6

Energy-Efficient Mobile Network Design and Planning

Yinan Qi, Muhammad Ali Imran and Rahim Tafazolli
Institute for Communication Systems (ICS), University of Surrey, Guildford, Surrey, UK

6.1 Introduction

The exponential increase of user data often requires planning and design of a data communication system and the operation and management of such a system. The latter issues are discussed in the next several chapters and this chapter mainly deals with planning and design of a wireless network. Network design and planning consists of three steps: network topology design, network synthesis and network realization [1–3]. In the first step, the locations and connecting patterns of the network nodes are determined. In the second step, the network parameters, such as size, are optimized subject to certain quality of service (QoS) requirements. In the last step, the network is realized with designed configuration parameters. These steps might be conducted in an iterative manner to gradually optimize the network. For a wireless network, network design and planning principles can be applied to both the core network and the radio access network (RAN). The main focus of this chapter is on the RAN. In the development of today's second and third generation cellular mobile radio networks, such as GSM and WCDMA, the main optimization objectives in network design and planning are focused on providing optimum throughput subject to coverage requirements and a required level of QoS. The "Green" aspect does not draw much attention and therefore is rarely addressed. However, the rapid evolution of telecoms industry is accompanied by a huge increase of energy consumption of mobile networks in recent years [4, 5]. Coupled with sharp rising cost of the energy resources in the past few years, the operators are forced to further exploit the potential of network design and planning to reduce both capital expenditure (CAPEX) and operational expenditure (OPEX), mainly arising from the energy consumption. In the efficient design and deployment of the fourth-generation and beyond fourth-generation cellular networks, for example 3GPP LTE/LTE-Advanced, it is proposed by industry consortium that the

Green Communications: Principles, Concepts and Practice, First Edition.
Edited by Konstantinos Samdanis, Peter Rost, Andreas Maeder, Michela Meo and Christos Verikoukis.

energy consumption will be taken into consideration as an additional optimization target for all the operators [6].

The aim of this chapter is to draw a fundamental and comprehensive picture of the state-of-the-art "green" network design and planning schemes, algorithms and methodologies, mainly focused on the first two steps aforementioned. The organization of the chapter is as follows: the potential of cell parameter optimization to achieve energy saving is introduced in the next section. Sections 6.3 and 6.4 discuss the most promising green solutions for urban and rural scenarios, respectively. The last section concludes this chapter and provides some insight into the future network architecture evolution path.

6.2 Deployment: Optimization of Cell Size

Studies have identified that the increasing energy consumption is mostly related with the higher BS site density to provide huge amount of data transmission capability and high level of coverage [7]. The transmission power can be significantly reduced with a smaller cell size due to the fact that the transmitted signals suffer lower path loss in the air. Meanwhile, considering the transmission power can merely be counted as a part of overall energy expenditure, smaller cells with higher cell deployment density clearly cause additional static power consumption in terms of circuitry power, site cooling power, and so on, which are almost irrelevant with transmission power. The saving potentials should be carefully designed to balance the trade-off between throughput and energy expenditure. Another important consideration when planning the network is the traffic, in particular, the spatial traffic variations. In a high traffic area, it is expected to deploy more cells to allocate more resources for satisfactory user experience. Otherwise, smaller number of large cells can be deployed to meet the traffic requirement. Some of the previous works [8, 9] have made a common assumption that the spatial traffic distribution is uniform to guarantee that the peak traffic can be handled with the required QoS throughout the entire region covered by the cellular system. However, the spatial traffic distribution is not uniform in realistic scenarios. The uniform traffic assumption causes inefficiency because the number of BSs deployed in the low traffic area is more than needed. Hence, a part of the energy consumed by those BSs is wasted. When designing the cellular network parameters, all the aforementioned aspects should be considered to offer realistic solutions.

6.2.1 System Model

Considering a hexagon cell with radius R as shown in Figure 6.1, the area of the cell is denoted as S and we assume that the entire cellular system is able to cover a large region with area U, where the spatial traffic density varies according to the locations. This model is commonly used in system level analysis for cellular networks.

6.2.1.1 Traffic Model Within a Cell

The multi-class MMPP/M/1/D-PS queue model (a single server processor sharing queue, with Markov-modulated Poisson arrival process, Markovian service time and finite capacity) has been widely used to study and dimension the telecommunication systems for more than a

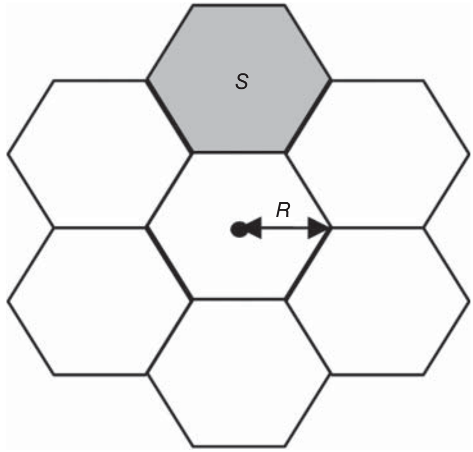

Figure 6.1 Cell definition

century [10, 11]. Surprisingly, the simplicity of the model does not impair the evaluation accuracy. The assumptions made in this model are as follows:

- No more than S sessions and D data connections can be admitted in a cell at the same time. The session arrival rate is λ_S and with s sessions and the data connection arrival rate is $s\lambda_D$.
- The user service time is exponentially distributed with mean value T; hence, the user service rate is exponentially distributed with mean value $\mu_s = 1/T$.
- In each data connection, the amount of information transferred is exponentially distributed with mean value G; hence, the data connection service time is exponentially distributed with mean value $\mu_d = B/G$, where B is the throughput.

In this traffic model, one of the parameters, that is, session arrival rate, is controlled by a Poisson process. It is typical in modelling the traffic types where time-varying arrival rates capture some of the important correlations between inter-arrival times and is applicable in services such as VoIP and FTP. However, this traffic model is not suitable to model the continuous traffic types, for example, video streaming.

6.2.1.2 Spatial Traffic Variation Model

Realistic traffic is envisaged to vary according to the locations, which implies that the session generating rate λ_S as well as the data connection generating rate λ_D are functions of the BS locations and can be expressed as $\lambda(r, \theta)$ in a polar coordinate system. For simplicity of the analysis, we assume that the spatial traffic is a step function as shown in Figure 6.2. The realistic traffic model might be far more complicated than this simple model. However, any realistic traffic model with traffic hotspots can be decoupled into similar step functions. Hence, it is adequate for the principle estimation of the energy savings and the optimization of network design and the planning parameters can be easily extended from this simple model. Note that

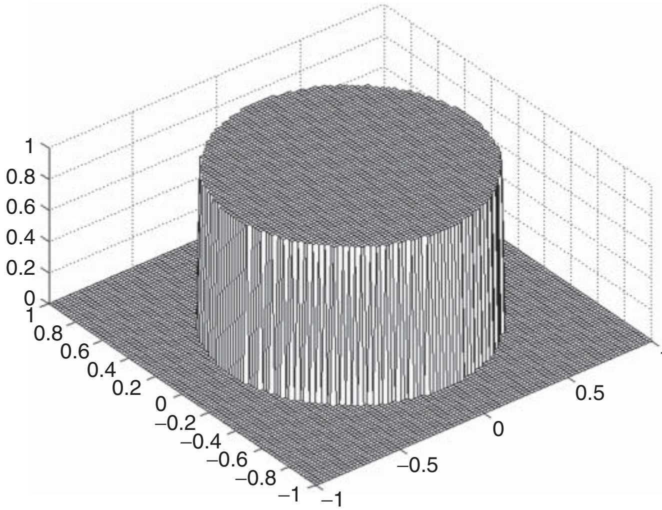

Figure 6.2 Normalized spatial traffic variation model

this model is a peak traffic model, which implies that although the traffic varies with time, it will never exceed the values in this model, guaranteeing required QoS at any time.

