Abstract

NEW METHODS OF EFFICIENT
BASE STATION CONTROL FOR
GREEN WIRELESS COMMUNICATIONS

SEONWOOK KIM
DEPARTMENT OF ELECTRICAL ENGINEERING AND
COMPUTER SCIENCE
COLLEGE OF ENGINEERING
SEOUL NATIONAL UNIVERSITY

This dissertation reports a study on developing new methods of efficient *base station* (BS) control for green wireless communications. The BS control schemes may be classified into three different types depending on the time scale — hours based, minutes based, and milli-seconds based. Specifically, hours basis pertains to determining which BSs to switch on or off; minutes basis pertains to *user equipment* (UE) association; and milli-seconds basis pertains to UE scheduling and radio resource allocation. For system model, the dissertation considers two different

models — heterogeneous networks composed of cellular networks and *wireless local area networks* (WLANs), and cellular networks adopting *orthogonal frequency division multiple access* (OFDMA) with *carrier aggregation* (CA). By combining each system model with a pertinent BS control scheme, the dissertation presents three new methods for green wireless communications: 1) BS switching on/off and UE association in heterogeneous networks, 2) optimal radio resource allocation in heterogeneous networks, and 3) energy efficient UE scheduling for CA in OFDMA based cellular networks.

The first part of the dissertation presents an algorithm that performs BS switching-on/off and UE association jointly in heterogeneous networks composed of cellular networks and WLANs. It first formulates a general problem which minimizes the total cost function which is designed to balance the energy consumption of overall network and the revenue of cellular networks. Given that the time scale for determining the set of active BSs is much larger than that for UE association, the problem may be decomposed into a UE association algorithm and a BS switching-on/off algorithm, and then an optimal UE association policy may be devised for the UE association problem. Since BS switching-on/off problem is a challenging combinatorial problem, two heuristic algorithms are proposed based on the total cost function and the density of access points of WLANs within the coverage of each BS, respectively. According to simulations, the two heuristic algorithms turn out to considerably reduce energy consumption when compared with the case where all the BSs are always turned on.

The second part of the dissertation presents an energy-per-bit minimized radio resource allocation scheme in heterogeneous networks equipped with multi-homing capability which connects to different wireless interfaces simultaneously. Specifically, an optimization problem is formulated for the objective of minimizing the energy-per-bit which takes a form of nonlinear fractional programming. Then, a parametric optimization problem is derived out of that fractional programming and the original problem is solved by using a double-loop iteration method. In each iteration, the optimal resource allocation policy is derived by applying Lagrangian duality and an efficient dual update method. In addition, suboptimal resource allocation algorithms are developed by using the properties of the optimal resource allocation policy. Simulation results reveal that the optimal allocation algorithm improves energy efficiency significantly over the existing resource allocation algorithms designed for homogeneous networks and its performance is superior to suboptimal algorithms in reducing energy consumption as well as in enhancing network energy efficiency.

The third part of the dissertation presents an energy efficient scheduling algorithm for CA in OFDMA based wireless networks. In support of this, the energy efficiency is newly defined as the ratio of the time-averaged downlink data rate and the time-averaged power consumption of the UE, which is important especially for battery-constrained UEs. Then, a component carrier and resource block allocation problem is formulated such that the proportional fairness of the energy efficiency is guaranteed. Since it is very complicated to determine the optimal solution, a low-complexity energy-efficient scheduling algorithm is developed, which approaches the optimal algorithm. Simulation results demonstrate that the proposed scheduling scheme performs close to the optimal scheme and outperforms the existing scheduling schemes for CA.

Keywords: Green communication, heterogeneous networks, carrier aggregation (CA), BS switching-on/off, user association, radio resource allocation.

Student Number: 2007-20941

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Green wireless communications target at energy-efficient operation of wireless communication systems to combat against ever-increasing energy consumption in cellular networks. The number of mobile subscribers has rapidly increased for the last decade years and recent forecasts predict that overall mobile data traffic will increase 13-fold in 2017 [2]. Such heavy traffic demands will cause a significant increase of energy consumption in radio access networks, which is directly linked to an increase of *operational expenditure* (OPEX) of the wireless network operators. Furthermore, the drastic increase in mobile data traffic volume may bring more greenhouse gas emissions. Therefore, due to the economic and ecological benefits, the energy efficient communication has recently received much attention and several international research projects devote great efforts to it [3, 4].

