arbitrary number of candidate cells.

A Monte Carlo simulator has been written in order to fully validate the results of the Markov model, resulting in excellent agreement with the analytical model and thus obtaining a mutual validation. One key result found is that the trunking gain achievable by TP_CS is greater in case of non-uniform capacity in the fixed network with respect to the uniform case. The proposed strategy is able to achieve the desired trunking gain while reducing the amount of radio degradation that we would have in case of using traditional cell selection schemes based only on radio metrics.

One powerful concept to cope with the multi-faceted complexity of current and future mobile networks is software defined network (SDN), whose opportunities are thoroughly discussed in Chaudet and Haddad [CH14]. SDN is a network paradigm that relies on the separation of the control and forwarding planes in IP networks. The interconnection devices take forwarding decisions solely based on a set of multi-criteria policy rules defined by external applications called *controllers*. It is possible to let multiple controllers manage each element of a given network, which allows creating independent networks on the same physical infrastructure.

The ability to separate the IP flows space in distinct subspaces is referred to as *slicing*. Implementing SDN requires at least being able to define slices and to limit interactions between these slices, and to let the network devices measure and report their status to the relevant controllers, which are non-trivial operations with the wireless medium.

The authors highlight the issues raised when trying to use the SDN paradigm in mobile networks: non-ideal separation between "slices" sharing the same band, coordination among different operators, interference, HO, changes in link conditions, etc. However, SDN can potentially bring tremendous benefits, listed below:

- *Improved end-user connectivity and QoS* by inter-operator cooperation (e.g., an AP of an operator routing packets of a customer of another operator when the latter is congested)-
- *Multi-network planning:*
  - SDN allows creating zone-specific controllers that transcend operators, suggesting channel selections and power control to participating ACs, based on sensed mutual interferences levels. Power control

helps reducing interferences by attaching users to the closest AC. The more ACs are available, the more efficient this process will be, and the capacity of wireless SDN to aggregate in a single VNet multiple BSs belonging to multiple ISPs eases the problem. Thus, the controller can create locally all the possible interferences graphs and select the one that maximises coverage while minimising interferences.
  - SDN can support soft HO between different technologies, by replicating packets in the core network and forwarding them through different radio interfaces.

- *Security:* the monitoring capacity of SDN allows detecting intrusions and malicious attacks; cooperation among APs enhances this capability.
- *Location*: if a terminal can connect to different networks, having a controller that supervises all networks will allow faster and more accurate determination of its position (even in indoor environments where satellite-based location is not available).

However, some issues shall be tackled prior to actual exploitation in commercial networks. There are confidentiality issues for the operators (e.g., they may not be willing to disclose technical details of their networks) and privacy issues for the users (e.g., a user may not want to let other operators know which websites he/she browses). For the time being, user data and network information can't be freely shared among different operators; however, a first step could consist in introducing the above discussed capabilities in different networks owned by the same operator.

One well-known technique to increase the capacity of mobile networks is the reduction of the cell sise. Having smaller cells, however, implies an increase in the HO rate, and hence of the probability of dropping a connection.

Previous experience shows that generally, in a small cell network (SCN), HO failure rate becomes excessive for speeds of about 30 km/h. The authors of Joud et al. [JGLR14] propose a solution allowing to use SCNs for speeds up to about 60 km/h without a significant degradation of HO performance.

Obviously, the most critical type of HO is the small cell to small cell handover (SS HO).

The principle of the proposed scheme is, therefore, to reduce the frequency of SS HOs and consequently handover failure (HOF) occurrence. This is achieved through dual connectivity with C-/U-plane split and DL reinforcement by means of cooperative multi-point (CoMP) in the SCN. The basic idea is transmitting the U-plane over more than one small cell for the users

beyond a certain speed limit. Groups of small cells are dynamically created and perform coordinated scheduling just for these users, being likely to have high HOF rates.

The simulated scenario consists of macrocells operating at 2 GHz, that provide complete coverage of the service area, plus small cells at 3.5 GHz, deployed at random locations (Figure 6.29).

For the higher mobility case, the UE is initially connected to the best macrocell and establish normal communication in both C-and U-planes.When the cell receives a report indicating an A3 event towards a small cell, the U-plane is transferred, the UE keeps receiving C-plane from the macro-cell. Therefore, it does not perform a conventional HO since radio resource control (RRC) messages are still sent/received though the macrocell. The U-plane is transferred to a group of small cells that will schedule the UE in a coordinated manner. Note that cooperative groups are not created for pedestrian (or low mobility in general) users, which still perform classic HO among all cells.

The presence of many UEs requiring cooperative groups might well reduce the small cell layer capacity dramatically: for this reason, the authors limited the number of UEs in cooperative transmission manner served by a small cell to $k = 3$, and the number of small cells that can serve a single UE to $n = 3$ as well (one serving cell and two co-serving cells).

