Chapter 2
Architectures and Interference Management for Small-Cell Networks

2.1 Requirements and Reference Model for Small-Cell Network Architectures

The successful deployment of heterogeneous small-cell networks relies upon how one can integrate small cells into the existing mobile access networks to provide seamless device-to-core network connectivity. Defined for hierarchical deployments with network elements installed in secure premises, the existing mobile network architectures in GSM, UMTS, cdma2000 and LTE standards cannot be trivially extended to include small cells. In the ad hoc small-cell deployments, it is also particularly challenging to gain access to the dedicated high-performance links for interconnection and proprietary management systems, as is the case in the current architectures. New network structures are therefore needed to support small-cell integration with the following minimum requirements [1].

- **Scalability**: Whilst the current mobile networks only allow some few hundreds of macrocells to connect to the next level of the hierarchy, it is expected that small cells are massively deployed with many thousands of units per one single network. This calls for an architecture that can support sufficient scalability within the same network.
- **Transparent integration**: Small cells should be easily and transparently integrated into the existing mobile networks. At the same time, the additional load on the legacy infrastructure should be kept to the minimum.
- **Security**: Deployed at end-user premises, small cells typically operate in an insecure environment. As such, any proposed small-cell network architecture must guarantee a sufficient level of security for both mobile networks and end users.
- **Limited backhaul capacity**: The new network architecture must take into account the fact that small cells connect with one another via shared broadband IP links with variable performance. This situation is very different from that in the existing mobile networks where dedicated interconnection links are available.
Since the implementation details of small-cell architectures can vary considerably, it is important to have a consistent design approach to promote compatibility. Towards this end, the Femto Forum has provided a reference architecture for small cells that includes all the network elements and interfaces. This generic reference is applicable to a vast majority of network architectures and it can be used to compare alternative approaches. Illustrated in Fig. 2.1, the main functional components in this reference architecture are described as follows [1–3].

**Small-cell access point**: At the customer premise, the key component is a low-power hardware device called small-cell access point (SAP). Mobile users located inside the premise communicate with the SAP over the radio links, and typically up to a dozen of which can be supported by one SAP. The SAP connects to the core network via a broadband access gateway, which can either be a stand-alone device or be integrated in the SAP. The air interface between the SAP and mobile users can be single-carrier (e.g., CDMA) or multi-carrier (e.g., OFDMA). The Fl interface is used by the SAP to control the operating parameters in the broadband access gateway.

**Broadband IP backhaul**: As home base stations are equipped with more powerful processing capabilities, the traditional network protocol has essentially collapsed. At the same time, the Internet Protocol (IP) rapidly replaces the hierarchical telecommunications-specific transport protocols. It is proposed that small cells use flat networks, i.e., Internet-like, as the backhaul to transport data from home devices to the core network. The reference architecture employs broadband IP access links (e.g., digital subscriber lines, cables, fiber to the home) as the backhaul.

**Small-cell gateway**: The direct connectivity between the core network and the SAP is maintained by the small-cell gateway (SCGW). Together with signaling protocol and channel conversions, SCGW aggregates and integrates traffic from a large number of small cells into the existing mobile networks. The SCGW
also implements security functions that authenticate and secure the connectivity with remote SAPs over the unsecured public broadband access links. The SCGW interfaces with the circuit-switch and packet-switch network segments of the mobile network operators (MNO) via Fb-cs and Fb-ps reference points, respectively. The SCGW–IMS network connectivity is supported by the Fb-ims interface.

With SCGW, the complexity and dimension of small-cell networks are hidden from the core network elements. It was earlier proposed that SAPs be kept simple and that all functions but radio be moved to SCGW. More recent solutions incline to support a flatter network by distributing much more functionalities to SAPs and keeping SCGW relatively simple. On the one side, SAPs support the front-end functions of Radio Network Controller (RNC), interact with end users, support mobility and perform radio resource management. On the other, SCGW supports back end RNC function, interfaces with core network and performs signalling aggregation. This approach allows for self-configured SAPs that support local services and local network access, enabling more cost-effective scalability.

**Small-cell management system:** Using Fm interface, the SAP management system (SAP-MS) can offer service provisioning and fault reporting of SAP devices. SAP-MS can handle tens of thousands of multi-vendor SAP units. Similarly, the SCGW management system (SCGW-MS) is expected to manage multiple SCGW devices via the Fg interface. The functions of SCGW-MS include traffic management, fault and alarm processing, and signaling protocol setting.

**Subscriber database:** The customer information such as SAP identity, network configurations and settings is stored in the subscriber databases. The SCGW accesses to these databases using the Fs and Fr interfaces.