Motivated by the necessity for reducing the energy consumption, a large amount of work has been reported on improving the energy efficiency in all available components of cellular networks, e.g., *base stations* (BSs), *user equipments* (UEs), and the core network [5, 6]. Specifically, it is of great importance to enhance the energy efficiency of the BSs because recent surveys reported that the BSs consume 60–80% of energy used in cellular networks [7]. Energy efficient communication is also important for battery constrained UEs because UEs operate based on battery power in the most practical cases.

To meet the increasing demand for mobile data traffic, wireless network operators are interested in integrating *wireless local area networks* (WLANs) and cellular networks such as 3GPP-LTE and WiMAX. In general, cellular data offloading is cost-effective and energy-efficient since WLANs can offer high data rate at lower energy consumption than cellular networks. However, excessive data offloading leads to revenue reduction of the cellular operators. In addition, using WLANs is not always more energy-efficient than using cellular networks for a UE which is near to the BS but not to the *access point* (AP) of WLANs. Therefore, it is necessary to study how to efficiently combine heterogeneous networks, considering the economic factors and the energy efficiency of the overall network.

Another way to satisfy high demand for wireless data is to combine multiple spectrums in licensed band, which is referred as to *carrier aggregation* (CA). The LTE-Advanced standard [8] describes that CA supports up to 100 MHz system bandwidth by aggregating up to five *component carriers* (CCs) of 20 MHz and allows a UE to use one or multiple CCs simultaneously. Thus, if a UE uses two CCs simultaneously to communicate with the BS, it can get roughly two times higher capacity than the single CC case. However, in case that the UE continuously uses CA,

Table 1.1: BS control strategies based on time scale.

Time scale	Hours	Minutes	Milli-seconds
BS control strategy	BS switching-on/off	UE association	UE scheduling, radio resource allocation

its energy consumption rapidly increases, which can be fatal to the battery powered UEs.

The dissertation investigates BS control strategies for greening wireless networks under the above two different system models. As shown in Table 1.1, the BS control strategies can be classified based on their operational time-scale. It is reasonable to consider the communication system to operate on different time scale because different type of parameters is needed in different time scale. It is assumed that the BS operation is determined at hours level, the UE association is performed at minutes level, and the UE scheduling or radio resource allocation is conducted at milli-seconds level. Specifically, three BS control scenarios are considered by combining the system model and time-scale based operation; the first scenario is a joint algorithm for BS operation and UE association in heterogeneous networks, the second scenario is a radio resource allocation algorithm in heterogeneous networks, and the third scenario is a UE scheduling algorithm for CA. Note that all those scenarios commonly aim at green communication. The detailed explanation on each scenario is as follows:

1) BS switching-on/off and UE association: Large amount of energy can be saved by switching-off under-utilized BSs at night since the traffic load exhibits wide-range fluctuations in time [9]. However, if some BSs are switched-off, the UEs connecting to those switched-off BSs have to be newly associated with other switched-on BSs. Hence, several hours based BS on/off operation should be considered jointly with several minutes based UE association.

In the case of heterogeneous networks, cellular data offloading through WLANs leads to not only the decrease of BSs' utilization but also the reduction of total energy consumption. However, if a majority of cellular data is served by APs, instead of BSs, the revenue of the cellular operators may considerably diminish, which makes it hard to maintain healthy cellular network and good cellular services. Hence, there is a tradeoff relationship between the energy consumption and the network revenue. Therefore, a joint algorithm for BS on/off and UE association needs to be developed to balance the energy consumption and the network revenue.

Based on time scale, the dissertation deals with the joint algorithm by decomposing it into a UE association algorithm and a BS switching-on/off algorithm. For the UE association problem, it presents an optimal UE association policy and, for the BS switching-on/off problem, it presents two heuristic algorithms based on the total cost function and the density of access points of WLANs within the coverage of each BS, respectively.

2) Radio resource allocation: Once the BS on/off operation is determined, several milli-seconds based radio resource allocation may be performed in heterogeneous networks consisting of an *orthogonal frequency division multiple access* (OFDMA) based cellular network and multiple WLANs. The dissertation presents a downlink resource allocation algorithm which minimizes the energy-per-bit of the entire network under minimum data rate requirements. It assumes that all UEs support multi-homing access, which means that each UE can communicate with the BS of cellular network and the AP of a WLAN simultaneously. Multi-homing access enables more energy-efficient communications by exploiting network diversity using multiple interfaces.

