Part I

Communication architectures and models for green radio networks

1 Fundamental trade-offs on the design of green radio networks

Yan Chen, Shunqing Zhang, and Shugong Xu

1.1 Introduction

There is currently a global concern about the rise in the emission of pollutants and energy consumption. The carbon dioxide (CO_2) footprint of the information and communications technologies (ICT) industry, as pointed out by [1], is 25% of the 2007 carbon footprint for cars worldwide, which is similar to that of the whole aviation industry. Within the ICT industry, the mobile network is recognized as being among the biggest energy users. The exponentially growing data traffic in mobile networks has made the issue an even grander challenge in the future. In a data forecast report provided by Cisco [2], it has been pointed out that the global mobile data traffic will increase 26-fold between 2010 and 2015. In particular, unexpectedly strong growth in 2010 has been observed mainly due to the accelerated adoption of smartphones. For instance, China Unicom's 3G traffic increased 62% in a single quarter from Q1 to Q2 of 2010, while AT&T reported a 30-fold traffic growth from Q3 2009 to Q3 2010. The unprecedented expansion of wireless networks will result in a tremendous increase in energy consumption, which will further leave a significant environmental footprint. Therefore, it is now a practical issue and demanding challenge for mobile operators to maintain sustainable capacity growth and, at the same time, to limit the electricity bill. For instance, Vodafone Group has announced the goal of reducing its CO_2 emissions by 50% against its 2007 baseline of 1.23 million tonnes, by the year of 2020 [3]. Figure 1.1 gives examples of the green demand from mobile operators worldwide.

As has been pointed out in [4], the radio access part of the wireless network accounts for up to more than 70% of the total energy bill for a number of mobile operators. Therefore, developing energy-efficient wireless architectures and technologies is crucial to meet this challenge. Research actions have been taken worldwide. It is now an important trend for the wireless designers to take energy consumption and energy efficiency into their design frameworks. Vodafone, for example, has predicted that energy-efficiency improvement will be one of the most important areas that demand innovation for wireless standards beyond LTE [5].

Green radio research is a large and comprehensive area that covers all layers in the design of efficient wireless access networks. There have been efforts devoted to traditional energy-saving ways, such as designing ultra-efficient power amplifiers, reducing feeder losses, and introducing passive cooling. However, these efforts are isolated and thus cannot make a global vision of what we can achieve in five or ten years for energy

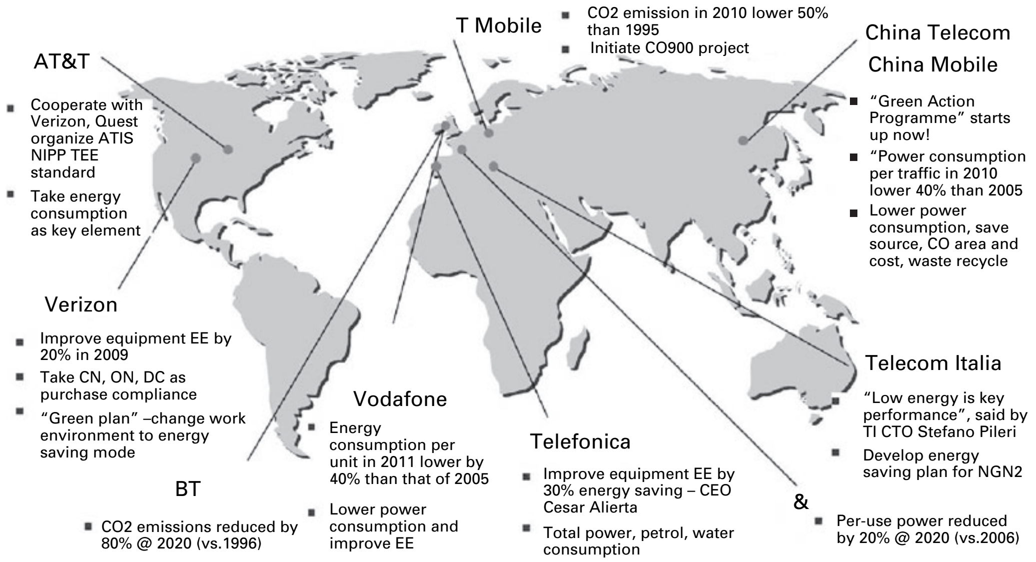

Figure 1.1 Global operators' demand on green communications.

saving as a whole. Innovative solutions based on top-down architecture and joint design across all system levels and protocol stacks are needed, which cannot be achieved via isolated efforts.

