Toward Energy-Efficient Operation of Base Stations in Cellular Wireless Networks

Kyuho Son  
University of Southern California, Los Angeles, USA

Hongseok Kim  
Sogang University, Seoul, Korea

Yung Yi  
Korea Advanced Institute of Science and Technology, Daejeon, Korea

Bhaskar Krishnamachari  
University of Southern California, Los Angeles, USA

CONTENTS
16.1 Introduction ...................................................... 436
16.2 Overview of Green Cellular Network Design .................... 437
16.3 Energy-aware heterogeneous deployment ........................ 440
   16.3.1 Micro base station deployment strategy ................. 441
       16.3.1.1 System model .................................. 441
       16.3.1.2 Problem formulation ............................ 442
       16.3.1.3 Key observations and algorithm ............... 442
       16.3.1.4 Numerical results .............................. 444
   16.3.2 Deployment structure: homogeneous vs. heterogeneous 446
       16.3.2.1 System model .................................. 447
       16.3.2.2 Problem formulation ............................ 448
       16.3.2.3 Optimal area power consumption vs. target
              10-percentile area throughput .................... 449
16.4 Base Station On/Off and User Association ...................... 450
   16.4.1 A quantitative case study ............................... 450
   16.4.2 System model and formulation ........................... 451
   16.4.3 Energy-efficient user association ....................... 455
   16.4.4 Traffic-driven base station on/off ..................... 456
       16.4.4.1 Greedy turning on algorithm ................... 456
       16.4.4.2 Design rationale ................................ 457
       16.4.4.3 Greedy turning off algorithm ................... 458

435
16.1 Introduction

For decades, abundant research in wireless networks has contributed to the improvement of energy-efficiency for battery-operated devices [1], e.g., prolonging the lifetime of sensor nodes and mobile terminals. However, more recently, potential harmful effects to the environment caused by CO₂ emissions and the depletion of non-renewable resources bring renewed focus on the need to develop more energy-efficient underlying network infrastructures that account for heavy energy usage [2].

It has been estimated that 3% of the world’s annual electrical energy consumption and 2% of CO₂ emissions are caused by the information and communication technology (ICT) industry [3]. According to this estimate, about a tenth of this can be attributed to cellular mobile communication systems. As
of 2008, the energy consumption corresponded to 60 billion kWh of electricity usage annually, about 40 million metric tons of CO$_2$ emissions each year. To put it into perspective this is equivalent to annual greenhouse gas emissions from about 8 million cars. Another consistent estimate states that there were over 600,000 base stations (BSs) in China deployed by three major operators which consumed about 20 billion kWh in 2007 [2].

From the perspective of cellular network operators, reducing energy consumption is not only a matter of social environmental responsibility towards being green and sustainable but also tightly related to their business survivability in coming years. They are spending huge operational expenditures (OPEX) to pay electricity bills. Moreover, it is expected to grow due to explosive growth in data demand and the possible increase of energy price in the near future. It is estimated that the energy consumption rises at 15-20% per year, doubling every five years in the field of ICT. This will result in a collective cellular network OPEX of $22 billion in 2013, according to a study from ABI Research [4]. Thus, reining back the spiraling OPEX is crucial to the continuing success of operators.

16.2 Overview of Green Cellular Network Design

Pushed by such needs of energy reduction, the operators have been seeking ways to improve energy-efficiency in all components of cellular networks including mobile terminals [5–7], BSs, and mobile backhaul networks. In particular, it has been reported that BSs are the key source of energy usage, contributing to 60-80% of the total power consumption$^1$ in cellular networks [8]. In this chapter, we therefore focus on energy-efficient operation of cellular BSs, which we refer to as green cellular base stations operation.

The goal of green cellular networks can be achieved in various ways:

- **Component level**: novel designs and hardware implementations, e.g., energy-efficient power amplifiers [9] and fanless cooler, or even cooling based on natural resources [10].

- **Link level**: energy-aware transmission and resource management schemes, e.g., power control [11] and user association [12].

- **Network level**: topological approaches from deployment to operation, e.g., smart deployment at the stage of network planning by using micro BSs [13–16] or relays [17, 18], and traffic-aware dynamic BS on/off [2, 8, 12, 19–21].

In this chapter, we focus on network/link level solutions with emphasis

$^1$Throughout this chapter, we interchangeably use the terms ‘energy’ and ‘power’ (really meaning average power) unless clear distinction is needed.
on BS operations and deployment topology designs, whose problem space can be seen from a variety of angles. As depicted in Figure 16.1, we explore the entire problem space based on the following two axis: (i) what we can control under what time scale, and (ii) what performance we need to sacrifice in saving energy. We used the word “sacrifice” because the key of green BS design lies in how to tune the tradeoff between energy and efficiency, where both energy and efficiency are measured by various metrics. A research paper in the community typically has corresponded to one or multiple points in this space. We will provide a taxonomy of the state-of-the-art research in the context of this problem space later in Table 16.1. We now elaborate on each axis here.

• (i) Control and time-scales. Saving energy in the cellular networks is differentiated by the control mechanisms we handle as well as their operating time scales. For example, deployment of energy-aware base stations may take the order of months or even years time scale, whereas the time-scale of transmission power control of mobile terminals or BSs is on the order of milliseconds.

• (ii) Efficiency metric and greening cost. Major system performance metrics of applications in cellular networks, which are also the ones to be sacrificed by power reduction, include throughput, delay, and blocking or outage probability. Throughput is an appropriate metric for best-effort data traffic, whereas delay and blocking/outage probability are mainly for real-time traffic.

Table 16.1 shows the survey of the recent papers which has studied greening effect and algorithms in wireless cellular networks, which are classified
TABLE 16.1 Classification of green cellular network problems.

<table>
<thead>
<tr>
<th>Work</th>
<th>Techniques</th>
<th>Metric</th>
<th>Keyword</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deployment</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BS On/Off</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Association</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Power control</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Relay</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CoMP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[22-26]</td>
<td>x - - - - -</td>
<td>$T$</td>
<td>E</td>
</tr>
<tr>
<td>[14,27-29]</td>
<td>x - - - - -</td>
<td>$T$</td>
<td>E</td>
</tr>
<tr>
<td>[13,30]</td>
<td>x x x - - -</td>
<td>$ASE$ or $T$</td>
<td>E</td>
</tr>
<tr>
<td>[2,16,19-21]</td>
<td>- x - - -</td>
<td>$T$</td>
<td>E</td>
</tr>
<tr>
<td>[31,32]</td>
<td>- x x - - -</td>
<td>$O$</td>
<td>E</td>
</tr>
<tr>
<td>[12,22]</td>
<td>- x x - - -</td>
<td>$D$ or $T$</td>
<td>E</td>
</tr>
<tr>
<td>[11]</td>
<td>- - - x - -</td>
<td>$T$</td>
<td>E</td>
</tr>
<tr>
<td>[17,18]</td>
<td>- - - x - -</td>
<td>$T$</td>
<td>E</td>
</tr>
<tr>
<td>[33,34]</td>
<td>- - - - x</td>
<td>$T$ or $O$</td>
<td>E</td>
</tr>
</tbody>
</table>

according to which metrics for system performance and energy consumption are taken with which focus on greening techniques.

Most of the papers use a model with two objectives chosen from the perspectives of system performance and energy consumption. A classical method to deal with such multi-objective cases are: maximizing $\text{system performance} - \eta \cdot \text{energy consumption}$ (similarly minimizing $\text{energy consumption} - \frac{1}{\eta} \cdot \text{system performance}$), or maximizing $\text{system performance}/\text{energy consumption}$, where $\eta$ is the weight that controls the tradeoff between two perspectives.

The system performance metrics they studied are denoted, in Table 16.1, by $T$ (Throughput, bit/sec), $D$ (Delay, sec), $O$ (Outage probability, %), and $ASE$ (Area Spectral Efficiency, bit/sec/Hz/m$^2$). The energy consumption metrics consist of $APC$ (Area Power Consumption, W/m$^2$), and $E$ (Energy, J). Dividing one metric by the other, the unified metrics such as $EE$ (Energy Efficiency, bit/J) and $AEE$ (Area Energy Efficiency, bit/J/m$^2$) are also often used in literature to capture the tradeoff between system performance and energy consumption.

Under the techniques categorized in Figure 16.1 and Table 16.1, we summarize several control mechanisms that are mainly discussed in this chapter.

(a) Deployment. Future cellular networks are expected to have a mixture of large and small cells. In such environments, the general problem of deployment is to determine where (site locations), how many and which type (macro/micro/pico/femto) of BSs need to be deployed in an energy-efficient manner.
(b) BS on/off. BSs are typically deployed and operated on the basis of peak traffic volume and also stayed turned-on irrespective of traffic load. Recent temporal traffic traces [2, 8] point out that BSs are largely under-utilized in late night or early morning over the daily time-scale and also under-utilized in weekends over the weekly time-scale. Dynamic BS operation which turns on and off depending on the imposed traffic may lead to huge energy saving.

(c) User association. Turning BSs on/off is naturally coupled with user association because a user may need to change its associated BS. A proper association mechanism, which determines the most energy-efficient BS among a set of active BSs, is necessary to fully exploit the amount of energy savings from traffic-aware BS operations.

(d) Power control. To mitigate interference, the current and future cellular systems are expected to employ dynamic interference management (IM) algorithms that perform dynamic BS transmission power control depending on the scheduled users over the multiple neighboring cells. It is also possible to achieve huge greening gain by reducing the transmitted power of BSs intelligently while maintaining the system performance.