### 2.2 Small-Cell Architectures in Wireless Network Standards

#### 2.2.1 3GPP UMTS Small-Cell Architecture

The 3GPP Universal Mobile Telecommunications Systems (UMTS) consist of a Core Network (CN) and a Universal Terrestrial Radio Access Network (UTRAN). In particular, the UTRAN has a hierarchical architecture comprising RNCs and Node Bs, and it is connected to the CN via the Iu interface. As shown in Fig. 2.2, UMTS architecture is consistent with the generic model given in Fig. 2.1 albeit with the following modifications [2, 3]:

- Mobile device is now termed user equipment (UE),
- Small-cell access point is called home node B (HNB),
- Small-cell gateway is now HNB gateway (HNB-GW),
- Security gateway function is separated from HNB-GW,
- Fa interface is replaced by Iu-h interface.
Deployed in the customer premise, the HNB is a low-power node that serves only one cell. The radio communication between the HNB and the UE is established via the Uu interface. In the core network, the HNB-GW plays the role of an RNC in that it concentrates multiple HNB connections on one side and connects to the MNO on the other side. While the connectivity between HNB-GW and HNBs is made possible with the Iuh interface, the HNB-GW employs Iu-cs and Iu-ps interfaces to connect with circuit-switch and packet-switch networks, respectively.

At the Iu-h reference point, a security gateway is deployed to protect the core network against security threats. Note that the security gateway can be implemented either as a separate physical element or be integrated to the HNB-GW. In this architecture, a new network element—HNB Management System (HMS)—is used to discover the HNB-GW, provide configuration data to HNBs, perform location verification of HNBs, etc.

It is worth noting the UMTS small-cell structure is able to offer architectural consistency. Since the HNB subsystem appears to core network as an existing Radio Network Subsystem (RNS), one can substantially reuse the existing network elements and protocols. At the same time, the HNB subsystem suffers from the existing limitations of RNS. Since a single HNB-GW can only address up to 65,535 unique HNBs, handover from the regular macrocell to HNBs is not supported due to the limited cell addresses. Although hard handover from HNBs to macrocell is possible as macrocells can be unambiguously identified using the Cell Global Identification, soft handover from and to an HNB is not yet supported.

### 2.2.2 3GPP LTE Small-Cell Architecture

Evolved-UTRAN (E-UTRAN) is an evolution of the 3GPP UMTS radio access technology, where Long Term Evolution (LTE) is the radio interface and Evolved
Packet Core (EPC) is defined to accommodate the high-speed LTE access. The E-UTRAN consists of multiple evolved Node Bs (eNBs), which connect with one another via the X2 interface to support handover and with the EPC via the S1 interface for traffic and control purposes. Each eNB connects to the mobility management entity (MME) via the S1-MME interface and to the Serving Gateway (S-GW) via the S1-U interface.

Figure 2.3 shows that small cells can be integrated into the LTE structure with consistency. Compared with the reference model in Fig. 2.1, the following new definitions are introduced [2, 3]:

- Small-cell access point is now termed Home evolved NodeB (HeNB),
- Small-cell gateway is called HeNB gateway (HeNB GW),
- Security gateway function is separated from HeNB GW,
- Small-cell management system is called HeNB management system (HEMS).

The functions supported by the HeNB are identical to those by the eNB in the UMTS case [see Fig. 2.2]. Similarly, the procedures that run between the HeNB and the EPC are the same as between the eNB and the EPC. In this architecture, the HeNB GW is used to allow the S1 interface between the HeNB and the EPC, thereby supporting a large number of HeNBs. While the HeNB GW appears to the MME as an eNB, the former appears to the HeNB as the MME. Therefore, a HeNB is architecturally indistinguishable from an eNB in EPC. The handover support from a HeNB to an eNB and vice versa is available, whereas that among the HeNBs is still under investigation.
2.2.3 3GPP2 CDMA2000 1x Small-Cell Architecture

The architecture for small-cell CDMA2000 1x deployment is shown in Fig. 2.4 [2, 3]. Different from the UMTS small-cell architecture, here the femto access point (FAP) includes a SIP user agent (SIP UA) to connect the 1x procedures on the mobile user side with the core network via a SIP/RTP interface. In this architecture, the femto security gateway (FSGW) maintains secure IP connectivity between the IMS core network and the FAP. On one side, an IPsec tunnel is established between the FSGW and the FAP via the Fx3 interface. On the other side, Fx1 interface transports RTP media packets to and from the FSGW, whereas the Fx2 interface implements the SIP signaling control.