Green research projects with holistic approaches and joint efforts from the industry and the academia have sprung up all over the world during recent years. For instance, the EARTH (Energy Aware Radio and neTwork tecHnologies) project [6]–[7] under the European Framework Program 7, started to develop green technologies at the beginning of 2010. In the UK, *GreenRadio* [8] is one of the Core 5 Programs in Mobile VCE that has been set up since 2009. Most recently, the *GreenTouch* Consortium sets its 5-year research goal to deliver the architecture, specification, and roadmap needed to reduce the end-to-end energy-consumption per bit by a factor of 1000 from the current level by the year 2020. In addition, there are also active discussions in standardization organizations, such as ETSI, ATIS, and 3GPP, on energy-efficiency metrics and measurement, as well as studies for base station level or network level savings.

Instead of a survey that reaches every aspect of the matter, or a report elaborating one specific green research point, this chapter focuses on the fundamental framework for green radio research and strings together the currently scattered research points using a logical "rope." In this chapter we propose four fundamental trade-offs to construct such a framework. These were first introduced in [9]. As depicted in Figure 1.2, they are

- *Spectrum efficiency–energy efficiency (SE–EE) trade-off*: given the bandwidth available, to balance the achievable rate and the energy cost;
- *Bandwidth–power (BW–PW) trade-off*: given the target transmission rate, to balance the bandwidth utilized and the power needed;

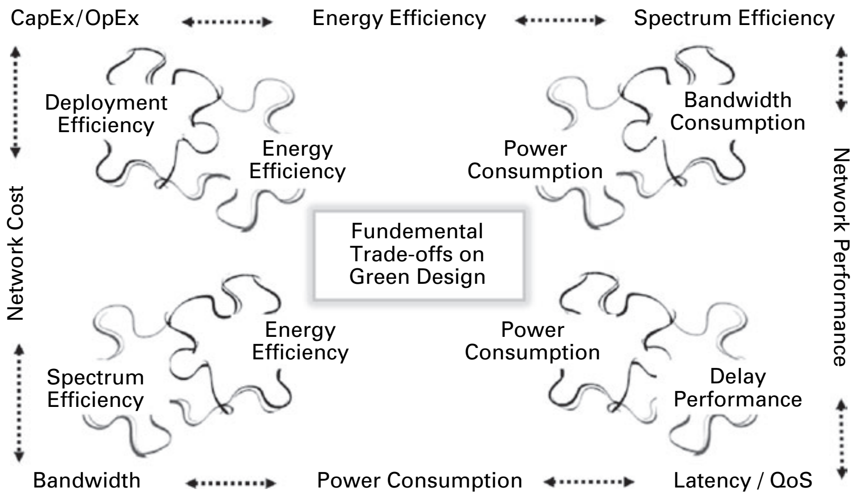

Figure 1.2 Four fundamental trade-offs form the core of green research.

- *Delay–power (DL–PW) trade-off*: to balance the average end-to-end service delay and the average power consumed in the transmission;
- *Deployment efficiency–energy efficiency (DE–EE) trade-off*: given the network traffic requirement, to balance the deployment cost, throughput, and energy consumption, in the network as a whole.

By means of the four trade-offs, key network performance/cost indicators are all strung together. In the rest of the chapter, we will elaborate in detail the definitions, justifications, practical concerns, as well as research directions for each of the trade-off studies. In particular, we shall show that in practical systems, the trade-off relations usually deviate from the simple monotonic curves derived from Shannon's formula, which brings a new design philosophy.