Simulations show that the proposed method allows an improvement of HO performance for fast moving users, at the cost of reduced throughput: more investigations are necessary to determine the optimum solution according to the actual traffic distribution and cellular layout.

6.7.1 Advances in Heterogeneous Cellular Systems

As services migrate from voice centric to data and multimedia centric, which requires increased link budget and coverage extension to provide uniform user experience, traditional networks optimised for homogeneous traffic face

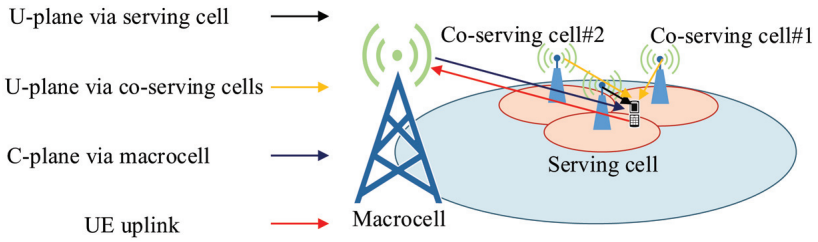


Figure 6.29 Coverage scheme proposed by Joud et al. [JGLR14]

unprecedented challenges to meet the various demands effectively. In this context, third generation (3G) 3rd generation mobile group (3GPP) LTE-A has started a new study to investigate HetNet [DMW+11] deployments as an efficient way to improve the system spectral efficiency as well as effectively enhance network coverage. Under this architecture, a number of wireless technologies such as small/micro cell deployment [SCO13a], Device-to-Device [DRW+09], and multicell cooperation [SWPR13] have been proposed and/or developed in order to enhance spectral and energy efficiency. Especially, in the LTE-A HetNet architecture, the deployment and configuration of small cells is considered a key technology to provide high capacity, good coverage, high spectral efficiency, and high energy efficiency. In the above context, some of the recent developments in resource allocation, energy efficiency, diversity, and adaptive antenna techniques for HetNet systems are discussed in the following subsections.

6.7.2 Optimal Resource Allocation

One of the main challenges in HetNet systems is how to optimally allocate the available resources such as frequency, power etc. in order to satisfy the users' service requirements as well as to guarantee the economic, environmental, and operational sustainability of the service provider. The resource allocation problem in wireless communication systems has evolved incrementally in the following three stages [NPGP13]: (i) performance management [SAR09], (ii) QoS management [ZY10], and (iii) interference management [KYRC07]. The first stage aims to optimise the meet a performance metric at the cost of denying the service to the worst users. In the second stage, the system tries to best-effort traffic users, subject to full satisfaction of the guaranteed traffic users. Subsequently, the third stage tries to manage the interference allocating by both frequency and power in an intelligent manner. Several authors have proposed different ways of handling resource allocation problem in HetNet systems [BD08, MH13, PNSR14].

The service providers point of view can be incorporated while designing a n acceptable resource allocation strategy in HetNet systems [NPGP13]. For a service provider, success or failure of the resource management task is directly related to the level of satisfaction of services demanded by users, on the basis of an efficient usage of available resources. To achieve this, the following two schemes can be considered while designing a model: (i) efficacy-oriented scheme which aims to find the best possible solution by improving the model performance in a variable time interval, and (ii) efficiency-oriented scheme which aims to find an acceptable solution in a

defined time interval. Regarding the strategy for allocating frequency blocks to users, the following two schemes are considered: (i) a random allocation strategy with reuse factor 1, which allocates all available resource blocks, and (ii) a random allocation strategy with a flexible frequency reuse factor. Furthermore, the proposed resource allocation scheme in [NPGP13] considers the following aspects: (i) the resource allocation task solves the problem using a centralised scheme, (ii) it uses a super frame level scale of time, (iii) The solution is based in a scenario with heterogeneous services and architecture, (iv) it establishes a balance between efficacy and efficiency, (v) it tries to satisfy the SINR required to offer the service demanded by each user deployed in the scenario, and (vi) three different strategies to schedule users in picocells. Moreover, the SINR ratio can be used as a key parametre in order to evaluate the resource allocation task.