The responsibility of the femto management system (FMS) includes configuring and managing the femtocell components via the newly-defined Fm interface. The femto AAA server authenticates the FAPs and shares security policy data with the FSGW. Using Fx4 interface, femto AAA server enables IPsec tunnels between the FAPs and the FSGW. Finally, the femto application server supports the interworking functions between the IMS network and the mobile carrier’s MAP network.

2.2.4 Air Interfaces: CDMA vs. OFDMA

CDMA is used for medium access in UMTS, CDMA2000 and high speed packet access (HSPA) wireless standards. In a CDMA system, UEs in all cells are allowed to simultaneously transmit over all available frequency bands (see Fig. 2.5a).
2.2 Small-Cell Architectures in Wireless Network Standards

Fig. 2.5 Radio resource sharing in CDMA and OFDMA. (a) CDMA: All users share the same frequency at the same time (b) OFDMA: One subchannel is given to at most one user at a time

These transmissions are differentiated by the use of orthogonal codes, i.e., spreading codes, assigned to individual UEs. At the transmitting side, user’s data signal is modulated with a spreading code to create a signal of a much larger bandwidth. At the receiving side, the cross-correlation of the received signal and the user’s spreading code is calculated. When the resulting cross-correlation reaches its maximum, the corresponding data signal can be extracted. Since increasing the number of CDMA users only raises the noise floor in a linear manner, the system performance gradually degrades for all users. Hence, there is no absolute limit on the number of users that can be accommodated by the system.

On the other hand, frequency-selective fading is one of the major impairments of wireless channels, particularly in multipath environments such as indoor and urban areas. Since the channel responses differ among different frequencies, it can be challenging to alleviate the distortion that broadband signals experience when transmitted over such channels. In this situation, orthogonal frequency-division multiplexing (OFDM) signals are preferred because they are more robust to this type of fading.

The basic idea of OFDM is to divide the transmitted bitstream into many sub-streams, to be sent over a large number of closely-spaced orthogonal subchannels. Each subchannel is represented by one subcarrier, and one substream of data is transmitted through one subcarrier. Since individual subcarriers are modulated with a conventional modulation scheme at a much lower symbol rate, each of the resulting narrowband signals experiences frequency-flat fading. The IEEE Wireless Interoperability for Microwave Access (WiMAX) standard uses OFDM in the physical layer, whereas the 3GPP LTE standard employs OFDMA in the downlink and single-carrier FDMA (SC-FDMA) in the uplink [4]. Different from CDMA where each UE occupies all the spectrum at all time, a UE in OFDMA systems is allowed to only use a subgroup of OFDM subchannels, as shown in Fig. 2.5b.
2.3 Interference Management in Small-Cell Networks

2.3.1 Interference Scenarios

In a small-cell heterogeneous network, the communication of two tiers of users results the following interference scenarios. The intra-tier interference situation is similar to what occurs in homogeneous networks, where a macrocell interferes with other macrocells and a femtocell interferes with other femtocells. However, due to the significant difference in the transmit power limits, the most severe interference happens in the cross-tier scenario as illustrated in Fig. 2.6. In Scenario A, a victim cell-edge MUE is strongly interfered by the downlink transmission of a nearby femtocell BS. In Scenario B, an MUE located far away from its serving macrocell BS transmits at a high power in the uplink to compensate the path loss. This transmission creates strong interference to a nearby victim femtocell BS.

The severity of cross-tier interference also depends on the way that the radio frequency is allocated. In the orthogonal frequency allocation, distinct sets of frequencies are assigned to small-cell users (or femtocell users) and regular users (or macrocell users). Although the cross-tier interference can be completely avoided in this way, the resulting spectral efficiency is low because the radio spectrum is not efficiently reused. In the partially shared spectrum allocation option, macrocells have full access to the overall spectrum while femtocells are permitted to share a
subset of such spectrum. To mitigate the strong cross-tier interference, some radio channels are specifically reserved to only macrocells in the form of escape frequencies.

The highest degree of freedom is available in the universally shared spectrum allocation strategy, where both femtocell and macrocell users in all cells are allowed to utilize the same frequency bands. Potentially offering the most efficient use of the limited radio resources, this approach is highly promoted for next-generation wireless networks, and thus it will be assumed throughout this brief. However, the increased cross-tier interference in this case calls for more sophisticated schemes to mitigate the adverse effects of interference, thus fully realizing the potential gains of universal frequency reuse.