Allowing traffic sharing and interworking between cellular network and WLAN, the dissertation formulates the resource allocation problem as a nonlinear fractional programming and then, solves it by using a double-loop iteration method. In addition, it develops suboptimal resource allocation algorithms by using the properties of the optimal resource allocation policy.

3) UE scheduling: Another several milli-seconds based BS control, UE scheduling may be conducted in OFDMA based wireless networks with CA. The CA may be classified into three types based on how to configure multiple CCs, among which, inter-band CA merging CCs within different bands is of main interest. In the inter-band CA case, it may not be desirable to allocate multiple CCs simultaneously to a UE. It is because the UE has to turn on additional *radio frequency* (RF) elements for the concurrent transmission, which significantly reduces the battery lifetime of the UE. Hence, it is necessary to devise a novel UE scheduling for CA considering UE power consumption.

In OFDMA based wireless networks with CA, the BS allocates CCs and *resource blocks* (RBs) to appropriate UEs at each time slot. The energy efficiency may be defined as the ratio of the time-averaged downlink data rate and the UE power consumption. The proportional fairness criterion may be considered to balance the tradeoff between maximizing the overall energy efficiency and preserving some degree of fairness. The dissertation presents a CC and RB scheduling algorithm to guarantee the proportional fairness of the energy efficiency in the above sense. Since it is very complicated to determine the optimal solution, it develops a low-complexity energy-efficient scheduling algorithm which turns out to approach the optimal scheme.

The remainder of this dissertation is organized as follows: Chapter 2 investigates how to determine a set of active BSs and how to perform UE association in heterogeneous networks. Chapter 3 presents a radio resource allocation scheme optimizing the overall energy efficiency with multi-homing capability. Chapter 4 provides an energy efficiency proportional fair scheduler in OFDMA based networks with CA. Finally, Chapter 5 concludes the dissertation.

Chapter 2

A Joint Algorithm for Base Station Operation and User Association in Heterogeneous Networks

2.1 Introduction

Reducing energy consumption of *base stations* (BSs) is of great importance in wireless communication systems. According to recent surveys, for operating a cellular network including BSs, *user equipments* (UEs), and the core network, about 80% of the total energy is consumed at the BS sites [10, 11]. Today, BSs are densely deployed to support peak time traffic load [12]. However, since the traffic load fluctuates over time [9], the utilization of BSs may be very low during off-peak hours, e.g., at night. Therefore, it is possible to save large amount of energy by turning off under-utilized BSs during such off-peak hours [13–17]. However, turning on/off of BSs results in a new problem of UE association. If some BSs are turned off, the UEs communicating with those turned-off BSs need to be newly associated with other BSs. Therefore, turning on/off of BSs should be considered jointly with UEs' association problem.

In the next generation wireless networks, integration of heterogeneous networks, including *wireless local area networks* (WLANs) and cellular networks such as 3GPP-LTE and WiMAX, is considered a promising architecture to meet the increasing demand for mobile data traffic. Since WLANs can offer high data rate at lower energy consumption than cellular networks do, offloading cellular data through WLANs leads to the reduction of BSs' energy consumption. However, if a majority of cellular data is offloaded through WLANs, the revenue of the cellular operators may drop exceedingly, thus making it hard to maintain good cellular services and healthy cellular network. Hence, there exists a tradeoff between reducing the energy consumption and maintaining the network revenue at a reasonable level. In [15], the authors propose efficient BS control mechanisms for energy-delay tradeoff. However, different from [15], we investigate the tradeoff between energy consumption and network revenue in heterogeneous networks.

In the following, we will discuss how to determine an optimal set of active BSs and how to perform optimal UE association in heterogeneous networks consisting of cellular networks and WLANs. We will first develop a total cost minimization problem that reflects the energy-revenue tradeoff. We will assume that the intervals of turning on/off a BS is in the order of hours while UE association is determined in much shorter time scale (e.g., several minutes). Based on this assumption, we decompose the problem into two subproblems, namely, UE association problem and BS switching-on/off problem, to make the total cost minimization problem tractable. For the UE association problem, we devise an optimal UE association policy and provide insights on the characteristics of the optimal policy. As to the BS switching-on/off problem, since the problem has a combinatorial form which is very difficult to solve, we develop two greedy BS operation algorithms which enable us to solve the problem in polynomial time.