6.2.1.3 Propagation Model and Coverage

Normally, the received signal suffers from path loss, slow fading and fast fading. A simple model taking path loss and slow fading into consideration is given as:

$$P_{\mathrm{rx}}(r) = K\left(\frac{r}{r_{\mathrm{ref}}}\right)^{-\gamma} \varphi P_{\mathrm{tx}} \tag{6.1}$$

where P_{tx}, P_{rx}, r_{ref}, γ and φ represent transmit and receive power, reference distance and propagation path loss exponent and a random variable modelling shadowing effects, respectively, and K is a constant which can be obtained through measurement and its value is given in Ref. [12].

Based on this propagation model, coverage, as another critical performance metric is defined as the fraction of cell area where the received power is able certain threshold P_{min}:

$$C = \frac{1}{S}\int_S r\,\mathrm{Pr}(P_{\mathrm{rx}}(r) \geq P_{\mathrm{min}})drd\theta \tag{6.2}$$

6.2.1.4 Quality of Service (QoS)

The QoS can be referred to several related metrics such as latency, jitter, packet dropping probability, and so on. In this chapter, QoS is mainly determined by the blocking probability that a new generated request is rejected because there are no available resources to be allocated to the new user. According to Ref. [13], the blocking probability p_{blk} can be easily calculated by applying the infinitesimal generator matrix based on the steady state probability as:

$$p_{\text{blk}} = \frac{\sum_{d=0}^{D} \pi_{d,S}\lambda_s + \sum_{s=0}^{S} \pi_{D,s} s\lambda_d}{\sum_{s=0}^{S}\sum_{d=0}^{D} \pi_{d,s}(s\lambda_d + \lambda_s)} \tag{6.3}$$

where $\pi_{d,\, s}$ is the steady state probability for state $(d,\, s)$ which indicates that the dth data connection request of the sth session arrives. The numerator represents the probability that the incoming session/data connection requests are rejected and the denominator stands for the summarised probability of all steady states.

6.2.2 *Optimization of Cell Parameters*

In the previous deployment strategies, the cell size is designed to satisfy the maximal traffic in the entire network, which is obviously not an energy-efficient solution. The basic idea of the new deployment strategy is to separate the entire service region according to the traffic level. For the normalized spatial traffic variation depicted in Figure 6.2, we divide the covered area into two parts: a dense zone where the traffic is intensive and a sparse zone where the traffic is relatively low. In the dense zone, the target blocking probability must be met, that is, $p_{\text{blk}} \le p_k^T$. However, in the sparse zone, if we maintain the same cell size, the blocking probability achieved might be significantly smaller than the target one due to the lower traffic demand, thus wasting a large amount of energy. Therefore, we can enlarge the cell size within the sparse zone so that the BSs receive more traffic requests but the coverage and QoS requirements will not be impaired. By doing this, the number of BSs used in the entire network is reduced as well as the overall consumed energy.

Within the dense zone, the radius of the cell is assumed to be R_d and the area of each cell is S_d; whereas in the sparse zone, the radius of the cell is $R_s = \mu R_d$ and the area $S_s = \mu^2 S_d$. A heuristic scheme can be employed to optimize the cell sizes in two zones with different traffic demands:

- **Step1**: For the dense zone, determine $R_{d,\,\text{BP}}$ according to the blocking probability requirement.
- **Step2**: Calculate the average number of data connections within a cell and $R_{h,\,\text{C}}$ according to the coverage requirement.
- **Step3**: Choose $R_d = \min\,(R_{d,\,\text{BP}}, R_{d,\,C})$.
- **Step4**: Choose a proper $R_s = \mu_0 R_d$ to satisfy the target blocking probability in the sparse zone.
- **Step5**: Calculate coverage of the cells within the sparse zone.
- **Step6**: If the coverage requirement is met, the algorithm is ended; else reduce $\mu_{n+1} = \mu_n - \Delta\mu$ and goes back to step 5.

In order to validate the feasibility of this scheme, we set up a simulation scenario with LTE link budget and use the FTP service model in Ref. [14]. The target blocking probability is set as 1% and we assume the entire covered area is a circle with radius 100 km and the dense zone is a circle with radius 20 km. The new scheme will be compared with conventional scheme where no separation is conducted and the network planning is merely based on the maximum

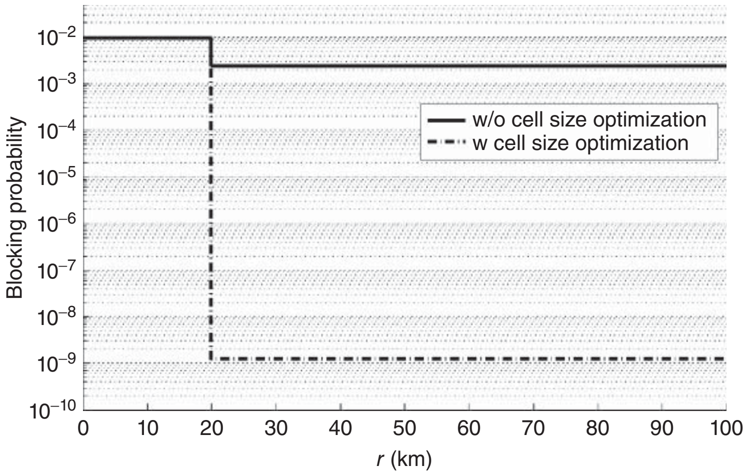

Figure 6.3 Blocking probability

traffic, tailoring to a uniform cell size. We define the energy saving power ratio Φ_{sav} to exploit the energy saving as:

$$\Phi_{sav} = \frac{P_{new}}{P_{uni}} \tag{6.4}$$

where P_{uni} is the overall power consumption for the uniform cell size.

We use the aforementioned scheme to find the optimal $\mu = 1.4$ and the corresponding $\Phi_{sav} = 0.82$, which means 18% energy consumption saving can be achieved because less number of BSs are deployed in the low traffic zone. The blocking probability is shown in Figure 6.3. If the cell size is unchanged in the sparse zone, the blocking probability sharply decreases due to the very low traffic demand. On the contrary, if the cell size is enlarged, more traffic is generated within one cell and the achieved blocking probability is closer to the target.

Now we come to the conclusion that by separating the whole network into two zones, the cell size can be optimized based on the spatial traffic distribution. With smaller number of BSs deployed in the sparse zone, the overall energy consumption is successfully reduced. Actually, the whole area can be divided into more zones according to the spatial traffic variation. In such a case, the network planning strategy is able to adapt itself to the spatial traffic variation in a more efficiently manner.

6.3 Network Design and Planning for Urban Areas

The planning of a network cannot be easily modified once the rollout is complete. It is difficult to relocate the already deployed base station sites of a legacy network. However, it does not mean that the network is not able to adapt itself to the change of the operation conditions. In one aspect, the network layout does not depend on the deployed but on the active BSs. Hence, it can be dynamically changed by adaptively switching on/off some BSs under different traffic circumstances. In another aspect, the densification of the current network can be tailored with additional macro BSs and/or micro BSs to reduce the inter-site distance (ISD) when the required traffic demand is growing. In the latter case, it is more important to leverage the deployment options for savings of energy by dynamic adaptation when the already deployed and the additionally deployed base station sites are fixed.