It is noteworthy that while CDMA systems provide resistance to narrowband interference, this property does not occur with broadband interference such as signals from other users. These signals remain as broadband interference even after the despreading process. With a unity spectral reuse factor where all UEs (either within the same cells or from different cells) share the same frequencies, interference is a critical problem in small-cell networks that is based on CDMA. With OFDMA being the air interface, intracell interference among UEs within the same cell can be suppressed. This is due to the assumption of exclusive subchannel assignment, i.e., one subchannel is used by at most one UE at a particular time (see Fig. 2.5b). However, aggressive frequency reuse allows a common spectrum to be shared among the UEs belonging to different cells. While interference averaging helps reduce the effect of interference in CDMA, it does not happen in OFDMA systems. Here, one interfering transmitter is enough to completely jam a given subchannel. It therefore remains challenging to effectively manage the ICI in OFDMA-based small-cell networks.

The successful rollout of small-cell wireless networks depends upon how the interference challenges are addressed. Optimized for the carefully-planned homogeneous networks, conventional approaches prove inefficient in managing the random and severe interference in small-cell scenarios. The stringent requirement of protecting macrocell performance imposes a new set of design constraints that may as well invalidate any available solutions. Moreover, the limited capacity for control and signaling also renders centralized mechanisms, which require the exchange of global network information, impractical in many situations.

### 2.3.2 Power Control for CDMA-Based Wireless Networks

#### 2.3.2.1 Conventional Wireless Homogeneous Networks

Consider a CDMA-based multicell wireless homogeneous network. Let $p_i \geq 0$ be the transmit power of user $i$ and $\sigma_i$ be the power of the additive white Gaussian noise (AWGN). Denote the channel gain from the transmitter of user $i$ to its receiver as $h_{i,j}$, and that from the transmitter of user $j$ to the receiver of user $i \neq j$ as $h_{i,j}$. 
Note that the “transmitter of user $i$” in the downlink is the BS that serves UE $i$, whereas in the uplink it is UE $i$. The received SINR of user $i$ can be written as:

$$\gamma_i = \frac{h_{i,i} p_i}{\sum_{j \neq i} h_{i,j} p_j + \sigma_i}. \tag{2.1}$$

As can be seen from (2.1), a large unwanted signal power $\sum_{j \neq i} h_{i,j} p_j$ may significantly decrease the SINR of user $i$, thereby degrading the quality of radio communication. Figure 2.7 illustrates two typical interference scenarios. In the downlink, UE 1 in cell 1 receives not only the intended signal from its serving BS 1 but also interfering signals from BSs 2 and 3. In the uplink, the signal transmitted by UE 1 to its BS 1 is interfered by those from UEs 2 and 3 in the two adjacent cells.

Power control has been proven to be very effective in dealing with interference in CDMA-based wireless networks. The most popular power control solution is probably the Foschini-Miljanic’s algorithm [5], which enables users to eventually achieve their fixed target SINRs by iteratively adapting their transmit power according to:

$$p_i[t + 1] = \frac{\gamma_{i,\text{min}}}{\tilde{\gamma}_i[t]} p_i[t]. \tag{2.2}$$

Here, $\gamma_{i,\text{min}}$ is the target SINR of user $i$, whereas $p_i[t]$ is the transmit power and $\tilde{\gamma}_i[t]$ is the measured SINR at the receiver of user $i$ at time $t$.

It is worth noting that the simple algorithm in (2.2) can be implemented distributively by individual users, without requiring any form of network cooperation. As long as the target SINRs are feasible, (2.2) converges to a Pareto-optimal solution at a minimal aggregate transmit power $\sum_i p_i$. However, there is one major
drawback in the Foschini-Miljanic’s algorithm. If there exists an infeasible SINR target, the transmit power computed according to (2.2) will eventually diverge to infinity as each user $i$ always attempts to meet its own required SINR at any cost. To deal with infeasible SINR targets, admission control algorithms are introduced in [6, 7].

The works in [8–11] investigate several other power control schemes from a game-theoretical point of view. The solutions devised from noncooperative games are appealing since they can be implemented in a decentralized fashion. In these games, individual users selfishly optimize their own performance, regardless of the actions of other users. Denote the utility (or payoff) function of user $i$ as $U_i(p_i, p_{-i})$, where $p_{-i}$ is the power vector of all the users except $i$. The objective of each user $i$ in the power-control game can be formally expressed as:

$$\max_{p_i \geq 0} U_i(p_i, p_{-i}).$$

Depending on the type of utility function $U_i(\cdot)$, a number of games can be formulated whose solutions to the individual problem (2.3) exhibit different convergence properties. In most cases and under certain conditions, the underlying games settle at a Nash equilibrium (NE) $p^* = \{p_i^*\}$, a stable and predictable state at which no user has any incentive to unilaterally change its transmit power level, i.e.,

$$U_i(p_i^*, p_{-i}^*) \geq U_i(p_i, p_{-i}^*), \quad \forall p_i \geq 0, \forall i.$$  