In the following sections, we present details on recent developments from analytical algorithms to practical applications regarding: (i) energy-aware heterogeneous deployment in section 16.3 and (ii) joint BS on/off and user association problem in section 16.4. Finally, in section 16.5, we present several open problems and other directions including a new paradigm and architecture towards green cellular networks.

16.3 Energy-aware heterogeneous deployment

Generally speaking, the problem of energy-aware deployment is to determine where (site locations), how many and which type (macro/micro/pico/femto) of BSs need to be deployed in an energy-efficient manner. There have been many efforts in literature [13, 14, 22–30] that deal with the general problem. In particular, this section introduces two case studies in an energy-aware heterogeneous deployment.

The first case study, that is presented in section 16.3.1, considers the use of micro BSs on top of the pre-existing deployment of macro BSs to upgrade the network capacity in a cost-effective way. A theoretical implication and a practical solution for the following questions are provided: what is the minimum number of additional micro BSs to meet the quality of service requirement

\[\text{Within the scope of this chapter, active means that a BS is ON state.} \]

\[\text{The key results of the two case studies we present in sections 16.3.1 and 16.3.2 were drawn from [13] and [14, 28], respectively.} \]
and where they should be deployed. The second case study in section 16.3.2 focuses on the predefined topology of regular grid, where only inter-site distance is a variable, and considers two deployment strategies: a conventional homogeneous and heterogeneous deployments. Though the model is simple, it allows us to systematically investigate the energy efficiency of the deployment strategies and to answer under what conditions, which deployment strategy is better.

### 16.3.1 Micro base station deployment strategy

#### 16.3.1.1 System model

**Network model.** A wireless cellular network where the sets of macro and micro BSs, denoted by $B_M$ and $B_m$, respectively, is considered. Throughout the section, subscript $M$ is used for macro BSs, and $m$ is for micro BSs. Let us denote by $b \in B = B_M \cup B_m$ the index of BSs. Our main focus is on downlink communication that is a primary usage mode for the mobile Internet, i.e., from BSs to mobile terminals (MTs). However, we would like to emphasize that some aspects of our work can be applied to the uplink as well.

**Link model.** The received signal strength from BS $b$ to MT at location $x$ can be expressed as $E_b(x) = p_b \cdot g_b(x)$, where $p_b$ denotes the transmission power of BS $b$, $g_b(x)$ denotes the channel gain from BS $b$ to location $x$, including path loss attenuation, shadowing and other factors if any. Note, however, that fast fading is not considered here because the time scale for measuring $g_b(x)$ is assumed to be much larger. Accordingly, the signal to interference plus noise ratio (SINR) at location $x$ can be written as:

$$\Gamma(x, B) = \frac{E_b(x) \cdot b(x, B)}{\sum_{b \in B, b \neq b(x, B)} E_b(x) + \sigma^2}, \quad (16.1)$$

where $\sigma^2$ is noise power and $b(x, B)$ denotes the index of the BS at location $x$ that provides the highest signal strength, i.e., $b(x, B) = \arg \max_{b \in B} E_b(x)$. Following Shannon’s formula, spectral efficiency at location $x$ is given by $C(x, B) = \log_2 (1 + \Gamma(x, B))$.

**Area spectral efficiency.** The notion of the area spectral efficiency (ASE), firstly introduced in [35], is considered as a performance metric. It is defined as the summation of the spectral efficiency over the reference area $A$:

$$S(B) = \frac{\sum_{x \in X} C(x, B) \cdot Pr(x)}{|A|}, \quad [\text{bps/Hz/m}^2] \quad (16.2)$$

where $Pr(x)$ is the probability of the MT being at a specific location $x$; $X$ is the set of locations included in the area $A$ satisfying $Pr(x) > 0$ for all $x \in X \subset A$. For simplicity, in this section, the user distribution is assumed to
be homogeneous\(^4\) such that the discrete set \(\mathcal{X}\) is a rectangular lattice with a small grid size and the probability of each location is the same.

### 16.3.1.2 Problem formulation

Consider an area of interest \(A\) served by a wireless network operator whose access network consists of only macro BSs \(B_M\). Suppose that the maximum required ASE \(S_{th}^t\) at the peak time \(t^* = \arg \max_t S_{th}^t\) during a day \(t \in [t_0, t_0+D]\) almost approaches to the one that can be provided by turning on all the macro BSs \(B_M\), i.e., \(S(B_M) \approx S_{th}^t\). Thus, the operator wants to upgrade its access network by micro BSs which are considered as the cost-effective way of incrementally increasing capacity inside the initial macro cell deployment.

**Minimal deployment problem.** The objective is to find a minimal deployment of micro BSs (i.e., minimizing the additional power consumption) while providing \(\zeta \geq 1\) times higher ASE than before the upgrade. This can be mathematically formulated as the following optimization problem:

\[
\text{(MDP1)} \quad \min_{B_m} P_m \cdot |B_m| \tag{16.3}
\]

\[
s.t. \quad S(B_M \cup B_m) \geq \zeta \cdot S(B_M) = \zeta \cdot S_{th}^t, \tag{16.4}
\]

where \(P_m\) is the total operational power consumption of micro BSs. Note that (MDP1) can be also interpreted as CAPEX minimization. It is basically a combinatorial problem, and that makes it difficult to find an optimal solution, especially, when the number of candidate locations is large.

### 16.3.1.3 Key observations and algorithm

Several interesting observations\(^5\) can be made from various real topologies which help us to gain insight and develop an efficient algorithm. Given the area that is covered by the existing set of macro BSs, let us focus on the deployment of one new micro BS. The contour plot in Figure 16.2(a) shows how much ASE a micro BS can improve according to the location of deployment.

**Observation 16.1.** As long as a new micro BS is placed not too close to the one of existing BSs that would interfere with each other, ASE can be expected to increase before the upgrade. Especially, the ASE increment becomes large as the distance from macro BSs increase.

The wireless network operators are supposed to deploy a micro BS at the location where ASE can be improved. Therefore, only such locations are considered as candidate positions for the micro BS deployment:

\[
\forall k \in K, \quad S(B \cup \{k\}) > S(B), \tag{16.5}
\]

---

\(^4\) Please refer to [13] for the further results of the heterogeneous user distribution.

\(^5\) The results provided here is from the topology of Korea [36], similar trends could be observed in the other topologies [2,37] as well. Please refer to [13] for more details.
Let us denote by $A_{i<j}$ the set of locations that have better SINR from BS $i$ than $j$, and further denote by $A_{i=j}$ the set of boundaries having the same SINR from both BSs $i$ and $j$. Then, the set of locations covered by BS $k$ (or simply, coverage) can be written as:

$$A_k(B) = \{x | x \in A \text{ s.t. } b(x, B) = k\} = \bigcap_{b \in B, b \neq k} A_{k>b}. \quad (16.6)$$

Figure 16.2(b) examines how much area the micro BS can cover according to the location of deployment and investigates the correlation with ASE increment. As can be seen, ASE increment has a distinct tendency to increase with the coverage of micro BS. Interestingly, it becomes sharper (i.e., smaller variance) as the coverage increases. In such locations giving high increment, the most likely candidates for the deployment, coverage and ASE increment almost surely has a near-monotonic relationship.\(^6\)

Observation 16.2. The larger area can be covered by the new micro BS, the higher ASE increment is likely to be expected.

Motivated by this observation, the following monotone relationship is assumed to be hold.

$$|A_k(B \cup \{k\})| \geq |A_{k'}(B' \cup \{k'\})| \Rightarrow S(B \cup \{k\}) - S(B) \geq S(B' \cup \{k'\}) - S(B'), \quad (16.7)$$

where $k$ (or $k'$) is the index of the micro BS.

\(^6\)Results from monotone test, which randomly picks two points having positive ASE increments in Figure 16.2(b) and checks whether the slope between these points are positive or not, (90.4~97.0% depending on the topologies) also support the following observation.
These two observations are intuitively understandable. Consider the area covered by the micro BS far from existing macro BSs. Since the signals from the macro BSs are weak, the micro BS will provide the highest SINR to a large extent area. In addition to this large coverage, the area originally had low spectral efficiency, resulting in the high increment of ASE.

**Constant-factor approximation greedy deployment algorithm.** Prior to introducing a natural greedy algorithm for (MDP1), a real-valued set function $F : \mathcal{B}_m \rightarrow \mathbb{R}$ is defined as follows:

$$F(\mathcal{B}_m) = S(\mathcal{B}_M \cup \mathcal{B}_m) - S(\mathcal{B}_M), \quad (16.8)$$

which returns the ASE increment by additionally installing the micro BSs $\mathcal{B}_m$.

<table>
<thead>
<tr>
<th>Greedy deployment algorithm for (MDP1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Initialize $\mathcal{B}^{\text{greedy}}_m = \emptyset$</td>
</tr>
<tr>
<td>2: do while $S(\mathcal{B}_M \cup \mathcal{B}^{\text{greedy}}_m) &lt; \zeta \cdot S^{*}$</td>
</tr>
<tr>
<td>3: $k^* = \arg\max_{k \in \mathcal{K}} F(\mathcal{B}^{\text{greedy}}_m \cup {k}) - F(\mathcal{B}^{\text{greedy}}_m)$</td>
</tr>
<tr>
<td>4: $\mathcal{B}_m \leftarrow \mathcal{B}_m \cup {k}$</td>
</tr>
<tr>
<td>5: end do</td>
</tr>
</tbody>
</table>

The greedy algorithm starts with the empty set $\mathcal{B}^{\text{greedy}}_m = \emptyset$, and iteratively adds the micro BS location having the highest increment among the set of candidate locations $\mathcal{K}$ until ASE reaches a target value, i.e., satisfying the constraint (16.4).