6.3.1 Adaptive On/Off Strategies to Change the Network Layout

The daily traffic profile has revealed that there are long periods in which the traffic load is low and the BSs are unnecessarily activated [14]. Different on/off schemes have been investigated to save energy, where the number of active BSs are adjusted dynamically adapting to the real-time traffic load. Therefore, unnecessary network nodes are switched off and the network layout is changed to save energy. It should be noted that the required QoS should not be compromised in the procedure of minimization of energy expenditure.

The on/off scheme can be simply categorized into two approaches: static and dynamic [15]. While the static approach only applies on/off scheme for a fixed potion of all network nodes in the period of one day, the dynamic approach is more adaptive in the sense that the on/off operations can be applied to the BSs depending on the traffic variations. With a simplified linear daily traffic variation pattern assumed as shown in Figure 6.4(a), the energy savings of two approaches are demonstrated in Figure 6.4(b) [15].

Two static on/off schemes, denoted as ½ and ¼, are demonstrated, where $1/x$ means that 1 out of x BSs are activated during the off operation period. As shown in the figure, the dynamic approach benefits from more flexibility and stronger adaptation capability and achieves higher percentage of energy savings. In addition, the presented results also clearly deliver an important message: the achieved energy saving is rather depending on the variation of the traffic, that is, the slope from peak to off-peak d, than one the absolute values of the traffic, that is, the peak traffic and the off-peak traffic.

6.3.2 Adaptive (De)sectorization

It has been pointed out that switching off the entire BS site might not be realistic because the transmission power of the BSs is assumed to be adjustable in a large dynamic range to meet the coverage requirement [16], which is normally impractical in real systems because of the power amplifier and RF link constraints. A more effective approach is adaptive (de)sectorization, where the BSs are able to adapt themselves to the temporal traffic variation by switching off some sectors and changing the beamwidth of the remaining sectors.

As shown in Figure 6.5, each BS consists of three sectors and each sector is defined as a hexagon with radius R. Sectorized antenna is employed in each sector.

Here, we use another commonly used metric of interest G-factor to evaluate coverage:

$$G_{\text{factor}} = \frac{P(BS_i \rightarrow UE)}{\sum_{j \neq i} P(BS_j \rightarrow UE) + P_{\text{therm}}} \tag{6.5}$$

where $P(BS_k \rightarrow UE)$ is the received power from BS k to the user equipment (UE), given in mW and P_{therm} is the thermal noise power given in mW. The long-term coverage can defined as:

$$C = \frac{1}{S_a} \int_{S_a} \Pr\{G_{\text{factor}}(x,y) \geq G_{\text{factor,min}}\} dxdy \tag{6.6}$$

where S_a is the area of a sector.

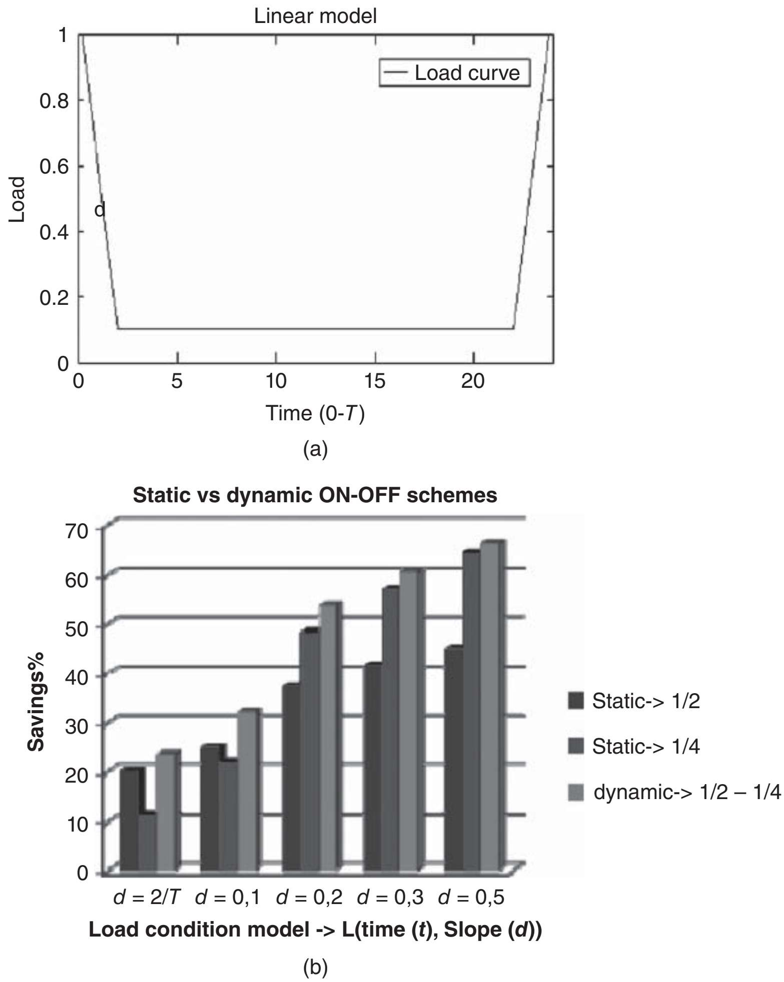

Figure 6.4 Energy savings and daily traffic profile (a) daily traffic profile (b) energy savings

When traffic demand is at high level, all three sectors within one BS are activated. When traffic demand is lower, one sector out of three sectors is switched off. Here for simplicity, we assume sector 1 is switched off. Handover is performed throughout the system. The users in the silent sector, that is sector 1, will be assigned to other sectors or even other BSs based on the long-term received signal strength, that is summation of path loss and shadowing loss. Apparently, those users in the silent sectors will suffer from lower coverage because of the weaker receive power. To maintain the coverage level of the silent sector, we need to change the antenna pattern of the other two remaining sectors.

Considering the 3D antenna pattern in Ref. [17], given by horizontal (azimuth) and vertical (elevation) orientations, the azimuth antenna pattern used in LTE system is given by:

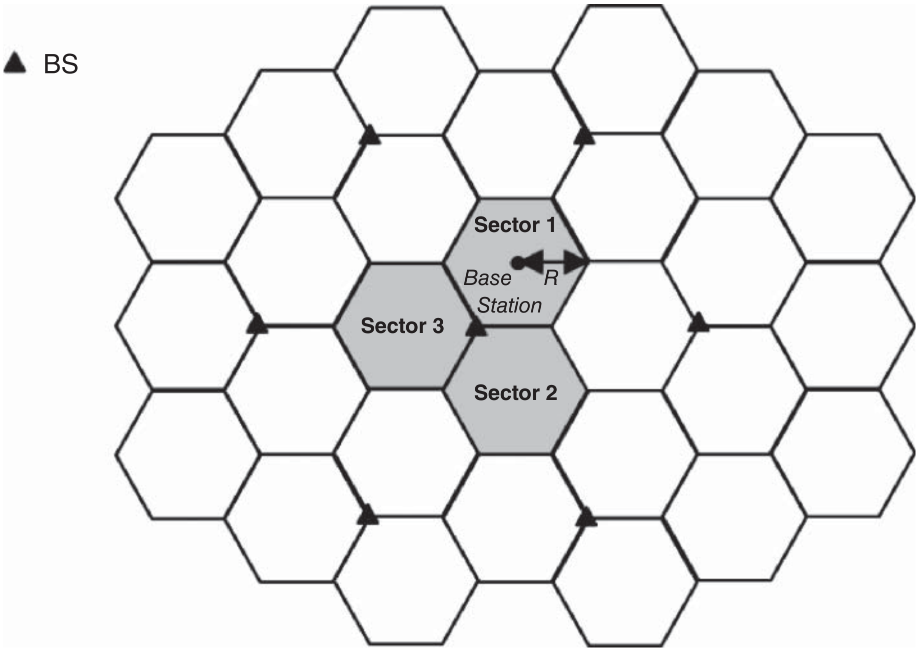

Figure 6.5 BS deployment

$$G_{az}(\Theta) = G_{max} - \min\left\{ 12\left(\frac{\Theta}{\Theta_{3\,dB}}\right)^2, G_{f2b} \right\}$$
$$G_{max} = 14dBi, \Theta_{3\,dB} = 65^\circ, G_{f2b} = 20\,dB$$
$$-180^\circ \leq \Theta \leq 180^\circ \tag{6.7}$$