6.7.3 Energy Efficiency

In the coming years, the growth of wireless networks will significantly increase the energy consumption of information and communication technologies by approximately 15–20% [CZB+10]. This will cause an increase of environmental pollution, and impose more and more challenging operational cost for the operators. In this context, green communications, which aim to improve the energy efficiency of future wireless communications, has received important attention for researchers and wireless industries. Among many wireless equipments, the RAN consumes the most energy and it is necessary to control the DL power since BSs are the primary energy consumers of cellular networks. Deployment of low power BSs in traditional macrocells has been considered to extend coverage and simultaneously reduce the load of the macrocellular networks [3GPP10]. The integration of picocells in cellular networks is a low power, low-cost solution to offer high data rates to outdoor customers and simultaneously reduce the load of the macrocellular network. From the study of the effects of joint macrocell and residential picocell deployment on the network energy efficiency by Claussen et al. [CHP08], it can be noted that introducing picocells within macrocells can reduce the total energy consumption by up to 60% in urban areas with high data rate user demand. However, the massive and unplanned deployment of small cells and their uncoordinated operation may result in harmful crosstier interference and raise important questions about energy efficiency. In the following subsections, we provide various approaches considered for energy efficiency in HetNet systems.

6.7.3.1 Users' social pattern perspective for energy efficiency

In the last several years, social networks such as Facebook, Twitter, Microblog, etc. have attracted billions of active users and the number of users are increasing exponentially. In such networks, some users may exhibit a kind of social pattern/behaviour and tend to operate in similar characteristics. Several research works have focused on the users and traffic characteristics to optimise cellular networks. For example, the study of connections among users is considered as one approach [CSB10] in the context of traditional social networking and another approach can be to consider spatial and temporal characteristics of users by taking into account of their social characteristics [XZG14]. Due to the social nature and human habits, the probability that users close in vicinity (e.g., in the coverage of one or several BS) have similar habits, pattern/behaviour and mobility rules may be high. Furthermore, understanding and modelling a user pattern is crucial for the design of LTE-A HetNet systems and services, especially for the deployment and configuration of small cells. The user pattern can be used as the basis for the performance optimisation of the LTE-A system.

In the above context, two energy-efficient transmission control schemes in cellular networks for real-time and best-effort services have been proposed in [ASGL13, HWZJ12] exploiting the user pattern/behaviour. User Social Pattern (USP) can be considered as an important parametre which basically characterises the general user behaviour, pattern and rules of a group of users as a social manner. To mathematically describe the USP, Gini coefficient used in statistics and economics can be considered as a reference. The Gini coefficient (also known as the Gini index or Gini ratio) is a measure of statistical dispersion intended to represent the income distribution of a nation's residents. Similar to incomes, users and traffics in HetNets can also be described and modelled [XZG14].

6.7.3.2 Stackelberg learning framework for energy efficiency

In spectrum-sharing HetNet systems, the cross-tier interference between macro-cells and small-cells and the co-tier interference among small-cells may greatly degrade the network performance. Without effective interference management, the overall energy efficiency of HetNet system will become even worse than that of a network with no small-cell deployments. Various approaches have been proposed to mitigate the above interferences. Kang et al. [KZM12] formulated the resource allocation in two-tier femtocell networks as a stackelberg game (SG), where the MBS protects itself by pricing the interference from femtocell users. Similarly, Chandrasekhar et al. [CAM+09]

proposed a distributed utility-based SINR adaptation algorithm to mitigate the cross-tier interference, under the assumption that macrocell BS knows the exact positions of overlaid users. However, in practice, the number and the locations of small-cells may be unknown, which results in unpredictable interference patterns. For such dynamic HetNet scenario, all cells tend to be autonomous. In this context, Lien et al. [LTCS10] applied the CR technique to identify available radio resource in a macrocell and authors in Chen [Che14a] formulate a SG to study the power allocation in HetNet systems with the aspect of energy efficiency. In the considered scenario, macrocells are considered as foresighted leaders while the small-cells as followers and the their objectives are to learn to optimise the power adaptation policies.

The SG is a strategic game in which there exists a leader that moves first, while a group of followers make their moves sequentially. In the macro-small cell coexistence scenario, the macrocell used can be referred to as the leader, while the small cell users as the followers. Accordingly, the formulated SG game can have two levels of hierarchy [Che14a]: (i) the leader behaves by knowing the reaction function of the followers, (ii) given the action of the leader, the followers play a non-cooperative sub-game. A Stackelberg equilibrium (SE) describes the optimal strategy for the macrocell user. In stochastic learning game, users are assumed to behave as intelligent agents in the formulated SG. The objective of each user is to maximise the expected utility, which reflects the users satisfaction of executing a specific power adaptation policy. Users with learning ability learn to adapt to the surrounding networking environment to maximise its individual expected utility. One of the possible ways of implementing adaptation mechanisms is Q-learning [WD92] where the users' power adaptation policies are parametreised through Q-functions that characterise the relative expected utility of a particular transmit power level. In Q-learning, users try to find the optimal Q-values in a recursive way.