**Theorem 16.1.** The ASE increment achieved by an optimal placement with the same number of micro BSs as the greedy algorithm cannot be more than a factor of $e/(e - 1)$ from the ASE increment achieved by the greedy algorithm.

$$\max_{|\mathcal{B}_m| = |\mathcal{B}^{\text{greedy}}_m|} F(\mathcal{B}_m) \leq \frac{e}{e - 1} F(\mathcal{B}^{\text{greedy}}_m), \quad (16.9)$$

where the constant $e$ is base of the natural logarithm.

This theorem can be proved following a similar procedure as in [38] taking into account the fact that the ASE increment function $F$ is submodular based on assumptions (16.5) and (16.7). For proof, please refer to [13]. In order to further reduce the computational complexity of the greedy algorithm, it is possible to check the locations near cell boundaries instead of all candidate locations.

**16.3.1.4 Numerical results**

For simulations, the deployment of macro BSs, shown in Figure 16.3(a), is considered. There are 10 macro BSs in $8 \times 8\text{km}^2$. In order to avoid edge effects, one the area of $5 \times 5\text{km}^2$ in the center is observed. To model the power consumptions, an affine power model [27] developed based on data sheets of
several GSM (Global System for Mobile Communications) and UMTS (Universal Mobile Telecommunications System) BSs with focus on component-level is adopted. Denote $p_{tx}^M$ and $p_{tx}^m$ are the average transmit (radiated) power for macro and micro BSs, respectively. Then the total operational power consumptions for macro and micro BSs, $P_M$ and $P_m$, are given by

\begin{align}
  P_M &= a_M \cdot p_{tx}^M + b_M, \\
  P_m &= a_m \cdot p_{tx}^m + b_m, 
\end{align}

where the relationship between the total operational power consumption and the transmit power is modeled in a linear fashion. The coefficient $a$ accounts for the power consumption that scales with the average radiated power due to amplifier and feeder losses as well as cooling of sites. The term $b$ denotes power offsets which are consumed independently of the average transmit power. These offsets are, amongst others, due to signal processing, battery backup, as well as site cooling. In Tab. 16.2, typical transmission and total operational powers for a macro BS with three sectors and two antennas per sector and a micro BS with a single sector with one omni-directional antenna are summarized.

On top of the deployment of macro BSs, Figure 16.3(a) illustrates the snapshots after 15 micro BSs additionally deployed by the proposed greedy deployment algorithm. It should be noted that the number associated with each of new micro BS are the order of the greedy deployment. As expected, the micro BSs tend to be placed in the boundaries of the cell because this makes the micro BSs cover the larger area, resulting in the more ASE increment.

Figure 16.3(b) shows the performance improvement according to the ad-
TABLE 16.2 Typical transmit power and total power consumption for macro BSs (3 sectors/2 antennas) and micro BSs (1 sector/1 antenna).

<table>
<thead>
<tr>
<th>BS type</th>
<th>$p_{tx}^M$ (W)</th>
<th>$p_{tx}^m$ (W)</th>
<th>$P_M$ (W)</th>
<th>$P_m$ (W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Macro BS</td>
<td>10 W</td>
<td>638 W</td>
<td>10 W</td>
<td>638 W</td>
</tr>
<tr>
<td>Micro BS</td>
<td>0.5 W</td>
<td>0.5 W</td>
<td>1 W</td>
<td>35 W</td>
</tr>
<tr>
<td></td>
<td>(22.6, 412.4W)</td>
<td>(22.6, 412.4W)</td>
<td>20 W</td>
<td>865 W</td>
</tr>
<tr>
<td></td>
<td>(22.6, 412.4W)</td>
<td>(22.6, 412.4W)</td>
<td>40 W</td>
<td>1317 W</td>
</tr>
<tr>
<td></td>
<td>(5.5, 32W)</td>
<td>(5.5, 32W)</td>
<td>0.5 W</td>
<td>35 W</td>
</tr>
<tr>
<td></td>
<td>(5.5, 32W)</td>
<td>(5.5, 32W)</td>
<td>1 W</td>
<td>38 W</td>
</tr>
<tr>
<td></td>
<td>(5.5, 32W)</td>
<td>(5.5, 32W)</td>
<td>2 W</td>
<td>43 W</td>
</tr>
</tbody>
</table>

TABLE 16.3 Additional total power consumption required for the target ASE increment. (from [13], ©2011 IEEE)

<table>
<thead>
<tr>
<th>BS type</th>
<th>Tx power</th>
<th>10%</th>
<th>15%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Macro</td>
<td>20 W</td>
<td>100W</td>
<td>9515W</td>
</tr>
<tr>
<td>Micro</td>
<td>2 W</td>
<td>645W</td>
<td>1476W</td>
</tr>
<tr>
<td>Micro</td>
<td>1 W</td>
<td>836W</td>
<td>1672W</td>
</tr>
<tr>
<td>Micro</td>
<td>0.5 W</td>
<td>1050W</td>
<td>2240W</td>
</tr>
</tbody>
</table>

ditional deployment of BSs having different transmission powers. Four types of BSs are considered: the macro BS with the transmit power of 20W and the micro BSs with transmit power of 0.5W, 1W and 2W, respectively. As can be clearly seen, there are diminishing returns on the normalized ASE increment. This is not only because the coverage of newly deployed BS will shrink but the amount of interference in the network increase as the number of BSs increases.

To meet the target ASE increment of 10%, while only five additional macro BSs are needed, 15, 22 or 30 micro BSs (three to six times more than macro BSs) are needed depending on their transmission powers. Nevertheless, the transmission power consumptions ($p_{tx}^M$ and $p_{tx}^m$) of additional micro BSs are much less than that of additional macro BSs. For example, while 100W is consumed by the macro BSs, only 30W, 22W, or 15W is consumed by the micro BSs. Reflecting the total power consumptions ($P_M$ and $P_m$), the advantage of micro BSs becomes more clear. Table 16.3 shows the required the additional total power consumptions for different target ASE increment. Compared to the case of macro BSs, deploying micro BSs can reduce more than 3kW and 6kW for the target ASE increments of 10% and 15%, respectively. This corresponds to about 70% energy savings.

16.3.2 Deployment structure: homogeneous vs. heterogeneous

In terms of energy consumption, it is often believed that network topologies with the high density deployment of small, low power BSs yield strong improvements compared to the low density deployment of few high power BSs [39]. However, when the static offset powers for site cooling, power supply, battery backup and so on are taken into account, it is no longer true, i.e., making cell sizes too small may lead to energy-inefficient solutions in
some cases. This section summarizes the results given in [14, 28] where it is considered a conventional \textit{homogeneous} deployment (macro BSs only) and an \textit{heterogeneous} deployment (macro + micro BSs). The performance of both deployments are compared under conditions where the inter-site distance $D$ varies in a predefined regular grid topology.

16.3.2.1 System model

\textbf{Deployment types.} The homogeneous macro network is simply modeled as a cloverleaf network layout consisting of three-sectorized macro BSs as shown in Figure 16.4(a). The layout of regular grid deployment can be characterized by an arbitrary inter-site distance $D$ varying in a certain range. Each macro BS serves an area denoted by $A$ (corresponding to the grey-shaded region). The area $A$ is referred to as \textit{cell}, whereas the geographic location of a BS is denoted as cell site or simply \textit{site}.

In the heterogeneous network, as depicted in Figure 16.4(b), a certain number of micro BSs are additionally placed in a specific location at the top of the above homogeneous macro network. For instance, if a micro BS is placed only at each corner marked in Figure 16.4(b) as a solid circle, then this can be considered as the particular case of single micro BS per macro cell.

\textbf{$\alpha$-percentile area spectral efficiency.} Earlier definition of ASE in (16.2) only considers the expectation of the achievable rates but is not concerned with the distribution of rates around in the system. In order to incorporate a fairness aspect into the notion of ASE, \textit{$\alpha$-percentile area spectral efficiency} is defined as the $\alpha$-quantile of the overall spectral efficiency in the reference cell divided by the cell size as follows:

$$S_\alpha = \frac{Q_\alpha[S]}{|A|}. \quad (16.11)$$
By scaling area spectral efficiency with the subcarrier bandwidth $B_{sc}$, the notion of $\alpha$-percentile area throughput per subcarrier is sometimes used as a more practical relevant measure, i.e., $T_{\alpha} = S_{\alpha} \cdot B_{sc}$.