where G_{max} is the boresight antenna gain, Θ is the angle between the sector and the mobile (BS-UE) line of sight and the sector boresight, $\Theta_{3\,dB}$ is the 3 dB angle, also defined as beamwidth and G_{f2b} is the antenna front to back ratio. The elevation pattern is:

$$G_{el}(\Phi) = G_{max} - \min\left\{ 12\left(\frac{\Phi}{\Delta_\Phi}\right)^2, G_{f2b} \right\}$$
$$\Delta_\Phi = 10^\circ, -180^\circ \leq \Phi \leq 180^\circ \tag{6.8}$$

where Δ_Φ is the elevation width and Φ is the downtilting angle. This downtilting angle is decided by the antenna height H_{ant} (typical value is 20–40 m, here is 20 m) and the distance D between the BS and the crossing point of the horizon and the antenna main lope orientation, that is $\Phi = \text{argtan}(H_{ant}/D)$. It is a key factor to the intra- and inter-cell interferences. It should be optimized to minimize the interference and maximize the long-term throughput, which

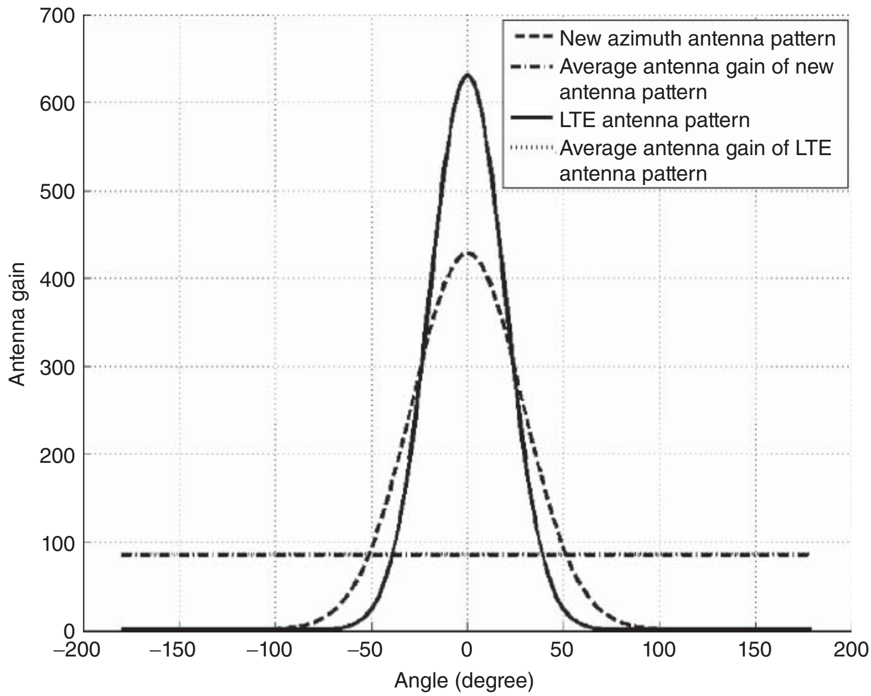

Figure 6.6 Azimuth antenna patterns

is a function of G_{factor}. For the sake of simplicity, we assume the downtilting angle keeps unchanged after sector 1 is switched off.

Even with optimized downtilting angle, turning off one sector will cause coverage problem. In order to cover the silent sector, it is natural to change the direction of other sections to offer a satisfactory coverage. For example, sector 3 antenna can be redirected from 180° to 120° and the beamwidth of two remaining sectors can be expanded from 65° to 95°. After adjustment, the new azimuth antenna pattern is shown in Figure 6.6.

In Ref. [18], a well-defined Macro cell power model based on measurement is presented, where the power consumption in watts is estimated individually for each subsystem. One BS consists of power amplifier (PA), main supply, DC part, RF link, Base band and cooling equipment. The power consumption of each component is shown in Figure 6.7. If one sector is switched off, the energy saving is not straightforward 1/3. PA and RF link consumption can be reduced by approximately 1/3. The other power consumption associated with DC, BB processing and cooling will not be significantly reduced. In a nutshell, the overall power consumption can be reduced by approximately 21%.

As we mentioned before, when minimizing the energy consumption the QoS should not be impaired. Turning off one sector not only brings the coverage problem, but it also increases the traffic demand of the remaining active sectors due to the new users handed over to them and might increase the blocking probability. The same MMPP/M/1/D-PS queue model aforementioned in the cell size optimization section can also be used here to analyze the blocking probability.

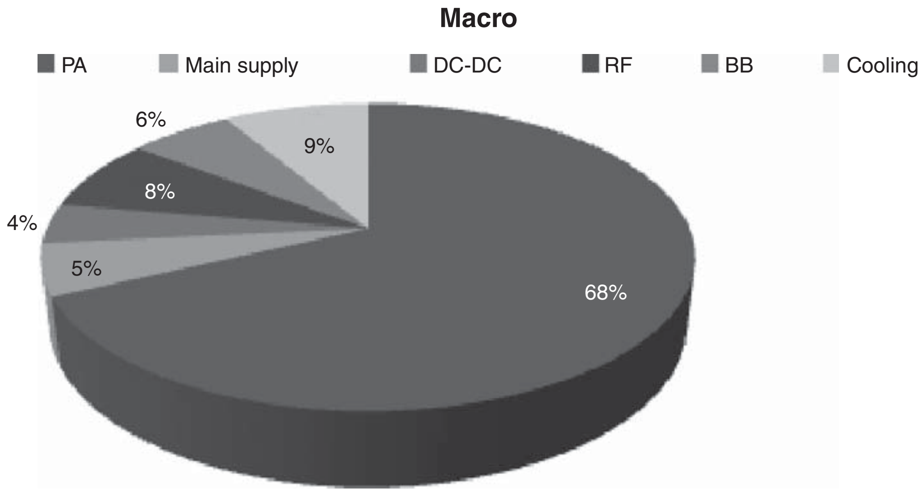

Figure 6.7 Power consumption breakdown

How the energy efficiency of an LTE system can be impacted by adaptive (de)sectorization is depicted in Figures 6.8–6.10 with following assumptions:

- Two tiers of BSs (19 BSs and 57 sectors) with central one as the reference one and inter-cell distance is 500m.
- UEs are placed randomly following a uniform distribution. One sector allows up to 50 UEs and 20 data connections.
- We assume OFDMA system without any cooperation among sites or any fractional frequency reuse patterns.
- LTE based link budget and propagation model.
- There are two main QoS constraints: the target blocking probability is 0.01 and the target coverage is 99%.

The G_{factor} map before and after switching off sector 1 is shown in Figure 6.8.