(2.4)

Although the achieved NE gives a stable operating point, it is by no means guaranteed to be Pareto-efficient. To improve the efficiency of the equilibrium solutions, various pricing schemes are developed in [12, 13]. A pricing mechanism can implicitly enforce the cooperation among users while, at the same time, maintaining the noncooperative nature of the games. With pricing, the total utility of user $i$ is:

$$U_{tot,i}(p_i, p_{-i}) = U_i(p_i, p_{-i}) - C_i(p_i, p_{-i}),$$

(2.5)

where $C_i(\cdot)$ denotes the cost imposed to user $i$. In each problem $\max_{p_i \geq 0} U_{tot,i}$, various choices of utility and cost functions are available. Typically, the resulting solution is some modified version of the SINR balancing algorithm (2.2).

By selecting proper utilities and a linear cost $C_i(p_i, p_{-i}) = a_i p_i$, [14, 15] show that noncooperative games with pricing can substantially enhance the NE if small deviations from the target SINRs are allowed. For instance, with $U_i(\gamma_i) = -(\gamma_i - \gamma_i^{\min})^2$ the transmit power can be updated according to [15]:

$$p_i[t + 1] = \left( \frac{\gamma_i^{\min}}{\hat{\gamma}_i[t]} p_i[t] - a_i \frac{p_i^2[t]}{\hat{\gamma}_i^2[t]} \right)^+,$$

(2.6)

where $(\cdot)^+ = \max(\cdot, 0)$. Numerical results show that the enhanced Nash solution of [15] converges even faster than the SINR balancing algorithm in (2.2).
Still, it is unclear how far the Nash solutions given by [14, 15] are to the global optima of the power control problems. Using a different pricing scheme that is linearly proportional to SINR, i.e., \( C_i(\gamma_i) = a_i \gamma_i \), [16] proves that the outcome of a noncooperative power control game in single-cell systems is a unique and Pareto-efficient NE. By setting dynamic prices for individual users and assuming noise-like ICI, various design goals can be met. In multicell communications where transmit powers of all users need to be jointly optimized across all cells, ICI cannot be simply treated as noise. The solutions by [16] are thus limited to single-cell scenarios.

Different from [5] where feasible SINR targets must be given a priori, [17] considers a decentralized joint optimization of SINR assignment and power allocation that is Pareto-optimal for multicell systems. It is argued that a fixed SINR assignment is not suitable for data-service networks, where target SINRs should instead be flexibly adjusted to the extent that the system capacity can still support. A high SINR is translated into better throughput and reliability, whereas a low SINR implies reduced data rates. In [17], a feasible SINR region is characterized in terms of the loads at BSs and the potential interference from UEs. With a re-parametrization via left Perron-Frobenius eigenvectors and a locally computable ascent direction, distributed Pareto-optimal solutions are derived for the uplink case.

### 2.3.2.2 Small-Cell Heterogeneous Networks

The results in [17] apply to homogeneous networks, in which there exist no differentiated classes of users with distinct access priority and design specifications. However, it is unclear how the proposed solutions account for the complicated coupling and strong interdependency among users in multi-tier heterogeneous networks. In such cases, the choices of target SINRs available to the lower-tier FUEs are much more limited. Also, strict QoS guarantees need to be enforced for the prioritized MUEs, and radio resources have to be dedicated to meet the demands of these users.

In the context of heterogeneous small-cell networks, [18–21] study various beamforming techniques to mitigate the undue cross-tier interference. Joint admission control and power management has also been examined in [22] for cognitive-CDMA networks. To protect the existing MUEs while enabling a scalable femtocell deployment, [23] proposes an uplink power control scheme for FUEs. Using open-loop and closed-loop techniques, this scheme adjusts the maximum transmit power as a function of the cross-tier interference level. Based on the actual interference at the macrocell BS, the proposed scheme can suppress the cross-tier interference. However, the devised solution is neither distributed nor Pareto-optimal.