**Area power consumption.** It may be unsuitable to observe only power consumption for comparing the networks with different site densities. This is because they may have different coverages. In order to assess the power consumption of the network relative to its size, the notion of area power consumption (APC) is introduced as the total power consumption in a reference cell divided by the corresponding reference area. With an average of $N$ micro BSs in a reference cell of size $A$, APC can be written by

$$P = \frac{P}{|A|} = \frac{P_M + N \cdot P_m}{|A|}, \quad [\text{Watt/m}^2].$$

Now let us investigate the effect of inter-site distance $D$ on APC, under a simplified propagation model [40] without considering shadowing and fading. The received signal strength decreases exponentially as the propagation distance $d$ increases, i.e., $P_{rx}/P_{tx} = K \cdot d^{-\beta}$, where $\beta$ is path loss exponent, and $K$ is a unitless constant which depends on the antenna characteristics. In order to guarantee the minimum signal level $P_{min}$ at the distance $R$ from the cell site, the required transmit power $P_{tx}$ can be given by

$$P_{tx} \propto P_{min} \cdot R^\beta.$$  \hspace{1cm} (16.13)

By substituting $D = \sqrt{3}R$ into (16.13), it can be obtained that the transmit power required for a certain coverage increases to $D^\lambda$. For the case of path loss exponent $\lambda = 2$, APC is not affected by the site distance $D$ because both numerator and denominator in (16.13) increase with $D^\lambda$. In general, for $\lambda > 2$, the following asymptotic results can be obtained [14]:

$$\lim_{D \to \infty} P(D) = \infty \quad \text{as well as} \quad \lim_{D \to 0} P(D) = \infty,$$

where the latter holds due to the nonzero constant terms $b_M$ and $b_m$ in (16.10). Thus, there exists an optimal inter-site distance $D^*$, which minimizes the area power consumption of the network.

**16.3.2.2 Problem formulation**

In order to find the minimal area power consumption required for a certain target area throughput, the following optimization problem is considered in [27,28]. The problem is to determine an optimal site distance $D^*$ that minimize the area power consumption while achieving a given target 10-percentile area throughput $T_{10}^{target}$ at least.

$$\min_{D \in [D, \infty)} P(D) \quad \text{s.t.} \quad T_{10}(D) \geq T_{10}^{target},$$

where $D$ is the minimum inter-site distance.
Remark 16.3.1. It is worthwhile mentioning that making the network denser (i.e., reducing \( D \)) helps to satisfy the target area throughput constraint because \( T_{10}(D) \) is strictly monotonically decreasing function\(^7\). However, as can be seen in (16.14), the area power consumption \( P(D) \) goes to infinity as \( D \) goes to zero. More specifically, it can be shown later in [14] that \( P(D) \) is a convex function. Thus, reducing \( D \) beyond a certain point will definitely increase the area power consumption.

16.3.2.3 Optimal area power consumption vs. target 10-percentile area throughput

Based on Remark 16.3.1, the following simple approach is able to solve (MDP2). For different numbers of micro BSs \( k = 0, 1, 2, \ldots \) per reference area \( A \) (\( k=0 \) means the case of macro BSs only),

Algorithm for (MDP2)

1. Determine the respective maximum inter-site distance \( \hat{D}_k \) that achieves the target 10-percentile area throughput \( T_{10\text{target}} \). Note that since the curves are strictly monotonically decreasing and thus all \( D_k \geq \hat{D}_k \) are feasible.

2. Determine the optimal distance \( D^*_k \) that minimize the area power consumption \( P(D) \).

3. If \( \hat{D}_k \) is larger than \( D^*_k \), then \( D^*_k \) is an optimal solution for problem (16.15). Otherwise, the optimal solution happens at the boundary, i.e., \( \hat{D}_k \).

Figure 16.5 shows the minimum area power consumption for different deployment strategies by varying the target 10-percentile area throughput. For each deployment strategy, the optimal power consumption remains constant until a certain target (i.e., the constraint in (16.15) is not active) and then increase almost linearly as the target area throughput increases. Note that for very low area throughput targets (< 40kbit/s/km\(^2\)), the pure macro BSs is most efficient. However, as the target area throughput increases, the benefit of micro BSs becomes more clear, i.e., the more micro BSs installed the more energy-efficiency the network becomes.

\(^7\)If \( D \) goes below a certain distance, say \( \overline{D} \), then the area throughput \( T_{10}(D) \) may decrease due to severe interference. However, such a region (i.e., \( D \in (0, \overline{D}) \) is not of our interest because it makes both the area throughput and power consumption worse.
16.4 Base Station On/Off and User Association

Recall that so far we have considered energy-aware deployment, which is an offline problem on the time-scale of months or even years. Once the BS deployment is done, the next problem is how to efficiently operate the BSs for energy conservation during the off-peak period. The solutions for the operation problem should be online distributed algorithms in order to be implemented in practical systems. There have been many efforts in literature [2,12,16,19–22,31,32] that deal with the operation problem. In particular, this section introduces our theoretical framework\(^8\) that encompasses dynamic BS operation and the related problem of user association together. For the total cost minimization problem that allows for a flexible tradeoff between flow-level performance and energy consumption, an efficient yet practical solution will be provided: an optimal energy-efficient user association policy and simple greedy-on and greedy-off BS switching on/off algorithms.

16.4.1 A quantitative case study

We shall begin with a motivational example [2] to provide an insight how much energy savings that could be obtained by turning off redundant BSs in cellular networks.

In order to quantitatively estimate potential energy savings, two sets of real data are combined together: the first is an anonymized temporal traffic trace of a period of one week from a cellular operator in a metropolitan area as

---

\(^8\)The key results of this section were drawn from [12].
shown in Figure 16.6; and the second is BS locations (total of 139 BSs in 128 locations in the area of 3.5×3.5km) from a central part of the city of Manchester, United Kingdom, as illustrated in Figure 16.7(a), obtained from a UK government-sponsored website [37]. As can be seen in Figure 16.7(b), which assumes an ideal circular range of 700 meters for each BS, there could be significant redundant overlap in cellular coverage. Putting together the temporal and spatial data-set, it has been shown that individual operators are able to save between about 8 to 22 percent of energy in such an urban deployment. Sharing BS resources together, a total reduction of about 29 percent can be achieved for the energy expended. This corresponds to about 78000-530000 kWh of total electricity savings for the region over a year (assuming the single BS power is between 800W and 1500W). According to [42], the cost in $ per KWh for electricity for industry is about 0.068 and for homes is about 0.116, so these translate to between about $5000 to more than $61000 of annual savings for the electricity bill for operating BSs. According to [43], these also translate to between 53.7 metric tons to 365 metric tons of CO$_2$ emissions. This is a substantial reduction in green-house gas emission as well as in the cost of operation.

16.4.2 System model and formulation

Consider a wireless cellular network in an area $\mathcal{R} \subset \mathbb{R}^2$ served by a set of BSs $\mathcal{B}$. Let $x \in \mathcal{R}$ denote a location and we use $i \in \mathcal{B}$ to index a typical $i$-th BS. Our focus is downlink communication, from BSs to MTs. File transfer requests are assumed to arrive following an inhomogeneous Poisson point process with arrival rate per unit area $\lambda(x)$ and file sizes which are independently
FIGURE 16.7 Real BS layout: (a) BS location data from a part of Manchester, UK (some BSs are collocated, in which case just one triangle is shown on the map), and (b) redundancy in the cellular coverage of Manchester area at 700m BS coverage. Note that this real dataset here is available online: [37] for the location of BSs [2] ©2011 IEEE.
Toward Energy-Efficient Operation of Base Stations in Cellular ...

distributed with mean $1/\mu(x)$ at location $x \in \mathcal{R}$, so the traffic load density is defined as $\gamma(x) = \frac{\lambda(x)}{\mu(x)} < \infty$. This captures spatial traffic variability. For example, a hot spot can be characterized by a high arrival rate and/or possibly large file sizes.

When the set of BSs $\mathcal{B}_{\text{on}}$ is turned on, the transmission rate of a user located at $x$ and served by BS $i \in \mathcal{B}_{\text{on}}$ is denoted by $c_i(x, \mathcal{B}_{\text{on}})$. For analytical tractability, it is assumed that $c_i(x, \mathcal{B}_{\text{on}})$ does not change over time, i.e., without considering fast fading or dynamic inter-cell interferences. Instead, $c_i(x, \mathcal{B}_{\text{on}})$ can be considered as a time-averaged transmission rate. This assumption is reasonable in the sense that the time scale of user association is much larger than the time scale of fast fading or dynamic inter-cell interferences. Hence, the inter-cell interference is considered static Gaussian-like noise, which is feasible under interference randomization or fractional frequency reuse [44–46]. It should be noted, however, that $c_i(x, \mathcal{B}_{\text{on}})$ is location-dependent but not necessarily determined by the distance from the BS.

The system-load density $\varrho_i(x, \mathcal{B}_{\text{on}})$ is then defined as $\varrho_i(x, \mathcal{B}_{\text{on}}) = \frac{\gamma(x)}{c_i(x, \mathcal{B}_{\text{on}})}$, which denotes the fraction of time required to deliver traffic load $\gamma(x)$ from BS $i \in \mathcal{B}_{\text{on}}$ to location $x$. A routing function $p_i(x)$ is further introduced to specify the probability that a flow at location $x$ is associated with BS $i$. Intuitively, $p_i(x)$ can be interpreted as the time fraction that a flow arrived at location $x$ is routed to BS $i$. We will see that, however, the optimal $p_i(x)$ will be either 1 or 0, i.e., deterministic routing (or user association) is the solution of our optimization problem, which is defined in (16.22). Then, the utilization of BS $i$, $0 \leq \rho_i \leq 1 - \epsilon$ is defined as

$$\rho_i = \int_\mathcal{R} \varrho_i(x, \mathcal{B}_{\text{on}}) p_i(x) dx, \forall i \in \mathcal{B}, \quad (16.16)$$

and where $\epsilon > 0$ is an arbitrarily small constant. We further denote the vectors containing the utilizations of all BSs by $\rho = (\rho_1, \cdots, \rho_{|\mathcal{B}|})$.