When sector 1 is switched off, the steering direction of sector 3 is changed and the beamwidth of the remaining two sectors is increased. The direct consequence is that one sector is receiving more interference because of the larger beamwidth, leading to G_{factor} degradation as well as the long-term system throughput. Figure 6.9 depicts the cumulative distribution function (CDF) of G_{factor} and it is indicated that the average value of G_{factor} is smaller when sector 1 is switched off. An interesting observation is that some users (the part in the circle) might actually benefit from this adaptive sectorization. Those users are previously strongly interfered by sector 1. Since sector 1 is off now, the interference is reduced and their performance is improved.

It should be noted that when traffic is further decreasing, we can switch off two sectors out of three sectors and change the antenna pattern to the omnidirectional to maintain the coverage. It will be shown in the following results that switching off two out of three sectors is beneficial when the traffic load is extremely low. The user arrival rate per km^2 is defined in Ref. [18]

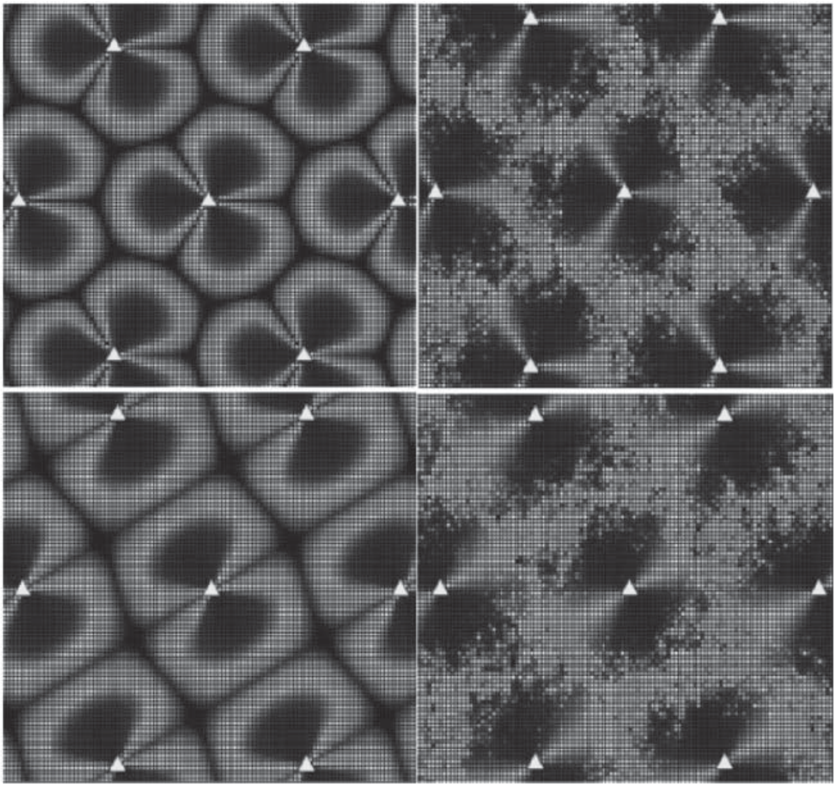

Figure 6.8 G_{factor} map (upper left: 3 sectors, pathloss only; upper right: 3 sectors, pathloss + shadowing; lower left: sector 1 is switched off, pathloss only; lower right: sector 1 is switched off, pathloss + shadowing)

and changes in 24 hours as shown in the upper part of Figure 6.10(a). Note that since the data connection arrival rate is constant, the temporal traffic variation is purely demonstrated by the fluctuation of the user arrival rate per km^2. The data connection arrival rate is set as 10 per UE per second per km^2. When the traffic demand in terms of user arrival rate is low, one sector is switched off to save energy. The coverage area of the remaining two sectors is then increased to 1.5 times of its original size, which means that the user arrival rate for each active sector is 1.5 times of its original value now. In the mean time, the average system throughput is degraded as shown in Figure 6.8 (the amount of reddish areas is decreased). As a consequence, the blocking probability increases.

However, if we carefully choose the switching off point, we can guarantee that the blocking probability is still below the target value (peak value of the reference/benchmark system). Obviously, the longer the sector is silent, the better energy efficiency we can achieve. Actually, the switching point is chosen where the blocking probability is exactly the same as the target value to maximize the energy saving. We define the energy saving ratio μ_{sav} as the ratio of the overall system energy consumption with sectors switching off over that without switching off. Figure 6.10(b) indicates that if we do not switch off sectors when the traffic demand is low,

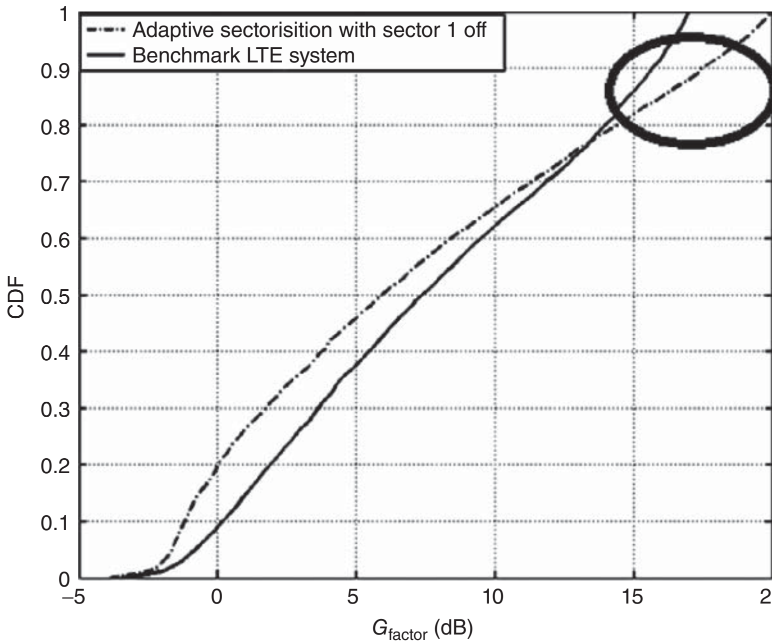

Figure 6.9 Cumulative distribution function (CDF) of G_{factor}

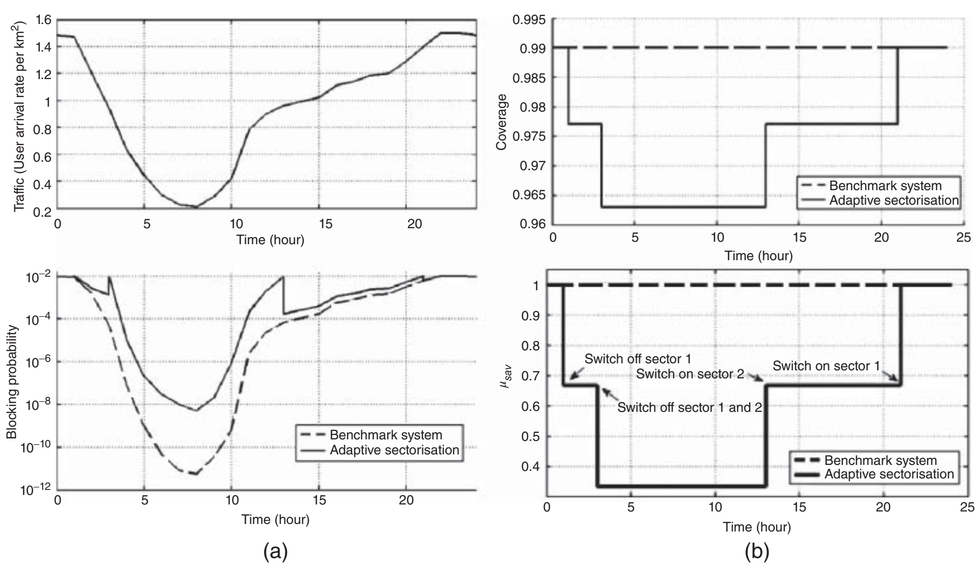

Figure 6.10 (a) Traffic demand and blocking probability, (b) coverage and saving ratio