The rest part of this chapter is organized as follows. We first describe the system model in Section 2.2. Then, we formulate a total cost function minimization problem in Section 2.3. We propose a UE association algorithm and heuristic BS operation algorithms in Section 2.4 and Section 2.5, respectively. Finally, we evaluate the performance of the proposed algorithms in Section 2.6.

2.2 System Model

We consider the downlink transmission in a *time division multiple access* (TDMA) based heterogeneous network consisting of a set $\mathcal{B}$ of BSs of cellular networks and a set $\mathcal{A}$ of APs of WLANs as shown in Fig. 2.1. We consider a region $\mathcal{K} \subset \mathbb{R}^2$ that is served by all the BSs and APs. We assume that file transfer requests follow a heterogeneous Poisson point process with an arrival rate $\lambda(x)$ and a mean file size $1/\mu(x)$ that are independently distributed at location $x(\in\mathcal{K})$. We define the traffic

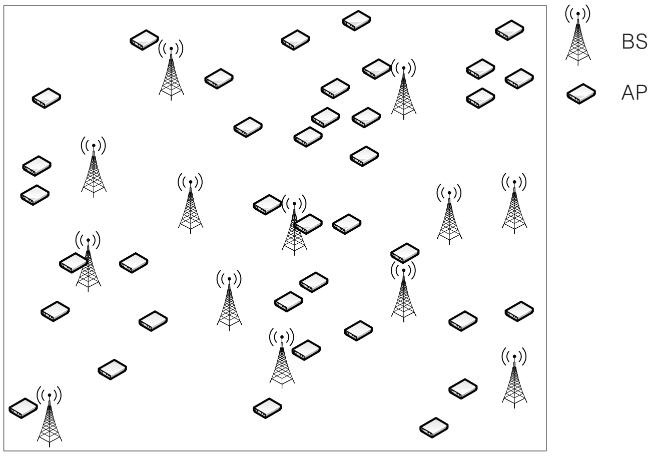


Figure 2.1: Heterogeneous networks with cellular networks and WLANs.

intensity at location x as $\gamma(x) \triangleq \lambda(x)/\mu(x)$.

We denote the set of active BSs by $\mathcal{B}_{\text{on}}$. We assume that each AP works on non-overlapping channels, so that no interference exists among APs.[1] In addition, since WLANs operate in an unlicensed band, no APs interfere with the BSs. Then, the maximum achievable data rate of a UE served by the ith BS or AP at location x is given by

$$C_i(x) = \begin{cases} \log_2\left(1 + \dfrac{P_i^{\text{out}} g_i(x)}{\sigma^2 + \sum\limits_{j \in \mathcal{B}_{\text{on}}, j \neq i} P_j^{\text{out}} g_j(x)}\right), i \in \mathcal{B}_{\text{on}}, \\ \log_2\left(1 + \dfrac{P_i^{\text{out}} g_i(x)}{\sigma^2}\right), i \in \mathcal{A}, \end{cases} \tag{2.1}$$

where P_i^{out} is the transmission (or radiated) power of the ith BS or AP, $g_i(x)$ is the

[1]For example, 4 and 19 non-overlapping channels are available in 2.4 GHz and 5 GHz bands, respectively, in Korea, to make it possible to assign non-overlapping channels to neighboring APs.

channel gain of a UE at location x from the ith BS or AP, and σ^2 is the noise power. We define the system load density as $\tau_i(x) \triangleq \gamma(x)/C_i(x)$, which is interpreted as the time fraction required to deliver the traffic intensity $\gamma(x)$ from the ith BS or AP to UEs at location x.