Two performance metrics are considered: the cost of flow-level performance such as file transfer delay and the cost of energy. The problem is to find an optimal set of active BSs ($\mathcal{B}_{\text{on}}$) and BS loads $\rho$ (i.e., user association) that minimize the total system cost function, given by

$$\min_{\mathcal{B}_{\text{on}}, \rho} \left\{ d_\alpha(\rho, \mathcal{B}_{\text{on}}) + \eta e(\rho, \mathcal{B}_{\text{on}}) \right\} \bigg| \rho \in \mathcal{F}(\mathcal{B}_{\text{on}}), \mathcal{B}_{\text{on}} \subseteq \mathcal{B}, \quad (16.17)$$

where $\mathcal{F}(\mathcal{B}_{\text{on}})$ is a feasible set of the load factor $\rho$ given the set of active BSs $\mathcal{B}_{\text{on}} \subseteq \mathcal{B}$ and $\eta \geq 0$ is the parameter that balances the tradeoff between the flow-level performance $d_\alpha(\rho, \mathcal{B}_{\text{on}})$ and the energy consumption $e(\rho, \mathcal{B}_{\text{on}})$, both of which will be explained shortly. In [12] we have shown that the feasible set $\mathcal{F}(\mathcal{B}_{\text{on}})$ is a convex set of $\rho$. The implication of $\eta$ is as follows; when $\eta$ is zero, the focus is only on the flow-level performance, however, as $\eta$ grows, more emphasis is given to energy conservation.

\[
    d_\alpha(\rho, B_{on}) = \begin{cases} 
        \sum_{i \in B_{on}} \frac{(1 - \rho_i)^{(1-\alpha)} - 1}{\alpha - 1}, & \alpha \neq 1 \\
        \sum_{i \in B_{on}} \log \left( \frac{1}{1 - \rho_i} \right), & \alpha = 1 
    \end{cases} 
\]  

where $\alpha \geq 0$ is a parameter specifying the desired degree of load balancing.

(ii) The cost function of energy: We use a general model for BSs consisting of two types of power consumptions: fixed power consumption and adaptive power consumption.

\[
    e(\rho, B_{on}) = \sum_{i \in B_{on}} \left[ (1 - q_i) \rho_i P_i + q_i P_i \right], 
\]

where $q_i \in [0, 1]$ is the portion of the fixed power consumption of BS $i$, and $P_i$ is the maximum operational power of BS $i$ when it is fully utilized, i.e., $\rho_i = 1$, which includes power consumptions of the transmit antennas, power amplifiers, cooling equipment, signal processor, battery backup, power supply, etc. Note that the first and second terms in (16.19) are the fixed and adaptive (i.e., proportional to the utilization) power consumptions, respectively.

When $q_i = 0$, BSs are assumed to consist of only energy-proportional devices. Such BSs would ideally consume no power when idle, and gradually consume more power as the activity level increases. This type of BSs will be referred to as energy-proportional BS. However, energy-proportional BSs are still far from reality because several devices in the BSs dissipate standby power while inactive. As an example, a class-A amplifier [48], which is a typical power amplifier for macro BSs and one of the most power consuming devices in BSs, has the maximum theoretical efficiency of 50%. This type of BSs, which consume the fixed power irrespective of its activity unless they are totally turned off, i.e., $q_i > 0$, will be referred to as non-energy-proportional BS. Note that when $q_i = 1$, this model can also capture a constant energy consumption model, which is widely used in many works in the literature [2, 8, 49–51].

Solving the general problem given in (16.17) is very challenging due to highly complex coupling of BS operation and user association. For analytical tractability, we shall make an assumption on time-scale separation that flow arrival and departure process and the corresponding user association process are much faster than the period on which the set of active BSs are determined.

Under this assumption, our general problem given in (16.17) can be decomposed into two subproblems, in which BS operation problem is solved at a slower time scale than user association problem. For any given set of active BSs $B_{on}$, the problem in (16.17) reduces to the following load balancing problem by ignoring the constant fixed power consumption term $\sum_{i \in B_{on}} q_i P_i$.

This problem can be also interpreted as user association problem because it
finds the optimal BS utilization vector $\rho$ by determining which BS should be associated to each MT.

**User association problem** [P-UA]:

$$\min_{\rho \in F(B_{on})} d_\alpha(\rho, B_{on}) + \eta \sum_{i \in B_{on}} (1 - q_i) \rho_i, \quad (16.20)$$

Then, the problem to solve is the following BS operation that finds the optimal set of active BSs $B_{on}$ on the longer time-scale.

**BS operation problem** [P-BO]:

$$\min_{B_{on} \subseteq B} G(B_{on}) + \eta \sum_{i \in B_{on}} q_i P_i, \quad (16.21)$$

where the function $G(B_{on})$ is defined as the obtaining optimal value from the underlying user association in (16.20), i.e.,

$$G(B_{on}) = \min_{\rho \in F(B_{on})} d_\alpha(\rho, B_{on}) + \eta \sum_{i \in B_{on}} (1 - q_i) \rho_i.$$

**Remark 16.4.1.** Note that [P-UA] and [P-BO] may have conflicting interest. [P-UA] tries to distribute traffic loads to improve the flow-level performance $d_\alpha(\rho, B_{on})$. On the other hand, to minimize $\sum_{i \in B_{on}} q_i P_i$, [P-BO] tries to concentrate traffic loads to a subset of BSs $B_{on}$ and turn off the other BSs.

Each of these two problems will be discussed in the consequent Sections 16.4.3 and 16.4.4.

### 16.4.3 Energy-efficient user association

Given the set of active BSs $B_{on}$, we focus on solving [P-UA] in (16.20), i.e., associating users with BSs in an energy-efficient manner, considering load-balancing. Let us denote the optimal BS load vector by $\rho^* = (\rho^*_1, \ldots, \rho^*_|B|)$, i.e., solution to the problem [P-UA], and further denote the optimal user association at location $x$ by $i^*(x)$. We now present the optimality condition of the problem that describes an optimal user association policy.

**Theorem 16.2.** [12] If the problem [P-UA] is feasible, then the optimal user association made by the MT located at $x$ to join BS $i^*(x)$ is given by

$$i^*(x) = \underset{j \in B_{on}}{\arg\max} \frac{c_j(x, B_{on})}{(1 - \rho^*_j)^{-\alpha} + \eta(1 - q_j)P_j}, \quad \forall x \in \mathcal{R}. \quad (16.22)$$

An online distributed algorithm that achieves the global optimum of [P-UA] in an iterative manner involves the following two parts.

**Mobile terminal:** At the start of the $k$-th iteration period, MTs receive BS loads $\rho^{(k)}$, e.g., through broadcast control messages from BSs. Then, a new

---

IEEE 802.16m facilitates this type of message structure [52, 53].
flow request from a MT located at $x$ simply selects the BS $i(x)$ using the deterministic rule given by

$$
i^{(k)}(x) = \arg\max_{j \in \mathcal{B}_{on}} \frac{c_j(x, \mathcal{B}_{on})}{(1 - \rho_j^{(k)})^{-\alpha} + \eta(1 - q_j)P_j}, \forall x \in \mathcal{R} \tag{16.23}
$$

**Base station:** During the $k$-th period, BSs measure their average utilizations after some period of time, i.e., when the system exhibits stationary performance. Then, BSs broadcast the average utilization vector $\rho^{(k+1)}$ for the next iteration.

This simple iteration provably converges to the global optimal point with a simple modification of the proof in [47].

**16.4.4 Traffic-driven base station on/off**

BSs typically consume large amounts of energy in power amplifier circuit, air conditioning unit, etc., irrespective of offered loads. As a simple intuition, the ratio of the overhead power to the total power ($= q_iP_i / (q_iP_i + (1 - q_i)\rho_iP_i)$) is close to 100% for small loads $\rho_i$. Thus, it would be definitely beneficial to turn off BSs with low activity, in conjunction with energy-efficient user association.

In this section, we show algorithms that reduce the energy consumption by solving the BS operation problem $[\text{P-BO}]$ in (16.21) determining the set of BSs that can be switched off.

Recall that the function $G(\mathcal{B}_{on}) = \min_{\rho \in \mathcal{F}(\mathcal{B}_{on})} d_\alpha(\rho, \mathcal{B}_{on}) + \eta \sum_{i \in \mathcal{B}_{on}} (1 - q_i)\rho_iP_i$ is obtained by the optimal user association policy in (16.22). The objective function in (16.20) is convex in $\rho$ given $\mathcal{B}_{on}$, but becomes a nonconvex and also discontinuous function when $\mathcal{B}_{on}$ is considered as a variable. Thus, this BS operation problem is a challenging combinatorial problem with $O(2^{|\mathcal{B}|})$ possible cases, which makes it very difficult to find an optimal solution through exhaustive search, especially, when the number of BSs is large. Thus, we show greedy-style heuristic algorithms, each of which has slightly different design rationale.

**16.4.4.1 Greedy turning on algorithm**

We first describe a greedy turning on algorithm, called GON, that iteratively finds BSs that have some benefit of delay reduction per their power usages.