the blocking probability is unnecessarily lower than the target, which means that resources are actually overpaid to provide much better QoS than expected. If we switch off one sector, the two active sectors are now serving all the users and there is a jump for the blocking probability because of the sudden increase of the user arrival rate. However, as we mentioned, by wisely selecting the switching point, the blocking probability is still below the target. When sector 2 is further switched off, the blocking probability jumps again but is still below the target level. Figure 6.10(b) shows that if we do not change the beamwidth of the remaining sectors, the users in the silent sector suffer from weak receiving signals and therefore the coverage target cannot be achieved. On the contrary, if we employ the adaptive (d)sectorization by widening the beamwidth when only one sector is switched off, the coverage level is well maintained (just 1% reduction). When two sectors are switched off, the omnidirectional antenna suffers from a 2% coverage reduction. The overall energy consumption is saved during the silent period of sector 1 (from time h1 to h21, total 20 hours) and the silent period of both sector 1 and 2 (from time h3 to h13, total 10 hours). Generally speaking, the overall energy consumption can be saved about 28% in each day (from h0 to h24).

6.3.3 Heterogeneous Network (HetNet)

In the previous section, we have shown that the network layout can be adaptively changed when traffic is lower than the peak value. However, the exponential growth in data traffic and number of connected terminals in contemporary life dramatically increases the peak traffic, which can only be tackled with densified planning with additional network nodes. Installation of new macro cells are neither convenient nor efficient. On the contrary, compact, low-power small cells are more cost-efficient to deploy and its plug-in-and-play property makes it suitable for improving field strength at macro cell edges, hotspot and indoor.

The mixture of macro cells and small cells, that is, heterogeneous network planning can be implemented in two different ways: first, the small cells can be deployed at the edge of the macro cells to tailor the density of the BSs to the required capacity per area as shown in Figure 6.11 [14]; second, the macro cells and micro cells can be mixed to offer non-standard deployment, where the current 3-sector hexagonal deployment will no longer be adopted [19].

When the micro cells are deployed at the edge of the macro cells, the area power consumption gains are depicted in Figure 6.12 for the four scenarios [14], where pure macro and micro deployment are also plotted for comparison purpose:

- Scenario 1: 1 micro per macro cell at the cell edge
- Scenario 2: 2 micro per macro cell at the cell edge
- Scenario 3: 3 micro per macro cell at the cell edge
- Scenario 4: 5 micro per macro cell at the cell edge.

The superiority of pure micro networks with regard to area power over throughput has to be related to its much higher area power consumption, that is, the optimum deployment does not provide excessive area throughput over the real capacity requirement. Given a throughput request, deployment of additional micro cells can only be efficient when this throughput request if relatively high. If the traffic request is low, the additional capacity gains brought by the micro cells are overwhelmed by the extra power consumption.

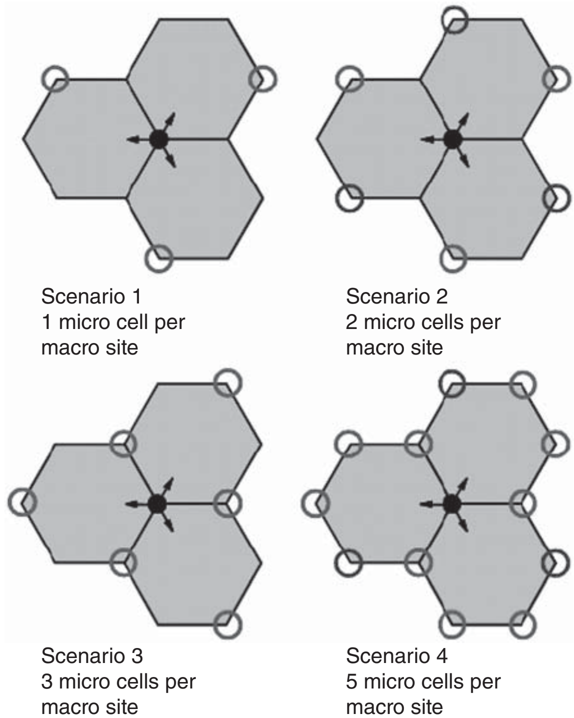

Figure 6.11 HetNet with micro cells at the edge of macro cells

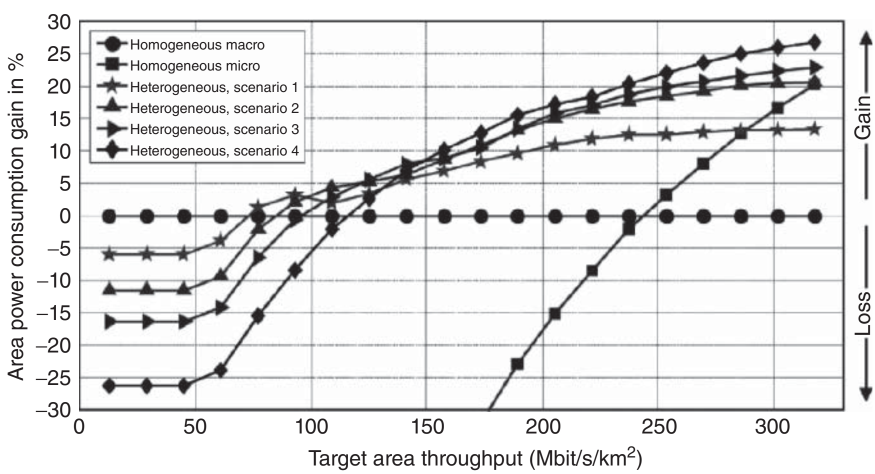

Figure 6.12 Area power consumption gain

As a consequence of even traffic demand distribution, arbitrary placement of base station sites and three-sectorized antenna configuration with 120° separation are valid. However, if the traffic distribution is not even, hexagonal deployment may lead to suboptimal network. In order to optimize deployment solution, an automatic algorithm was developed in Ref. [18] that performs the following main steps, in order to minimize energy consumption: (i) places base station sites, (ii) sets sector antennae main lobe direction (while keeping 120° separation for the three sectors) and (iii) reduces maximum output power of sectors. The algorithm is assumed to perform in an environment where the possible base station site locations form a discrete set of coordinates (non-arbitrary placement, e.g. modelling rooftops); traffic demands are assumed to be given in kbps units at finite number of discrete locations and it is assumed that antenna directions and output powers of sectors can be tuned. It has been shown in Ref. [19] that this non-hexagon deployment can achieve roughly 30% energy saving when compared with conventional hexagon deployment.

6.4 Network Design and Planning for Rural Areas

The green techniques introduced in the last section are more suitable for urban area where the BSs are deployed with relatively high densification. In rural area, where the cell size is large and the traffic level is low, coverage is a more important facto to be considered. In such a case, relaying technique is well known to be able to improve the data transmission in the cell edge and/or to provide coverage in new areas. Installation of new relay nodes (RNs) instead densifying the macro only networks can be more energy efficient. However, the efficiency improvements are highly depending on the geographic planning of the relay nodes and the capabilities of the relay nodes.