For CDMA-based wireless heterogeneous networks, power control games are formulated and analyzed by [24, 25]. In particular, [25] considers the interference scenario depicted in Fig. 2.8, where \( p_i \) denotes the transmit power of the BS that serves UE \( i \). Denoted as UE 0, the MUE is required to solve the following problem:

\[
\max_{0 \leq p_0 \leq P_{\text{max}}} U_0(p_0, \gamma_0|p_{-0}) = -(\gamma_0 - \gamma_0^{\text{min}})^2.
\]  

(2.7)
It is worth noting that the choice of utility function in (2.7) does not always guarantee the minimum SINR required by the MUE. Rather, only a “soft” SINR is provided. On the other hand, FUE $i$ is to solve the following individual problem:

$$
\max_{0 \leq p_i \leq P_{\text{max}}} U_i(p_i, \gamma_i | p_{-i}) = R(\gamma_i, \gamma_i^{\text{min}}) + \frac{C(p_i)}{I_i(p_{-i})},
$$

(2.8)

with reward function $R(\cdot) = 1 - \exp\left[-\bar{a}_i(\gamma_i - \gamma_i^{\text{min}})\right]$ and penalty function $C(\cdot) = -h_{0,i}p_i$. Here, $I_i(\cdot)$ is the interference power at the receiver of user $i$, and $\bar{a}_i, \bar{b}_i$ are constants. Note that because $C(\cdot)$ depends on the actual cross-tier interference $h_{0,i}$, explicit information about the cross-channel gains is required in the proposed algorithm. Due to the random fluctuations caused by shadowing and short-term fading effects, it can be quite challenging to estimate these channel values in practice.

### 2.3.3 Joint Subchannel-Power Allocation in OFDMA Networks

#### 2.3.3.1 Conventional Wireless Homogeneous Networks

Compared to CDMA, OFDMA—the multiuser version of OFDM—provides three dimensions of diversity, i.e., time, frequency and users, for a more efficient allocation of radio resources. As there are multiple subchannels available in OFDMA, the resource optimization in this case faces another major technical difficulty, i.e., the subchannel assignment that allots radio frequencies to different UEs in multiple cells. To solve this combinatorial problem alone, direct search methods usually require a prohibitive computational complexity. Radio resource
management for OFDMA-based networks relies upon efficient solutions that jointly optimize and assign powers and OFDM subchannels. Upon dividing the available spectrum into multiple subchannels, the SINR of UE \( k \) in cell \( m \) on subchannel \( n \) is expressed as:

\[
\gamma_{m,k}^{(n)} = \sum_{s \neq m} h_{m,k}^{(n)} p_{m}^{(n)} h_{s,k}^{(n)} + \sigma_k^{(n)} ,
\]

where \( p_{m}^{(n)} \) is the transmit power of BS \( m \) on subchannel \( n \), \( h_{m,k}^{(n)} \) the channel gain from BS \( m \) to UE \( k \) on subchannel \( n \), and \( \sigma_k^{(n)} \) the power of AWGN at the receiver of UE \( k \) on subchannel \( n \).

Using noncooperative game theory, [26] solves the competition for radio resources in a multicell OFDMA-based network. Assuming that the interference from other UEs is fixed, the solution to the pure noncooperative game for individual UEs is of an iterative waterfilling type. In this case, it may happen that some undesirable NE with low performance is obtained or, even worse, there exists no NE at all. Moreover, if the cochannel interference is severe on some subchannels, the NE may not be optimal for the entire system. Motivated by this observation, [26] introduces the concept of a “virtual referee.” By mandatorily changing of the game rules whenever needed, this referee can help improve the outcome of the formulated game. For example, it may reduce the transmit power of the UEs whose channel conditions are unfavorable. Those generating significant interference to other UEs may as well be prohibited from using certain subchannels. In doing so, the remaining cochannel UEs can share the corresponding subchannels in a more effective manner.

The study in [27] considers the problem of joint power allocation and subchannel assignment in the downlink of a multicell OFDMA network. Contrary to [26], the players in the formulated noncooperative game are the BSs, not the UEs. The players are responsible for allotting subchannels to the UEs within their cells, and deciding how much power to be distributed over those subchannels. Denote by \( p = [p_{m}^{(n)}]_{m,n} \geq 0 \) the network power vector that contains the transmit powers \( p_{m}^{(n)} \) for all BSs \( m \) and all subchannels \( n \). Also denote by \( \rho_{m} = [\rho_{m,k}^{(n)}]_{k,n} \) the channel assignment matrix of BS \( m \), where \( \rho_{m,k}^{(n)} = 1 \) if subchannel \( n \) is assigned to UE \( k \) in cell \( m \) and \( \rho_{m,k}^{(n)} = 0 \) otherwise. The utility function of BS \( m \) is defined as:

\[
U_m(p, \rho_m) = \sum_k \sum_n \rho_{m,k}^{(n)} \log \left( 1 + \frac{\sum_{s \neq m} p_{m}^{(n)} h_{m,k}^{(n)} h_{s,k}^{(n)} + \sigma_k^{(n)}}{\sum_{s \neq m} p_{s}^{(n)} h_{s,k}^{(n)} + \sigma_k^{(n)}} \right) - \alpha_m \sum_n p_{m}^{(n)} ,
\]
where $a_m > 0$ is the price per unit of power. Given a network power vector $p$, it is shown that BS $m$ assigns subchannel $n$ to UE $k^*$ if

$$k^* = k(m, n) = \arg\max_k \log \left(1 + \frac{p_{m}^{(n)} h_{m,k}^{(n)}}{\sum_{s \neq m} p_{s}^{(n)} h_{s,k}^{(n)} + \sigma_k^{(n)}}\right). \quad (2.11)$$

Certainly, $\rho_{m,k^*}^{(n)}(p) = 1$ in this case.