**Greedy on algorithm (GON)**

1. Initialize $\mathcal{B}_{on} = \mathcal{B}_{init}$
2. while $\mathcal{B}_{on} \neq \mathcal{B}$
3. Calculate $M_{\text{GON}}(i) = \frac{G(\mathcal{B}_{on}) - G(\mathcal{B}_{on} \cup \{i\})}{q_iP_i}$, $\forall i \in \mathcal{B} \setminus \mathcal{B}_{on}$
4. Find the BS $i^\ast = \arg\max_{i \in \mathcal{B} \setminus \mathcal{B}_{on}} M_{\text{GON}}(i)$
5. if $M_{\text{GON}}(i^\ast) > \eta$, then $\mathcal{B}_{on} \leftarrow \mathcal{B}_{on} \cup \{i^\ast\}$
6. else, stop the algorithm.
7. end while
We introduce a metric $M_{\text{GON}}(i)$ for BS $i$ that represents the turn-on benefit per fixed power consumption for BS $i$. The GON starts with an initial set of BSs $\mathcal{B}_{\text{init}}$ and iteratively finds the best BS as a candidate among the set of inactive BSs $\mathcal{B} \setminus \mathcal{B}_{\text{on}}$ that has the highest $M_{\text{GON}}$ (step 4). Then, the algorithm finally adds the selected BS to the list of BSs to turn on, only if its metric is greater than $\eta$ (step 5), or stops otherwise. Note that the criterion $M_{\text{GON}}(i^*) > \eta$ is directly obtained from the condition that additionally turning on BS $i^*$ is beneficial (i.e., minimizing the total system cost), given by:

$$G(\mathcal{B}_{\text{on}}) + \eta \sum_{i \in \mathcal{B}_{\text{on}}} q_i P_i > G(\mathcal{B}_{\text{on}} \cup \{i^*\}) + \eta \left( \sum_{i \in \mathcal{B}_{\text{on}}} q_i P_i + q_{i^*} P_{i^*} \right).$$

### 16.4.4.2 Design rationale

Consider the following problem that is closely related to (16.21):

$$\min_{\mathcal{B}_{\text{on}} \subseteq \mathcal{B}} G(\mathcal{B}_{\text{on}}) \text{ subject to } \sum_{i \in \mathcal{B}_{\text{on}}} q_i P_i \leq C, \quad (16.24)$$

where we essentially move the power consumption cost in the objective function into the constraint of power consumption with some nonnegative budget $C$. For a given $\eta$, we can find $C = C(\eta)^{10}$, such that the same optimal solutions are achieved for (16.21) and (16.24), in which $\eta$ is interpreted as a Lagrange multiplier of the dual formulation of (16.24).

We transform (16.24) into:

$$\max_{\mathcal{A} \subseteq \mathcal{B} \setminus \mathcal{B}_{\text{init}}} H(\mathcal{A}) \text{ subject to } c(\mathcal{A}) = \sum_{i \in \mathcal{A}} c(i) \leq \tilde{C}, \quad (16.25)$$

where $\mathcal{A} = \mathcal{B}_{\text{on}} \setminus \mathcal{B}_{\text{init}}$, $H(\mathcal{A}) = G(\mathcal{B}_{\text{init}}) - G(\mathcal{B}_{\text{init}} \cup \mathcal{A}) = G(\mathcal{B}_{\text{init}}) - G(\mathcal{B}_{\text{on}})$, $c(i) = q_i P_i$ and $\tilde{C} = C - \sum_{i \in \mathcal{B}_{\text{init}}} c(i)$. If it can be shown that $H$ is a non-decreasing submodular set function, then a variant greedy algorithm of GON, where the only difference lies in the stopping condition (step 5), can be shown to achieve a constant factor $(1 - 1/e)$ approximation$^{11}$ of the optimal value of the problem (16.25).

Submodularity, informally, is an intuitive notion of diminishing returns, which states that adding an element to a small set helps more than adding that same element to a larger set. Formally, it is defined as follows.

**Definition 16.1.** A real-valued set function $H$, defined on subsets of a finite

---

$^{10}$ $C(\eta)$ is a non-increasing function of $\eta$.

$^{11}$ The submodular maximization problem (SMP) in (16.25) is in general a NP-hard problem. It has been proved that the greedy algorithm of SMP can achieve a constant factor $(1 - 1/e)$ approximation and its ratio is an optimal in the sense that no other polynomial algorithms with better constant approximation ratio exist. We refer the readers to [38,54,55] for the details.
set $S$ is called submodular if for all $A_1 \subseteq A_2 \subseteq S$ and for all $s \in S \setminus A_2$, if it satisfies that

$$H(A_1 \cup s) - H(A_1) \geq H(A_2 \cup s) - H(A_2).$$

(16.26)

### 16.4.4.3 Greedy turning off algorithm

One can think of another greedy algorithm, called GOFF (Greedy Off), which can be interpreted as the opposite of GON. The GOFF, unlike GON, starts from the entire BSs $B$ and finds a solution by iteratively removing the BS with the lowest turn-off detriment per fixed power consumption. Note that GOFF does not have the issue of choosing $B_{\text{init}}$.

**Greedy off algorithm (GOFF)**

1: Initialize $B_{\text{on}} = B$
2: while $B_{\text{on}} \neq \emptyset$
3: Calculate $M_{\text{GOFF}}(i) = \frac{G(B_{\text{on}} \setminus \{i\}) - G(B_{\text{on}})}{q_i P_i}, \forall i \in B_{\text{on}}$
4: $i^* = \arg \min_{i \in B_{\text{on}}} M_{\text{GOFF}}(i)$
5: if $M_{\text{GOFF}}(i^*) < \eta$, then $B_{\text{on}} \leftarrow B_{\text{on}} - \{i^*\}$,  
6: else, stop the algorithm.
7: end while

### 16.4.5 Discussion: GON and GOFF

#### 16.4.5.1 Interpretation

We now discuss the implication of the metric $M_{\text{GOFF}}(i)$ used in GOFF (or $M_{\text{GON}}(i)$ used in GON). Note that this metric can be interpreted as the network-wide impact per unit power cost. GOFF tends to choose and remove the BS that will bring the small impact on the network when turned off, whereas GON tends to choose and add the BS that will bring the large impact on the network when turned on. From the BS $i$’s perspective, the following internal and external factors, coupled in a complex manner each other, affect the metric $M_{\text{GOFF}}(i)$ and the choice of the final set $B_{\text{on}}$.

- **Internal factors of BS $i$:** Traffic loads imposed on BS $i$ is one of the dominant factors. Turning-off BS $i$ with high utilization will cause high impact on neighboring BSs because the large amount of traffic loads needs to be transferred (or handed over) to its neighboring BSs with potentially low signal strengths.

- **External factors around BS $i$:** When turning-off the BS $i$, its network-wide impact also depends on the neighboring environment, e.g., the number of,
the distance to, and the utilization of the neighboring BSs. As the number is small, the distance is far, and/or the utilization is high, we can expect high network-wide impact.

It should be noted that GON and GOFF require information about spatial system-load density $\varrho$ in order to compute their metrics. This is because $G$ inside the metrics depends on $\rho$, where $\rho$ is defined as the integral (or summation) of $\varrho$ over the space in (16.16). This information can be obtained by exchanging signaling messages or use the predetermined traffic profile over a period (e.g., one day) as in [8]. In the next subsection, we introduce other purely heuristic algorithms that are more operator-friendly in the sense that no signaling or measurement overhead is necessary, yet with possibly slight performance degradation compared to GON and GOFF.

### 16.4.5.2 Other heuristic algorithms

For other heuristic algorithms, one can exploit the distances between BSs, such that BSs distant from (resp. close to) each other are turned on (resp. off). Motivated by this fact, another algorithm is proposed, i.e., distance-based greedy heuristics based on GON and GOFF, called GON-DIST and GOFF-DIST, by simply modifying the metrics in the step 3 of GON and GOFF as follows:

$$M_{\text{GON-DIST}}(i) = \left[\prod_{j \in B_{\text{on}}} d(i,j)\right]^{1/|B_{\text{on}}|}, \quad \forall i \in B \setminus B_{\text{on}}, \quad (16.27)$$

$$M_{\text{GOFF-DIST}}(i) = \left[\prod_{j \in B_{\text{on}}, j \neq i} d(i,j)\right]^{1/|B_{\text{on}}|}, \quad \forall i \in B_{\text{on}}, \quad (16.28)$$

where the geometric mean of the distances to the other BSs are used for the distance metric.

Another greedy algorithm is called GOFF-UTIL$^{13}$, that chooses the most underutilized BS by modifying the metric in the step 3 of GOFF as follows:

$$M_{\text{GOFF-UTIL}}(i) = \rho_i, \quad \forall i \in B_{\text{on}}, \quad (16.29)$$

### 16.4.6 Numerical results

The energy-efficient user association and BS operation algorithms are verified through extensive simulations under various practical configurations. A network topology composed of five macro BSs and five micro BSs in $2 \times 2$ km$^2$ as shown in Figure 16.9 is considered for our simulations. A real 3G BS deployment topology consisting of heterogeneous environments (urban, suburban and rural areas) is also considered in subsections 16.4.6.2 in order to

$^{13}$Note that it does not make sense to have GON-UTIL policy since BSs that are turned off cannot have utilizations by the definition.
provide more realistic simulation results. Among several typical levels of maximum transmission powers for BSs given in [27], the intermediate values are used for the simulations, i.e., 43dBm and 30dBm for macro and micro BSs, respectively. Based on the linear relationship between transmission and operational power consumptions in Table 16.2, the maximum operational powers for BSs could be calculated, i.e., 865W and 38W for macro and micro BSs, respectively.