The basic idea behind relaying is to receive help from some radio nodes, called relays, to perform more spectrum and energy-efficient communications [20, 21]. RNs can be specifically devoted network nodes or other user devices in the vicinity. The relay node can either be used as a simple repeater which just amplifies and forwards its reception without any further processing, namely amplify-and-forward (AF) [20], or be equipped with more sophisticated baseband processing capability to be able to decode, re-encode and forwards received messages [21]. The latter one is called decode-and-forward (DF). These two schemes have already been part of LTE-Advanced to offer the possibility to extend coverage and increase capacity, allowing more flexible and cost-effective deployment options [22]. Other than these two schemes, compress-and-forward has drawn considerable attentions recently, where the relay node compresses and forwards its observation to the destination, benefiting from receive diversity [23]. In order to further exploit the uncertainty of wireless media, relaying schemes can be combined in a hybrid fashion and achieve improved spectrum efficiency because of the flexibility, [24, 26].

The link level results reveal that the S–D distance and the position of RN play an important role in energy efficiency (EE) performance [19]. When extended to a cellular network scenario, multiple RNs can be deployed in each sector. In such a case, we need to find the optimal number of deployed relay nodes (RNs) and their locations.

Figure 6.13 shows the SINR map with 2 and 4 relay nodes in each sector, respectively. The SINRs of the vicinity areas of the relay nodes are enhanced. When the number of relay nodes increases, a larger area is covered by the relay nodes, which means UEs are more likely to receive help from the relay nodes.

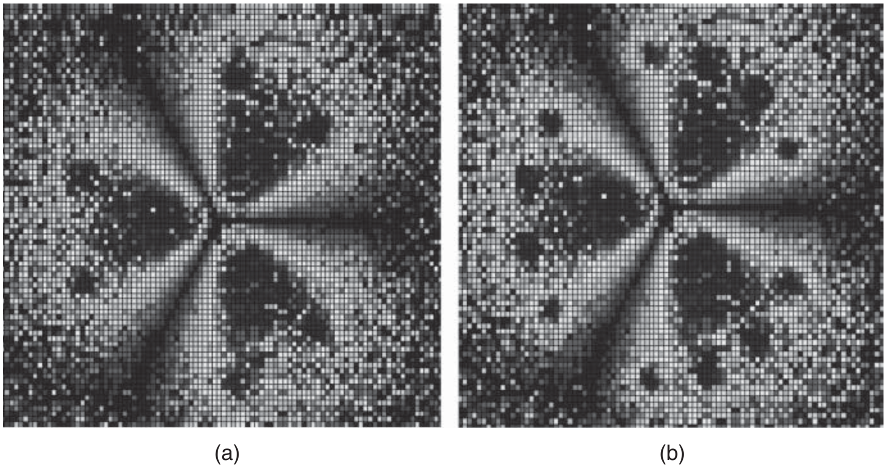

Figure 6.13 SINR map with 2 and 4 relay nodes in each sector

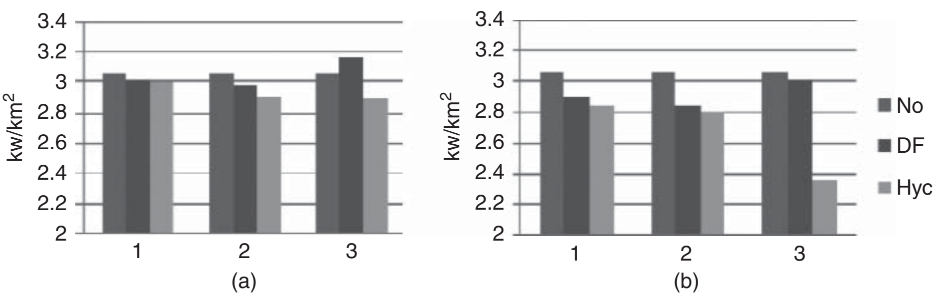

Figure 6.14 Energy efficiency: (a) one RN, (b) 4 RNs in (1) short, (2) middle, and (3) long BS-RN distance

As shown in Figure 6.14, hybrid relay achieves better EE performance than DF and the best EE performance is achieved at the cell edge. With four relays, although the overall energy consumption is increased compared to the one relay case during the time of transmission, the required transmission time for UEs served by RNs is reduced, increasing the idle time of the BSs. In this regard, more UEs can obtain help from the RNs and lead to lower energy consumption. However, this does not mean the number of relays can be increased without limit. At some point, the disadvantage of more offset energy consumption, which is the energy consumption even when transmission power is zero, will outweigh the benefit of improved throughput and transmission power, thus causes more energy consumption. Based on these results, we can conclude that the hybrid relay system that enables a pair of terminals (BS and RNs) to exploit spatial diversity shows significant improvement in energy efficiency performance in terms of consumed energy per km^2. However, compared with direct transmission, the cooperative strategy only shows significantly improved energy efficiency when the RN is not deployed in the very close area of the BS.

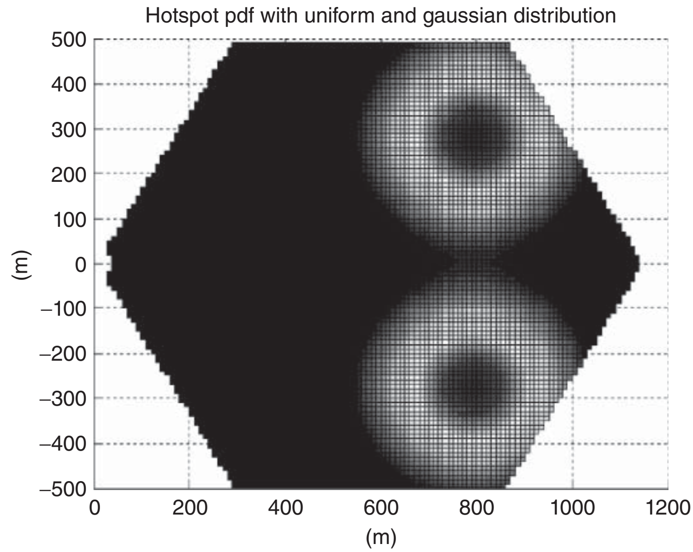

Figure 6.15 Traffic distribution with two hotspots covered by the RNs

Relay can be used to support also hotspot configurations which can be a realistic representation of the scenario where relay nodes are deployed to collect a large amount of traffic concentrated in a small area. In this case, the traffic distribution has been shaped assuming that half of the traffic of the cell is uniformly distributed and the remaining half is concentrated under the two relay nodes. This part of the traffic is distributed according to a spatial Gaussian distribution whose variance has been chosen so that 95% of this traffic falls in the area that is served by the relay nodes. The resulting distribution in the case of a cell with ISD 500 m is shown in Figure 6.15.

The results are shown in Figure 6.16 [25]. The column relative to BS SoTA uses the power model provided in Ref. [18] for state-of-the-art 2010 Base stations. It can be seen that RNs serve the hotspot of traffic offers a significant improvement, lowering the overall energy consumption index of 20.8% in the case of 2-hop relay nodes, and of 5.6% in the case of multicast relay nodes. Even higher gains can be achieved when also base stations are upgraded with a more efficient power model. When newer base stations are deployed alone, a good increase in energy efficiency is reached, but the improvements granted by relay nodes still stack with this higher performance, lowering the corresponding energy efficiency index of an additional 28.4% in the case of 2-hop relays and of 2.5% in the case of multicast relays.

6.5 Conclusions and Future Works

In this chapter, the detailed designing choices as well as the network planning solutions have been explained based on theoretical derivations, performance evaluations and system level studies. The discussed solutions address the high-energy efficiency targets of the future mobile cellular networks which feature very high traffic demanding, highly fluctuating daily traffic profile, and so on.