1.2 Insight from Shannon's capacity formula

Shannon's capacity formula [10] establishes a bridge between the maximum achievable transmission rate R and the received power $P^{(r)}$ for the point-to-point additive white Gaussian noise (AWGN) channel, i.e.

$$R = W \log_2 \left(1 + \frac{P^{(r)}}{W N_0}\right), \tag{1.1}$$

where N_0 is the noise power density at the receiver and W is the system bandwidth. Though Shannon's ground-breaking formula has been known for more than half a century, people mainly look at it from the channel capacity point of view. However, as we will show later in this section, the formula actually gives us a fundamental insight into the energy-related trade-offs in the wireless point-to-point link transmission. In this section, we shall formally introduce the definitions of the trade-offs and sketch their behavior predicted by Shannon's capacity formula.

The following are the equivalent transformations of the above formula, which will be used in the characterization of the different trade-offs.

$$\frac{R}{W} = \log_2\left(1 + \frac{R}{W}\frac{E_b^{(r)}}{N_0}\right). \tag{1.2}$$

$$\frac{1}{T_b} = W\log_2\left(1 + \frac{1}{T_b}\frac{E_b^{(r)}}{WN_0}\right). \tag{1.3}$$

In the equations above, $E_b^{(r)}$ stands for the average energy per bit and T_b denotes the average transmission time per bit. They are introduced through the relations $E_b^{(r)} = P^{(r)}/R$ and $T_b = 1/R$. Further, considering a constant attenuation on the transmitted signal, denoted as a simple function of the transmit power $P^{(t)}$, namely $f(P^{(t)}) = \kappa_0 P^{(t)}/d^\alpha$, where κ_0 and α are the attenuation coefficient and exponent, respectively, we have

$$\frac{R}{W} = \log_2\left(1 + \frac{R}{W}\frac{E_b^{(t)}}{N_0}\frac{\kappa_0}{d^\alpha}\right). \tag{1.4}$$

1.2.1 SE–EE trade-off

Spectrum efficiency (SE), defined as the system throughput for unit bandwidth, i.e. bits/sec/Hz, is a widely accepted criterion for wireless network optimization. The peak value of SE is always among the key performance indicators of standardization evolution such as 3GPP. For instance, the target downlink SE of 3GPP increases from 0.05 bps/Hz to 5 bps/Hz as the system evolves from GSM to LTE. On the contrary, energy efficiency (EE), defined as the data rate achievable per unit of transmitted power, i.e. bits/sec/Watt, namely bits/Joule, was previously ignored by most of the research efforts and has not been considered by 3GPP as an important performance indicator until very recently.

Shannon's groundbreaking work on reliable communication over noisy channels showed that there is a fundamental trade-off between SE and received/transmitted EE. Informally speaking, a lower transmission rate leads to a lower transmitted power, for the same system bandwidth. Given the definitions above, SE can be expressed as $\eta_{SE} = R/W$ and the received EE as $\eta_{EE}^{(r)} = 1/E_b^{(r)}$. From (1.2), the SE–EE trade-off can be characterized by

$$\eta_{EE}^{(r)} = \frac{\eta_{SE}}{(2^{\eta_{SE}} - 1)N_0}, \tag{1.5}$$

which is depicted on the left-hand side (LHS) of Figure 1.3, where $N_0 = -174$ dBm. Seen from both the mathematical relation and the figure, η_{EE} converges to a constant, $1/(N_0 \ln 2)$, when η_{SE} approaches zero. On the contrary, η_{EE} approaches zero when η_{SE} tends to infinity. Similarly, considering the relation in (1.4), the transmit EE-SE trade-off