We denote the region that can be served by any AP in $\mathcal{A}$ by

$$\mathcal{K}_{\mathcal{A}} = \left\{ x \in \mathcal{K} | P_i^{\text{out}} g_i(x)/\sigma^2 \geq \text{SNR}_{\text{th}}, \forall i \in \mathcal{A} \right\}, \tag{2.2}$$

where SNR_{th} is the *signal-to-noise ratio* (SNR) threshold that determines whether a UE at location x is in the coverage of any AP. UEs in $\mathcal{K}_{\mathcal{A}}$ can communicate with either BS or AP while other UEs can communicate only with BS. We define a feasible set ψ of load vector $\boldsymbol{\rho}(= (\rho_0, \rho_1, \cdots, \rho_{|\mathcal{A}\cup\mathcal{B}_{\text{on}}|}))$ for both BSs and APs by

$$\begin{aligned}\psi = \big\{ \boldsymbol{\rho} \big| \rho_0 = \sum_{i \in \mathcal{B}_{\text{on}}} \textstyle\int_{\mathcal{K}_{\mathcal{A}}} \gamma(x) u_i(x)\, dx, \\ \rho_i = \textstyle\int_{\mathcal{K}} \tau_i(x) u_i(x)\, dx,\ \forall i \in (\mathcal{A} \cup \mathcal{B}_{\text{on}}), \\ 0 \leq \rho_i \leq 1 - \epsilon,\ \forall i \in (\mathcal{A} \cup \mathcal{B}_{\text{on}}), \\ \sum_{i \in (\mathcal{A} \cup \mathcal{B}_{\text{on}})} u_i(x) = 1,\ \forall x \in \mathcal{K}, \\ 0 \leq u_i(x) \leq 1, i \in (\mathcal{A} \cup \mathcal{B}_{\text{on}}) \text{ and } \forall x \in \mathcal{K} \big\},\end{aligned} \tag{2.3}$$

where ρ_0 indicates the amount of traffic load that is served by BSs even if the traffic could be served by APs, $\rho_i (i \neq 0)$ indicates the amount of offered load of the ith BS or AP, $u_i(x)$ indicates the probability that a UE at location x is associated with the ith BS or AP, and ϵ is an arbitrarily small positive constant. Then, the feasible load vector $\boldsymbol{\rho}$ is characterized by the UE association vector $\boldsymbol{u}(x) = (u_1(x), \cdots, u_{|\mathcal{A}\cup\mathcal{B}_{\text{on}}|}(x))$.

2.3 Problem Formulation

Based on the system model, we consider a cost function minimization problem in the heterogeneous networks. The cost function is formulated as the difference of total energy consumption $E(\mathcal{A}, \mathcal{B}_{\text{on}}, \boldsymbol{\rho})$ and network revenue $R(\rho_0)$. Accordingly, minimizing the cost function is equivalent to minimizing the total energy consumption and maximizing the revenue jointly. Then, the optimization problem which finds the optimal set $\mathcal{B}_{\text{on}}$ of active BSs and the load vector $\boldsymbol{\rho}$ (i.e., UE association) is given by

$$\underset{\mathcal{B}_{\text{on}}, \boldsymbol{\rho}}{\text{minimize}} \quad \varphi(\mathcal{A}, \mathcal{B}_{\text{on}}, \boldsymbol{\rho}) \, (= E(\mathcal{A}, \mathcal{B}_{\text{on}}, \boldsymbol{\rho}) - \eta R(\rho_0)) \tag{2.4a}$$

$$\text{subject to} \quad \boldsymbol{\rho} \in \psi, \tag{2.4b}$$

where $\varphi(\mathcal{A}, \mathcal{B}_{\text{on}}, \boldsymbol{\rho})$ is the total cost function and η (≥ 0) is a parameter that balances the tradeoff between the energy consumption and revenue. When η is zero, it only considers the energy consumption but, as η increases, it pays more weight on the revenue side.

We assume that the revenue is generated when the BSs deliver the data that can be offloaded by APs. Hence, the cost function of revenue is proportional to ρ_0, taking the expression

$$R(\rho_0) = \rho_0 \delta, \tag{2.5}$$

where δ indicates the revenue per unit traffic load.

Total energy consumption of the ith BS may be divided into two parts: Static and dynamic. Static power $P_i^{\text{BS,sta}}$ is consumed irrespective of whether the BS is in transmit mode, whereas dynamic power $P_i^{\text{BS,dyn}}$ is consumed only when it is in transmit mode. Similarly, when the jth AP is in idle state, constant power $P_j^{\text{AP,idle}}$ is consumed but, when in transmit state, $P_j^{\text{AP,tx}}$ is consumed. Hence, total energy consumption is given by

$$\begin{aligned} E(\mathcal{A}, \mathcal{B}_{\text{on}}, \boldsymbol{\rho}) = & \sum_{i \in \mathcal{B}_{\text{on}}} \left(\rho_i P_i^{\text{BS,dyn}} + P_i^{\text{BS,sta}} \right) \\ & + \sum_{j \in \mathcal{A}} \left(\rho_j P_j^{\text{AP,tx}} + (1 - \rho_j) P_j^{\text{AP,idle}} \right), \end{aligned} \tag{2.6}$$

where the first term indicates the energy consumed by the BSs and the second term consumed by the APs.