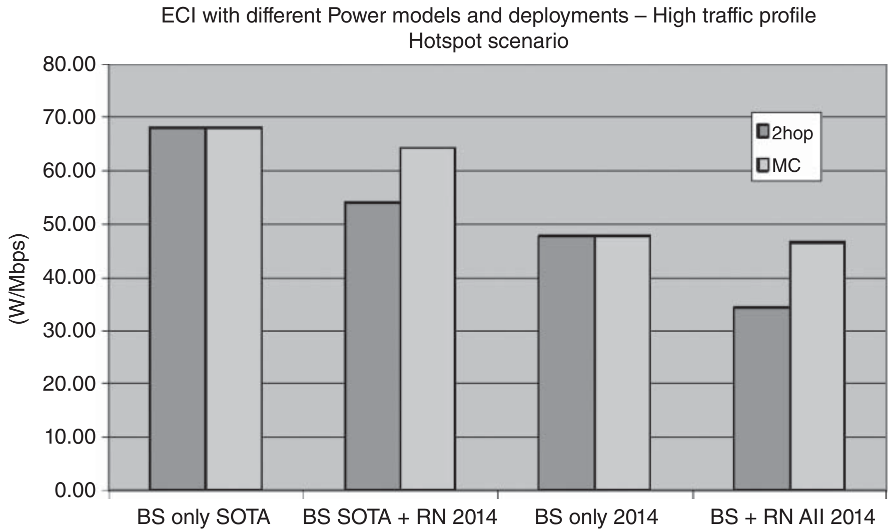

Figure 6.16 Overall ECI in the hotspot scenario with different deployments]

We have shown that the network design parameters, such as cell size, can be optimized according to the spatial traffic variations to save energy. After rollout, the network layout is difficult to change, but there are opportunities to adaptively adjust the network layout by applying some on/off schemes according to daily traffic variations. In particular, the (de)sectorization scheme, where the BS can switch off some of sectors to save energy during low traffic periods, is shown to be promising.

In the literature, it is often stated that the usage of small cells reduces transmit power. However, it is explained in this chapter that the densification of the current network with additional small cells can only be energy efficient in high traffic. Furthermore, it is shown that non-hexagon deployed heterogeneous network might be optimal when the spatial traffic is not even.

Furthermore, relaying technique is investigated for rural scenario. The hybrid relaying scheme with capability of switching between DF and CF are achieving considerably energy savings if their locations are carefully chosen. Additionally, relay can also be used in hotspot to improve energy efficiency.

Some new solutions are proposed for future architecture evolution. One of them is the idea of separation of control and data plan as shown in Figure 6.17. In this split concept, the user terminals exchange data traffic with small cells and signalling traffic with large base stations which act like a umbrella covering multiple small cells. This concept provides additional degree of freedom and has potentials to improve the resource usage efficiency and energy efficiency because the small cells no longer need to transmit/receive signalling messages frequently and, therefore can be put into "deep sleep" mode to save energy.

Another proposed idea is the federated network where the network is virtualized in a federated manner through infrastructure sharing among operators, with the capability in the provisioning of ample opportunity for competition and differentiation among involved actors. The designed architecture and management mechanisms exploit increasing variation of traffic

Figure 6.17 Control/data plane splitting

distribution among operators and ensure transparent service provisioning to end users. These potential candidates for future architectures are currently drawing more and more attention from both academia and industry.

References

[1] A. Penttinen, Chapter 10 – Network Planning and Dimensioning, Lecture Notes: S-38.145 – Introduction to Teletraffic Theory, Helsinki University of Technology, Fall, 1999.
[2] AriaNetworks white paper, "Seven best practices for successful mobile network planning," [online]. Available: http://www.aria-networks.com.
[3] S. Xie, "Planning, design and building large-scale networks at campus," in Proc. ICCRD, pp. 495–498, 2011.
[4] https://www.ict-earth.eu/Earth project Summary leaflet.
[5] Press Release, EU Commissioner Calls on ICT Industry to Reduce Its Carbon Footprint by 20% as Early as 2015, MEMO/09/140, 2009.
[6] http://www.greentouch.org.
[7] C. Lamour, "Energy consumption of mobile networks," The Basestation eNewsletter, 2008.
[8] F. Richter, A. J. Fehske, and G. P. Fettweis, "Energy efficiency aspects of base station deployment strategies for cellular networks," in Proc. of VTC2009-Fall, Anchorage, Alaska, USA, pp. 1–5, 2009.
[9] M. A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Optimal energy savings in cellular access networks," in Proc. of ICC2009 Communications Workshops, Dresden, Germany, pp. 1–5, 2009.
[10] K. S. Neier-Hellstern, "The analysis of a queue arising in overflow models," IEEE Trans. Commun., vol. 37, no. 4, pp. 367–372, 1989.
[11] L. N. Singh and G. R. Dattatreya, "A novel approach to parameter estimation in Markov-modulated Poisson processes," in 2004 IEEE Emerging Technologies Conference (ETC), Richardson Texas, 2004.
[12] F. Richter, A. J. Fehske, and G. P. Fettweis, "Energy efficiency aspects of base station deployment strategies for cellular networks," in Proc. of VTC2009-Fall, Anchorage, Alaska, USA, pp. 1–5, 2009.
[13] K. S. Neier-Hellstern, "The analysis of a queue arising in overflow models," IEEE Trans. Commun., vol. 37, no. 4, pp. 367–372, 1989.
[14] EARTH project, "D2.2 – definition and parameterization of reference systems and scenarios," [online]. Available: https://www.ict-earth.eu/publications/publications.html.
[15] EARTH project, "D3.1 – most promising tracks of green network technologies," [online]. Available: https://www.ict-earth.eu/publications/publications.html.
[16] Y. Qi, M. Imran, and R. Tafazolli, "On the energy aware deployment strategy in cellular systems," in Proc. of PIMRC 2010, pp. 363–367, 2011.
[17] Technical Specification Group Radio Access Network, "Evolved universal terrestrial radio access (E0UTRA); LTE radio frequency (RF) system scenarios," 3rd Generation Partnership Project (3GPP), Tech. Rep. TS 36.942, 2008–2009.
[18] EARTH project, "D2.3 - Energy efficiency analysis of the reference systems, areas of improvements and target breakdown" [online]. Available: https://www.ict-earth.eu/publications/deliverables/deliverables.html.

[19] EARTH project, “D3.3 – Final report on green network technologies” [online]. Available: https://www.ict-earth.eu/publications/deliverables/deliverables.html.
[20] J. N. Laneman, D. N. C. Tse, and G. W. Wonell, “Capacity diversity in wireless networks: efficient protocols and outage behavior,” IEEE Trans. Inform. Theory, vol. 50, no. 12, pp. 3062–3080, 2004.
[21] J. N. Laneman, Cooperative diversity in wireless networks: algorithms and architectures, PhD thesis, Massachusetts Institutes of Technology, 2002.
[22] 3GPP TR 36.814 v1.5.1 (2009-12), Further Advancements for E-UTRA, Physical Layer Aspects.
[23] M. Kats and S. Shamai, “Relay protocols for two colocated users,” IEEE Trans. Inform. Theory, vol. 52, no. 6, pp. 2329–2344, Jun. 2006.
[24] Y. Qi, R. Hoshyar, M. A. Imran, and R. Tafazolli, “H^2-ARQ-relaying: spectrum and energy efficiency perspectives,” IEEE JSAC, vol. 29, no. 8, pp. 1547–1558, 2011.
[25] Y. Qi, M. Ali Imran, S. Dario, and F. Roberto, “On the development opportunities for increasing energy efficiency in LTE-advanced with relay nodes,” in Proc. WWRF, 2012.
[26] Y. Qi, M. Ali Imran, R. Demo Souza, and R. Tafazolli, “On the Optimization of Distributed Compression in Multi-Relay Cooperative Networks,” accepted by IEEE Trans. on Vehicular Technology, 2015.