6.7.3.3 Capacity and energy efficiency

There exists a trade-off between energy efficiency and system capacity when deploying small cells. The deployment of low power BSs may result in maximising the system capacity, however, a high number of lightly loaded small cells increases the network energy consumption. Therefore, to evaluate different cell topologies for reducing the energy consumption, it is important to use adequate energy efficiency metrics [CKY10]. The energy efficiency is usually measured by the traffic capacity divided by the power consumption of the BSs and it is considered as a key performance indicator. The impact

of deploying picocells on the capacity and energy efficiency of macro-picocell two tier HetNets has been studied in Obaid and Czylwik [OC13]. Subsequently, an adaptive power allocation based on an iterative-water filling scheme is presented in order to control the DL transmit power for macro and pico BSs.

In a linear power model proposed by Richter et al. [RFMF10], the total power consumption of each BS changes linearly with respect to the average transmit power. Therein, the power consumption of different BSs is modelled as the sum of two parts. The first part describes the static power consumption, which is consumed by the regular operation of BS. The second part is the dynamic power consumption, which depends on the cell load. Let P^{M} and P^{P} denote the power consumption of macro and pico BSs, respectively. Then the relation between the transmit power and the total power consumption of each macro and pico BS is given by Richter et al. [RFMF10]

$$P^{\mathrm{M}} = N_{\mathrm{sec}} N_{\mathrm{ant}} (\alpha_{\mathrm{M}} P^{\mathrm{M}}_{\mathrm{Max}} + P_{\mathrm{CM}}) \tag{6.8}$$

$$P^{\mathrm{P}} = \alpha_{\mathrm{P}} P^{\mathrm{P}}_{\mathrm{Max}} + P_{\mathrm{CP}}, \tag{6.9}$$

where N_{sec} and N_{ant} denote the number of sectors in a macro site and the number of transmit antennas of a macro sector's BS, respectively, α_{M} and α_{P} represent the power consumption coefficients that scale with the transmit power due to amplifier, cooling of sites and feeder losses, P_{CM} and P_{CP} denote the fixed amount of powers consumed by the BS due to signal processing, battery backup and other auxiliary equipments.

Regarding energy efficiency metrics, although several metrics have been proposed in the literature [CKY10], the traffic capacity divided by the power consumption of the BSs can be considered a suitable energy efficiency (η_{E}) metric while taking into account of the power consumption of the BSs, which can be defined as by Obaid and Czylwik. [OC13]

$$\eta_{\mathrm{E}} = \frac{\text{System Capacity (bits/s)}}{\text{Power Consumption (W)}}. \tag{6.10}$$

6.7.4 Diversity and Adaptive Antenna Techniques

Distributed antenna system (DAS) is playing an important role to provide higher data rate and mass wireless access service for broad area coverage. The purpose of indoor DAS is to split the communication cell to different areas by several remote units [SRR87]. Therefore, the line of sight scenario is more frequently presented if the DAS is employed so as to improve the coverage

[NFSS03]; meanwhile the DAS also increases the received diversity. The LTE Femtocell home BS (eNodeB) is a low-power cellular BS that uses licensed spectrum and is typically deployed in residential, enterprise, metropolitan hotspot, or rural settings. It provides an excellent user experience through enhanced coverage, performance, throughput and services based on location [ACD+12]. Combined with the DAS system [OZW10], the LTE femtocell BS can be installed indoors with a flexible configuration to provide the indoor users who need multimedia services with better user experience. The LTE femtocell DAS eNodeB can be divided into two types [Tia]: (i) a combined eNodeB, which employs a combined adaptive process hub unit (HU) to the system, and (ii) un-combined eNodeB, which does not adopt the combined adaptive process HU in the system. Once the eNodeB adopts the combined HU, more remote units can be allocated to the distributed system. Since more signals from different transmitted channels can be combined to input to the eNodeB by using the combined HU, the combined eNodeB actually improves the diversity of eNodeB without combined HU.

In the above context, authors in Tian et al. [YTB13] study the 3GPP LTE femtocell BS evaluation test-bed considering the LTE system working in band 13 defined by 3GPP, with a centre frequency of 782 MHz for the UL. The baseband LTE signal is sampled at 15.36 MHz, and an oversampling factor of four is used, giving a sampling frequency of 61.44 MHz. The number of input channels to the HU is the same as the number of distributed remote units. The FPGA unit receives CSI values and computes a set of weights to apply to the four received signals in order to produce a combined signal for the BS. The signals are initially sampled at low intermediate frequency by the on-board analog to digital converters. Furthermore, the performance of two maximum ratio combining (MRC) combined techniques, space–frequency and space-only methods, are measured and analysed in Tian et al. [YTB13]. For the case of employing the BS with combined HU, the space–frequency MRC combined algorithm always provides a better performance than the space-only combined algorithm. However, the space-only combined algorithm is simpler to implement and cheaper than the space–frequency method; therefore, the space-only method is expected to replace the space–frequency method in industry if the performance loss is small.