For the traffic model, we assume that file transfer request follows a Poisson point process with an arrival rate $\lambda(x)$. Each request has exactly one file that is log normally distributed with mean $1/\mu(x) = 100$ kbyte. In modeling propagation environment, the modified COST 231 path loss model with macro BS height $h = 32m$ and micro BS height $h = 12.5m$ is used. Other parameters for the simulations follow the suggestions in the IEEE 802.16m evaluation methodology document [56]. We consider the average delay experienced by a typical flow as our system performance metric, i.e., setting the degree of load balancing parameter as $\alpha = 2$ for the cost function of level performance. For the cost function of energy, the portion of fixed power consumption $q_i$ ranges between 0 and 1 to include several types of BSs from energy-proportional BSs to non-energy-proportional BSs.

### 16.4.6.1 Energy-delay tradeoff for energy-proportional BSs

Energy-proportional BSs ($q_i = 0$) is considered to investigate the performance obtained purely by the proposed energy-efficient user association algorithm. Figure 16.8 shows the energy-delay tradeoff curves by varying the energy-delay tradeoff parameter from $\eta = 10^{-5}$ to $10^{9}$ for the different values of arrival rate $\lambda(x)$. As can be expected, energy savings at the cost of delay increase when $\eta$ grows. The percentage of maximum energy saving (moving from $\eta = 10^{-5}$ to $\eta = 10^{9}$) is about 50%. This result is obtained under homogeneous traffic distribution, i.e., $\lambda(x) = \lambda$ for all $x \in \mathcal{R}$. Note that similar trends can be also observed in inhomogeneous traffic distribution, please refer to [12] for more details.

In order to examine the details of where these energy savings come from, Figs. 16.9 (a) and (b) illustrate the snapshots of cell coverage for two extreme cases: low $\eta = 10^{-5}$ and high $\eta = 10^{9}$. By comparing these two figures, we can clearly see that the micro BS, which is more energy-efficient than the macro BSs, will have large coverage for the case of high $\eta$ (i.e., giving more emphasis on conserving energy). In other words, more MTs are likely to be associated with and served by the energy-efficient micro BSs that are indexed by 6 to 10 in the figures. However, as the traffic loads are concentrated in the micro BSs, the large utilizations at the micro BSs will results in the increase of per-flow delay.
FIGURE 16.8 Energy-delay tradeoff for the case of energy-proportional BSs ($q_i = 0$) by varying the energy-delay tradeoff parameter $\eta = 10^{-5} \sim 10^0$. As $\eta$ increases, energy saving can be obtained at the cost of delay increase. (from [12], ©2011 IEEE)

(a) Low $\eta = 10^{-5}$

(b) High $\eta = 10^0$

FIGURE 16.9 Snapshots of cell coverage by the energy-efficient user association algorithm. As $\eta$ increases, the energy-efficient micro BS (indexed by 6 to 10) will have larger coverage [12] ©2011 IEEE
FIGURE 16.10 Real 3G BS deployment map (30 BSs in 20 x 10 km$^2$). (from [12], ©2011 IEEE)

FIGURE 16.11 Energy-delay tradeoff of different algorithms for the case of non-energy-proportional BSs ($q_i = 0.5$) by varying the energy-delay tradeoff parameter from $\eta = 10^{-5}$ to $10^0$. While greedy algorithms perform close to the optimal solution when $\eta$ is small, there is a gap when $\eta$ is large. Especially, GON-DIST and GOFF-DIST have large performance gaps under the inhomogeneous traffic distribution [12] ©2011 IEEE.
16.4.6.2 Energy-delay tradeoff for non-energy-proportional BSs

As a numerical example, non-energy-proportional BSs \( q_i = 0.5 \) is considered and the performance is obtained by both the proposed energy-efficient user association and BS operation algorithms. Consider GON, GOFF, GON-DIST, GOFF-DIST and GOFF-UTIL as the BS operation algorithm, and compare their performance with the optimal solution obtained by exhaustive search. Figure 16.10 depicts the map of BS layout [36] that is chosen for more realistic simulations. It is a part of real 3G network operated by one of the major mobile network operators in Korea (the source has to be anonymized to preserve confidentiality). There are totally 30 BSs within \( 20 \times 10 \) km\(^2\) rectangular area. Particularly, this partial map is chosen to include a scenario that three environments (urban, suburban and rural) coexist together.

Figure 16.11(a) shows the energy-delay tradeoff curves of different algorithms by varying the energy-delay tradeoff parameter \( \eta = 10^{-5} \sim 10^0 \) in urban area with the homogeneous traffic distribution of \( \lambda(x) = 10^{-4} \) for all \( x \in \mathcal{R} \). This offered load corresponds to about 10% of BSs utilizations when all BSs are turned on. Recall the real traffic measurement report [2] showing that the time fraction when the traffic is below 10% of peak during the day is about 30% in weekdays and about 45% in weekends. As can be seen from Figure 16.11(a), energy is saved at the cost of per-flow delay increase. The greedy algorithms perform close to the optimal solution when \( \eta \) is small. For example, from \( \eta = 10^{-5} \), up to \( \eta = 10^{-2} \) for GON and GOFF, \( \eta = 10^{-3} \) for GOFF-UTIL, and \( \eta = 10^{-4} \) for GON-DIST and GOFF-DIST, respectively, the solution is very close to the optimal (i.e., Exhaustive ≥ GON = GOFF ≥ GOFF-UTIL ≥ GON-DIST = GOFF-DIST). However, there is a performance gap when \( \eta \) becomes large.

Figure 16.11(b) shows the energy-delay tradeoff curves under the inhomogeneous traffic distribution. As an example of inhomogeneous traffic loads, a linearly increasing load along the diagonal direction from right bottom to left top is considered. They are normalized over the space so as to have the same amount of total traffic as the homogeneous traffic loads have. Similar tradeoff curve can be observed in inhomogeneous traffic distribution as well. The greedy algorithms GON, GOFF and GOFF-UTIL still perform close to the optimal solution up to \( \eta = 10^{-3} \), however, GON-DIST and GOFF-DIST start to deviate much from the optimal solution after \( \eta = 10^{-4.5} \).

There is a reason why such GON-DIST and GOFF-DIST based on the distance do not work well under the inhomogeneous traffic distribution: turning on (resp. off) the BSs distant from (resp. close to) each other is no longer reasonable because the BSs distant from (resp. close to) each other but located in the area of low (resp. high) traffic loads may not be beneficial to turn on (resp. off). It is noteworthy that GOFF-UTIL has the almost comparable performance to that of GON and GOFF under both homogeneous and inhomogeneous traffic distribution. This is a desirable observation for wire-
TABLE 16.4 Effect of BS density on energy savings.

<table>
<thead>
<tr>
<th>BS Density</th>
<th>Urban</th>
<th>Suburban</th>
<th>Rural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum Energy</td>
<td>Greedy-on</td>
<td>66.2%</td>
<td>32.4%</td>
</tr>
<tr>
<td>Savings</td>
<td>Greedy-off</td>
<td>72.8%</td>
<td>32.5%</td>
</tr>
</tbody>
</table>

less network operators who do want to implement a simple yet efficient BS operation algorithm.

16.4.6.3 Effect of BS Density on Energy Savings

We also examine how much energy savings can be achieved according to the density of BS deployment. To this end, we vary the BS density by adopting the BS topology from three different environments in Figure 16.10: urban (15 BSs in $4.5 \times 4.5$ km\(^2\)), suburban (15 BSs in $12 \times 6$ km\(^2\)) and rural (8 BSs in $9 \times 9$ km\(^2\)) areas. Table 16.4 shows the effect of BS deployment density on the maximum energy savings. As expected, much energy saving in the urban and suburban environments, but almost no or low energy saving in the rural environment. This is because the degradation of signal strength is significant in the rural environment when traffic loads are transferred from the switched-off BS to neighboring BSs.

16.5 Other Issues and Discussions

16.5.1 Open problems and other directions towards green cellular networks

Even though recent papers have started to investigate ways to increase the energy efficiency of cellular networks, we are still at an early stage of this research. Therefore, we would like to encourage the community to give it greater attention by pointing out several open problems and other directions including new paradigm and architecture towards green cellular networks.

Multi-operator cooperation. There is also room for greater energy saving if different operators can pool their BSs together and accept each other’s traffic as roaming traffic when the BSs are switched off [2,57]. This kind of operator cooperation to ensure pooled coverage is particularly helpful in metropolitan areas where each operator sets up a dense deployment. Such a scenario has also previously been examined by Marsan et al. [57], who consider the simple setting of a single cell with two operators and show that considerable energy savings can be obtained in such a setting. In realizing cooperation, physical sharing may not be a substantial concern. In many cases, in urban areas, operator’s BSs tend to be closely situated, or even co-located. The main challenge
is the complexity in network operation, with respect to issues such as cross-
operator user authentication and billing, introduced by a novel fine-granularity
roaming. This yields a rich set of problems to be investigated from a game the-
etric and economic perspective. Under what conditions would self-interested
operators agree to cooperate with others? What kind of profit-sharing agree-
ments will provide an adequate incentive for all participants? It also needs to
be examined whether such operator agreements can be potentially abused to
create oligopolies that hurt customers and how these can be ameliorated in
the interest of green operation that offer other benefits to society.