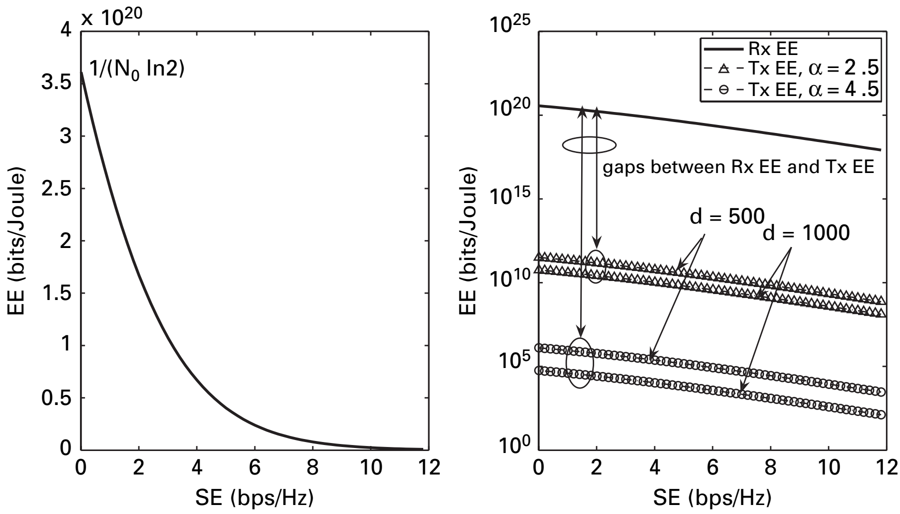

Figure 1.3 Illustration of the SE–EE trade-off. On the LHS, the figure shows the trade-off relation between SE and received EE from Shannon formula, while on the RHS, the figure depicts the transmit EE as function of the path-loss exponent α at different distance d.

can be expressed as

$$\eta_{EE}^{t} = \frac{\eta_{SE}}{(2^{\eta_{SE}} - 1)N_0} \cdot \frac{\kappa_0}{d^{\alpha}}, \tag{1.6}$$

as shown on the right-hand side (RHS) of Figure 1.3. The gaps between the received EE and the transmit EE depend heavily on the transmission channel degradation, i.e. the path-loss exponent α and the transmission distance d.

1.2.2 BW–PW trade-off

Bandwidth (BW) and power (PW) are both fundamental but limited resources in wireless communications. From the Shannon's capacity formula in (1.1) and (1.4), the relation between the transmit power, P^t, and the transmission bandwidth, W, for a given transmission rate, R, can be expressed as

$$P^t = W N_0 (2^{\frac{R}{W}} - 1) \cdot \frac{\kappa_0}{d^{\alpha}}. \tag{1.7}$$

The expression above exhibits a monotonic relation between PW and BW, as sketched in the LHS of Figure 1.4. The fundamental BW–PW trade-off shows that, to transmit at a given data rate, the expansion of the transmission bandwidth is preferred in order to reduce transmit power and thus achieve better energy efficiency. From (1.7), in the extreme case, the minimum power consumption is as small as $N_0 R \ln 2$ if there is no bandwidth limit.

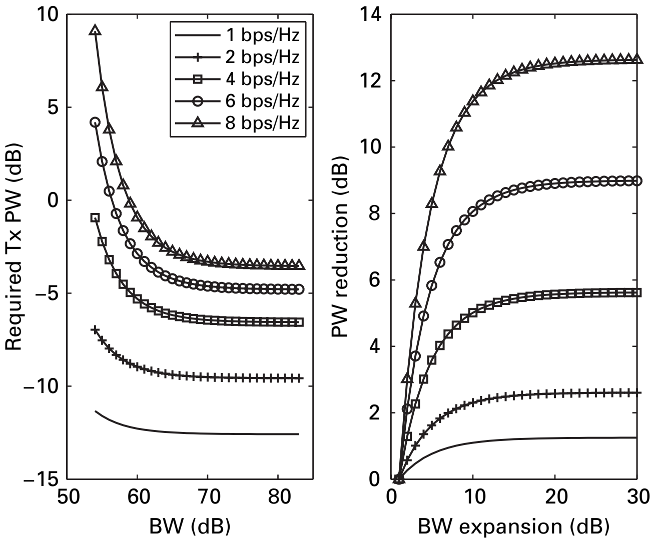

Figure 1.4 Illustration of the BW–PW trade-off derived from Shannon formula. The difference between the set of curves is the initial SE values before any BW expansion. The left figure gives the absolute value of the required transmit power while the middle one shows the PW reduction gain. The right figure depicts the PW reduction gain at 10 dB BW expansion at different initial SE values. $\kappa_0/d^\alpha = -140$ dB.