Once a fixed optimal subchannel assignment $\rho_m^*$ is found, the optimal power allocation is derived as:

$$p_m^{(n)} = \left(1 + \frac{\sum_{s \neq m} p_{s}^{(n)} h_{s,k^*}^{(n)} + \sigma_k^{(n)} h_{m,k^*}^{(n)}}{a_m + \lambda_m}\right)^{+}, \quad (2.12)$$

where $\lambda_m \left(\sum_n p_m^{(n)} - P_{\text{max}}\right) = 0$ with $\lambda_m \geq 0$ being the Lagrange multiplier for the maximum total power constraint $P_{\text{max}}$ at BS $m$. The allocations in (2.11) and (2.12) are performed iteratively until an equilibrium is finally reached. As proven in [27], such an iterative algorithm is guaranteed to converge to a unique NE under certain conditions. Usually, the stable operating points provided by the game-theoretical solutions do not globally maximize the network sum rates.

Different from [28–35] where the radio resources are allocated in a heuristic manner, [36] takes an optimization approach to solve the following problem of coordinated scheduling and power allocation in multicell OFDMA-based networks:

$$\max_{p; k = [k(m,n)]_{m,n}} \sum_m \sum_n w_{k(m,n)} r_{m,k(m,n)}^{(n)} \quad \text{s.t.} \sum_{k} p^0_{k(m,n)} \leq P_{\text{max}}. \quad (2.13)$$

Here, weight $w_{k(m,n)} \geq 0$ accounts for the priority of UE $k(m,n)$, $p^0_{k(m,n)} \geq 0$ is the transmit power vector of all UEs on subchannel $n$, and $r_{m,k(m,n)}^{(n)} = \log \left(1 + \gamma_{m,k(m,n)}^{(n)}(p^{(n)})\right)$ is the corresponding throughput. The first proposed scheme—a multicarrier extension of the SCALE algorithm [37]—is proven to converge to a solution that satisfies the necessary optimality conditions of the nonconvex combinatorial problem (2.13). Using Lagrangian duality, the second scheme provides an optimal solution if the number of OFDM subchannels is very large [38, 39]. The third scheme is an improved iterative waterfilling algorithm, adapted to this multicell scenario. It is noted that all the solutions developed in [36] depend on a central unit to collect and process the complete channel state.
information. To alleviate the high complexity required by such solutions, [40] proposes a distributed low-complexity scheme based on the concept of a “reference user” to solve (2.13).

Considering the downlink of an OFDMA network, [41] addresses the problem of maximizing the weighted sum of the minimal UE rates of coordinated cells. In this case, the objective in (2.13) is modified as:

\[
\sum_{m=0}^{\infty} \frac{w_{m}}{w_{m}} \min_{k \in \mathcal{X}_{m}} \sum_{n}r_{m,n}(n),
\]

(2.14)

where \(w_{m} \geq 0\) denotes the weight assigned to the smallest UE rate of cell \(m\), and \(\mathcal{X}_{m}\) the set of all UEs belonging to cell \(m\). Similar to [27], the centralized algorithm proposed by [41] alternatively optimizes the subchannel assignment and power allocation so that (2.14) keeps increasing until convergence. At each iteration, the allotment of subchannels is updated by resolving a mixed integer linear program for each cell. The optimal allocation of powers is found by a duality-based numerical algorithm. However, if a minimum rate constraint is strictly imposed to guarantee the QoS of some certain UE, the solutions in [36, 40, 41] are no longer applicable.

2.3.3.2 Small-Cell Heterogeneous Networks and Cognitive Femtocells

A joint subchannel and binary power allocation algorithm is developed in [42], where only one transmitter is allowed to send signals on each subchannel. Based on Lagrangian dual relaxation, [43, 44] propose various joint subchannel and power allocation schemes for OFDMA femtocells. It is assumed that the intra-tier inter-femtocell interference is negligible, whereas the cross-tier interference from the macrocell to femtocells is a constant. While these assumptions remarkably simplify the analysis, they are often not the case in practice. Moreover, network optimization for the existing macrocell is not considered at all in [43, 44].