**Component-level deactivating or deceleration.** As discussed in section
16.4, shutting down some underutilized BSs has substantial potential to ob-
tain energy savings. However, despite very low traffic, there is a certain group
of BSs that should not be turned off not to create coverage holes.\(^{14}\) Further,
switching BSs off may bring degradation in user experience, for example, more
uplink power consumption for file uploading. Due to these technical challenges
of entirely turning off BSs, component-level techniques \([58, 59]\) has been re-
cently proposed, which are more conservative yet can reduce conserve energy
consumption effectively. In \([58]\), the authors adjusted the number of radio units
and transmit antennas in BSs by deactivating them during low load period.
In \([59]\), the authors considered cooperation of a component-level deceleration
technique into BS operation. This technique, called dynamic voltage frequency
scaling (or simply speed-scaling) \([60]\), allows a central processing unit (CPU)
to adapt its speed for energy conservation based on incoming processing de-
mand.\(^{15}\) They also investigated its impact on the design of network protocols
in cellular networks.

**Coordinated multi-point transmission and reception (CoMP).** The
primary constraint of green cellular operation is to preserve coverage and ser-
vice quality although certain BSs are turned off during low activity period. One
can potentially increase transmission power when some BSs are turned off to
increase the coverage area of the remaining BSs. Another novel approach that
can play a role in maintaining coverage is the use of coordinated multi-point
transmission and reception (CoMP) being developed in the context of LTE-
Advanced systems \([61]\). The basic idea in CoMP, a macro-diversity scheme
based on network MIMO (multiple-input and multiple-output), is to improve

---

\(^{14}\) According to the classification in \([22]\), there are two different classes of BSs: class-A
(non-redundant) and class-B (redundant) BSs. The former are the BSs that have been
deployed to provide the coverage to the users, usually at an early stage of the deployment.
The latter BSs that have been additionally deployed to mainly increase the capacity of the
network at a later stage. While class-B BSs may be turned off, class-A BSs cannot because
their contribution to the coverage is particularly high.

\(^{15}\) Note that examples of in-BS processing are increasingly abundant from flow classifica-
tion, signal processing (OFDM modulation, coding, etc.) to even security and multimedia
conversion. It is also worthwhile mentioning that DVFS lowers heat dissipation as well. As
a consequence, it can reduce the power consumption in cooling equipment (not to mention
in CPU) contributing to a considerable amount of total energy consumption, where this
exerting influence is often linear.
the signal quality by coordinating the transmission and reception at neighboring cells. CoMP can improve the coverage of high data rates and/or the cell-edge throughput, which may allow the systems to turn off more BSs [34]. However, overheads such as message exchanges between cooperating BSs via backhaul and complex signal processing at the BSs are the major challenges to overcome. Taking into account additional energy consumption for such overheads, it is essential to investigate the net-energy-efficiency of CoMP [33].

**Multi-hop relaying.** The future cellular networks such as LTE-Advanced systems have very challenging requirements because there is a growing demand for coverage and capacity. Instead of deploying additional BSs to meet such requirements, it is sometimes beneficial to install relays. The rationale of relaying is that the path loss between BSs and MTs can be reduced significantly by breaking a long weak single-hop link into shorter strong multi-hop links. They have low power consumption due to their small size and have a wireless backhaul, which enables deployment flexibility and is much cheaper than deploying new BSs with a fixed wired backhaul. Furthermore, multi-hop relaying will clearly be useful to ensure that the dynamic shutting down of BSs, while saving energy, does not leave coverage holes. All these reasons allow relays to become one of candidate technologies for green cellular networks. Examples of recent studies on relaying in cellular networks include: an investigation into how relays can increase the spatial reuse and therefore provide the required data rates while reducing energy consumption [62], a joint optimization of relay placement and sleep/active problem [17], and a relay caching mechanism [63] to improve the energy efficiency, especially for multimedia applications.

**Transmission power control.** The current and future cellular networks are expected to become more heterogeneous with a mixture of macro and small cells, where interference is a major obstacle that can impair the potential gain of small cells and its pattern is highly diverse [64]. As the number of small cells increases, the portion of users at cell edges also grows, resulting in the increase of the number of users suffering from low throughput due to severe interference. To mitigate interference, starting from a static interference management (IM) having specific reuse patterns, e.g., a traditional frequency reuse, fractional frequency reuse (FFR) [65], or its variation [45, 66], the research community has paid extensive research attention to dynamic IM algorithms that performs dynamic BS power control depending on the scheduled users over the multiple neighboring cells [67–73]. Their proposals are more performance-oriented in the sense that with a given power budget the objective is to maximize the efficiency (typically measured by throughput or utility). However, in fact with significantly less amount of powers, it may be possible to achieve a near-optimal performance mainly because the environment is typically interference-limited and/or operated under the high SINR regime, where, for example, the achievable rate for users under such regimes slowly decreases as SINR decreases. Recently, in [11], Kwak et al. studied the impact of sharing the power budget, and concluded that sharing the power
budget more spatially and temporally enables us to save significant amount of energy.

**Radio over fiber.** Radio over fiber (RoF) refers to a technology where analog radio signal is directly transmitted over an optical fiber link for wireless access deployment [74,75]. To elaborate, a base station server is connected to antennas at remote sites via optical fiber, and each remote site serves as a (small) cell. RoF is well aligned with the upcoming wireless access network in several aspects such as network capacity, energy efficiency and economic perspective. To increase the network capacity, the cell size needs to be small, and then the base station can be made simple by moving the main functionality, i.e., the scheduler, into a base station server. Then, the base station server controls a number of (small) base stations. Hence, the base station at the remote site can be made of only antennas and power amplifiers without complex functionalities. When a base station server controls a number of remote sites, it further contributes to making multi-cell cooperation easier, for example, in realizing inter-cell coordinated scheduling and interference management. From greening perspective, RoF is energy efficient because removing the scheduling and digital processing part of the base station can significantly reduce the energy consumption at each cell site by reducing the circuit power. From economic perspective, RoF lowers CAPEX of base station deployment and OPEX of base station operation; simple infrastructural architecture reduces CAPEX and low electricity bill thanks to energy efficiency decreases OPEX. The way of realizing RoF depends on how to transport radio signal over optical fiber; RoF is classified either intermediate frequency (IF) over fiber or radio frequency (RF) over fiber depending on whether up/down conversion at the remote site is required (IF over fiber) or not (RF over fiber).

### 16.5.2 Ongoing projects and consortia

This section summarizes major international projects and consortia pertaining to green networks conducted by research institutes, universities and industries in recent years.

- **GreenTouch** [76], led by Bell Labs, is a worldwide research consortium comprised of 15 founder members from leading research institutions and university (e.g., INRIA, CEA-LETI, Stanford, MIT, etc.) and world’s largest equipment manufacturers and network operators (e.g., Samsung, Freescale, AT&T, China Mobile, Telefonica, etc.). Their goal is to increase energy efficiency in networks by a factor of 1,000 within five years by designing fundamentally new network architectures and creating enabling technologies on which they are based.

- **EARTH** (Energy-Aware Radio and neTwork tecHnology) [77,78] is a major European research project starting in 2010 with 15 partners from 10 countries, focusing on the energy efficiency in the next generation wireless access networks. The goal is to achieve at least 50% reduction within
the next two-and-a-half years to provide directly applicable solutions for environmentally sustainable and cost efficient broadband wireless services.

- **GreenRadio** [79] is one of the three working areas under Mobile VCE (Virtual Centre of Excellence). The goal of GreenRadio is to secure 100x reduction in energy requirements for high data rate services by investigating both architectural and individual technological approaches.

- **Cool Silicon** [80] is a research cluster supported by German government since September 2008. It currently focuses on three lead projects, CoolReader, CoolSensornet and CoolComputing, all aiming at massively increasing the energy efficiency in computing, cellular communications, and sensor networks.

---

**Acknowledgments**

We would like to acknowledge collaborations on green cellular network papers with Eunsung Oh at Korea Institute of Energy Technology Evaluation and Planning, Xin Liu at the University of California at Davis and Zhisheng Niu at Tsinghua University. We thank Mahasweta Sarkar and Santosh Nagaraj at San Diego State University for their helpful discussions. We also would like to thank Fred Richter, Albrecht J. Fehske, Patrick Marsch, and Gerhard P. Fettweis at the Dresden University of Technology, Germany for their generosity in allowing us to use their energy-aware deployment results in Section 16.3.2.

This work was supported in part by NSF NetSE grant 1017881, CNS-1116874, the Sogang University Research Grant of 2011, and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2011-0015042).

---

**Bibliography**


Toward Energy-Efficient Operation of Base Stations in Cellular ... 469


Green Communications: Theoretical Fundamentals, Algorithms ...


[56] IEEE 802.16m-08/004r5: IEEE 802.16m Evaluation Methodology Document (EMD), IEEE Std. 802.16m, 2009.


---

**Author contact information**

Kyuho Son is with Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089, USA, Email: kyuhoson@usc.edu. Hongseok Kim is with Department of Electronic Engineering, Sogang University, Seoul 121-742, Korea, Email: hongseok@sogang.ac.kr. Yung Yi is with Department of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 305-701, Korea, Email: yiyung@kaist.edu. Bhaskar Krishnamachari is with Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089, USA, Email: bkrishna@usc.edu.