Figure 1.4 depicts the BW–PW trade-off from three different angles. Firstly, the left-most figure shows the relation between the required transmit power and the system bandwidth, the trend of which behaves exactly as equation (1.7) predicts. The middle figure shows the PW reduction as function of the BW expansion. From (1.4), the reduction in the transmit power is the same as that in the received power. It can be observed from the figure that increasing the BW by ten (10 dB) brings considerable gain in PW reduction, no matter what the initial SE of the system is. Larger than 10 dB BW expansion, however, only adds marginal gain. Moreover, the higher the initial SE, the larger the PW reduction gain. It can be found from the right-most figure that expanding the BW 10 times brings less than 3 dB PW reduction gain to a system with the initial SE at 2 bps/Hz, but offers a larger than 10 dB gain to the system with the initial SE larger than 8 bps/Hz.

1.2.3 DL–PW trade-off

The metrics such as EE, SE, and BW, as described in the two trade-offs above, are important system performance criteria but cannot be directly observed by end users. Delay (DL) is different to these metrics and is usually taken as a measure of quality of service (QoS) and user experience. According to the scope of the definition, there are different types of delay. Two major ones are the physical (PHY) delay, defined as the time spent during the physical layer transmission, and the medium-access-control (MAC) delay, defined as the sum of both waiting time in the MAC layer data queue and transmission time in the PHY layer.

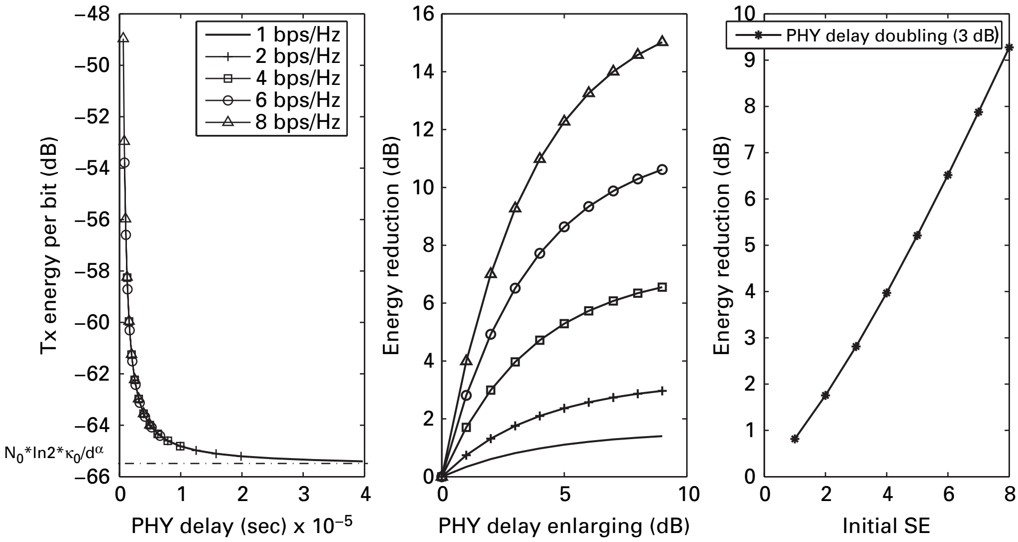

Figure 1.5 Illustration of the PHY DL–PW trade-off derived from Shannon's formula. The middle figure shows the gain of energy reduction as a function of the PHY delay increasing. The right figure shows the energy reduction gain provided by doubling the PHY delay at different initial SE values. $\kappa_0/d^\alpha = -140$ dB and $W = 200$ kHz.