In [45], the joint allocation of radio resource blocks and transmit powers is investigated for the downlink of OFDMA-based femtocells. The formulated exact-potential game is shown to always converge to an NE when the best-response adaptive strategy is applied [46]. Also taking a game-theoretical approach, [47] models macrocell BSs and femtocell BSs as the leaders and followers in a Stackelberg game [46]. In the hierarchical competition, a Stackelberg equilibrium, whose performance is better than that of an NE, is proven to exist under some mild conditions. As previously discussed, there is an ultimate need to protect the preferential MUEs in a mixed macrocell/femtocell network. This critical issue, however, has not been adequately addressed in [45, 47].

On the other hand, it has been confirmed that much of the licensed radio spectrum remains idle at any given time and location [48]. Spectrum utilization can thus be significantly improved by allowing (unlicensed) secondary users (SUs) to access spectrum holes unoccupied by (licensed) primary users (PUs). Cognitive radio [49–51] is promoted as an efficient technology to exploit the existence of spectrum
portions unoccupied by PUs. While PUs still have a priority access to the radio spectrum, SUs are permitted to have a restricted access, subject to a constrained degradation on the PUs’ QoS.

Spectrum pooling is an opportunistic access approach that enables public access to the already licensed frequency bands [52, 53]. The basic idea is to merge spectral ranges from different spectrum owners into a common pool, from which SUs may temporarily rent spectral resources during the idle periods of PUs. Here, the licensed system does not change while SUs access unused radio resources. In spectrum-pooling radio systems, OFDM is recognized as a highly promising candidate for SU transmission. This is mainly because of its flexibility in dynamically allocating the unused frequencies among SUs, and its ability to monitor PU spectral activities at no extra cost. However, OFDM transmission may cause mutual interference between PUs and SUs, due to the non-orthogonality of the respective signals [54, 55].

Several recent works propose that cognitive radio (CR) be used in heterogeneous small-cell networks, in that cognitive FUEs are allowed to opportunistically access the radio spectrum licensed to MUEs [56–58]. The roles of MUEs and FUEs in macrocell/femtocell settings correspond to those of PUs and SUs in CR networks, respectively. The existing results on radio resource management for OFDM-based CR networks can thus be applicable to two-tier cognitive femtocell networks.

In [59], an optimal power allocation scheme is devised to maximize the downlink capacity of a single SU, while guaranteeing that the interference induced to the PU is below a specified threshold. Similarly, [60] aims to maximize the CR link capacity, taking into account the availability of OFDM subchannels and the total interference limits at PUs. Extending the results in [59, 60] to multiuser scenarios, [61] aims at maximizing the discrete sum rate of a secondary network, constrained on the interference imposed to PU frequency bands. Subject to the per-subchannel power constraints (due to PU interference limits), [62] proposes a partitioned iterative water-filling algorithm that enhances the capacity of an OFDM CR system.

Zhang and Leung [63] attempts to solve the problem of resource allocation in multiuser OFDM-based CR systems. The main objective of [63] is to provide a satisfactory QoS to both real-time and non-real-time applications, despite the rapid variations in the available resources caused by the PUs’ activities. In [64], the issue of downlink channel assignment and power control for FDMA-based cognitive networks has also been addressed, where BSs make opportunistic spectrum access to serve fixed-location UEs within their cells. Suboptimal schemes are derived to maximize the total number of supportable UEs, while guaranteeing the minimum SINR requirements of SUs and protecting the PUs.

To deal with the combinatorial OFDM subchannel assignment problem, the Lagrangian dual framework in [38] has proven to be especially useful. Considering networks with the coexistence of multiple primary and secondary links through OFDMA-based air-interface, [65] utilizes such an optimization framework to develop centralized and distributed algorithms. The design goal of [65] is to improve the total achievable sum rate of secondary networks, subject to interference constraints specified at PUs’ receivers. Also based on Lagrangian duality, [66] studies the coexistence and optimization of a multicell CR network overlaid with
a multicell primary network. The weighted sum rate of SUs over multiple cells is maximized in this case. For the downlink of a spectrum underlay OFDMA-based CR network, [67] proposes a joint subchannel-power allocation scheme that maximizes the CR network capacity. Here, the ICI among different CR cells is also controlled. With Lagrangian duality, the primal problem is decomposed into multiple dual subproblems, each of which is solved by an efficient algorithm. For Lagrangian dual framework to apply, the “frequency-sharing” condition must be strictly satisfied [38, 68].

References