Let us start with the simpler one, the PHY delay, for which the Shannon's capacity formula reveals most of the characteristics. For point-to-point transmission over AWGN channels, formulas (1.3) and (1.4) tell us the average energy per bit required to transmit a data bit in time T_b can be calculated as

$$E_b^t = N_0 T_b W \left(2^{\frac{1}{T_b W}} - 1\right) \cdot \frac{\kappa_0}{d^\alpha}. \tag{1.8}$$

The above expression shows a monotonically decreasing relation between received energy per bit and PHY delay, as sketched on the left of Figure 1.5. The middle figure of Figure 1.5 shows that the higher the initial SE, the more energy reduction gain can be obtained from enlarging the PHY delay. For instance, doubling the PHY delay reduces the average transmit energy per bit by less than 2 dB for the initial SE of 2 bps/Hz but more than 6 dB for that of 6 bps/Hz. This is true for single symbol transmission or continuous symbol transmission (full buffer). However, the relation may change when we consider bursty data blocks, as will be shown later in Section 1.3.

The MAC delay, on the other hand, is closely related to the upper layer traffic arrivals and statistics. By Little's law [11], the average delay has a direct relation with the average queue length in the data queue. As a result, the design of transmission schemes shall cope with both channel uncertainties, traffic variations, and queue dynamics, which makes the characterization of DL–PW trade-off more complicated. Shannon theory alone is not enough to characterize the DL–PW in these scenarios. Other theoretical analysis tools are needed, such as queueing theory [11] and control theory [12]. Moreover, as technologies evolve, the types of future wireless services become diverse enough to have heterogeneous delay requirements. Therefore, in order to build a green radio, it is important to know when and how to trade tolerable delay for low power.

1.2.4 DE–EE trade-off

Deployment efficiency (DE), a measure of network throughput per unit of deployment cost, namely bits/$ or Mbits/$, is an important network performance indicator for mobile operators. The deployment cost consists of both capital expenditure (CapEx) and operational expenditure (OpEx). For radio access networks, the CapEx mainly includes infrastructure costs, such as base station equipment, backhaul transmission equipment, site installation, etc., while the key drivers for the OpEx are electricity bill, site and backhaul lease, and operation and maintenance costs. The scope of the EE definition in the previous trade-offs can either be for a single base station or for a network; the EE concept involved in the DE–EE trade-off is a metric for the whole network, namely a measure of network throughput per unit of network energy consumption, i.e. bits/Joule.

The two different metrics often lead to opposite design criteria for network planning. For example, to save the expenditure on site rental, base station equipment, and maintenance, network planning engineers tend to "stretch" the cell coverage as much as possible. However, the path loss between the base station and mobile users will degrade by 12 dB whenever the cell radius doubles if the path-loss exponent is four, which induces a 12 dB increase in the transmit power to guarantee the same signal strength for those users at the cell edges. Some simple calculations give the result that to provide cellular coverage for a given area, increasing the number of base stations will save the total network transmit power by the same factor.

Table 1.1 helps to understand the inner logic. Assume the reference cell radius is d_0, β and γ are two coefficients associated with the cell size shrinking scenario where $0 \leq \beta, \gamma \leq 1$, inter-cell interference is not considered, and the transmit power for all users is kept the same, derived from the SE requirement η_{SE} of the cell-edge user. Figure 1.6 further depicts the DE and EE performance at different β. The DE and EE values in the figure are normalized by that of the reference scenario. An implicit assumption is that the total traffic served by different scenarios on the given area A is the same. The left-most figure shows that the improvement in EE via cell size shrinking depends heavily on the wireless channel environment, e.g. the path-loss exponent α. The larger the α is (faster degradation of the transmitted energy), the more benefit small cells could bring. As shown in the middle figure, the value of γ impacts the DE performance. Here, $1 - \gamma$ can be interpreted as the average cost reduction ratio per base station. Note that the increase in the number of cells adds extra cost in the backhaul and site maintenance. The constant offset in γ is added to account for that. Finally, the right-most figure shows how the network EE trades off DE. Note that when transmitting in free space ($\alpha = 2$), the trade-off relation no longer holds.

1.2.5 Summary

In the previous four subsections, we have elaborated the definitions of the four fundamental trade-offs as well as their behavior predicted by the Shannon's capacity formula.