The optimization problem given in (2.4) is very difficult to solve because the process of switching-on/off BSs is highly coupled with the process of associating the related UEs to new BSs. In order to make the problem tractable, we decompose the problem into two subproblems. If we assume that the time scale for BS operation is much larger than that for UE association, we may decompose the overall problem into a combination of BS operation process and UE association process as described in Algorithm 1. The decomposed process indicates that a BS switching-on/off algorithm is performed at every T_h hours and a UE association algorithm at every T_m minutes. We additionally perform BS switching-on/off algorithm if some BSs get over-loaded due to a sudden increase of traffic load. We will deal the two algorithms in two subsequent sections: UE association algorithm in Section 2.4 and BS operation algorithm in Section 2.5, respectively.

Algorithm 1 BS operation and UE association algorithm

1: Every T_h hours,
2: Execute BS switching-on/off algorithm;
3: Every T_m minutes,
4: Execute UE association;
5: **if** Some BSs are over-loaded, **then**
6: Execute BS switching-on/off algorithm;
7: Execute UE association;
8: **end if**

2.4 UE Association Algorithm

We first solve the optimization problem in (2.4) for a given set $\mathcal{B}_{\text{on}}$ of active BSs. Then, we present an optimal UE association policy that each UE associates with an AP or a BS depending on energy efficiency and revenue.

LEMMA 2.1 *For a given set $\mathcal{B}_{on}$, the problem in (2.4) is a convex optimization problem.*

Proof : It is trivial to prove that the feasible set ψ is convex by applying the proof in [15]. The objective function is also convex since it is linear with respect to the load vector $\boldsymbol{\rho}$. Therefore, the problem in (2.4) becomes a problem of minimizing the convex function under convex constraints, which is a standard convex optimization problem. ■

LEMMA 2.2 *It is the optimal UE association policy to associate a UE located at x with the i^*th AP or BS satisfying*

$$i^* = \operatorname*{argmin}_{j \in (\mathcal{A} \cup \mathcal{B}_{\text{on}})} \pi_j(x), \tag{2.7}$$

where

$$\pi_j(x) = \begin{cases} \dfrac{P_j^{\text{BS,dyn}}}{C_j(x)} - \eta\delta, & x \in \mathcal{K}_{\mathcal{A}},\ \forall j \in \mathcal{B}_{\text{on}}, \\ \dfrac{P_j^{\text{AP,tx}} - P_j^{\text{AP,idle}}}{C_j(x)}, & x \in \mathcal{K}_{\mathcal{A}},\ \forall j \in \mathcal{A}, \\ \dfrac{P_j^{\text{BS,dyn}}}{C_j(x)}, & x \notin \mathcal{K}_{\mathcal{A}},\ \forall j \in \mathcal{B}_{\text{on}}. \end{cases} \tag{2.8}$$

Proof : As the problem (2.4) is a convex optimization problem by Lemma 2.1, to prove optimality of the association, it is sufficient to prove the following inequality [18]:

$$\nabla\varphi(\mathcal{A}, \mathcal{B}_{\text{on}}, \boldsymbol{\rho}) \cdot \triangle\boldsymbol{\rho}^* \geq 0, \tag{2.9}$$

where $\nabla\varphi(\mathcal{A}, \mathcal{B}_{\text{on}}, \boldsymbol{\rho})$ indicates the gradient vector of $\varphi(\cdot)$ with respect to $\boldsymbol{\rho}$ and $\triangle\boldsymbol{\rho}^* = \boldsymbol{\rho} - \boldsymbol{\rho}^*$ for the optimal load vector $\boldsymbol{\rho}^*$. The above inner product may be expanded as follows:

$$\begin{aligned} &\nabla\varphi(\mathcal{A}, \mathcal{B}_{\text{on}}, \boldsymbol{\rho}) \cdot \triangle\boldsymbol{\rho}^* \\ &= \sum_{i=0}^{|\mathcal{A} \cup \mathcal{B}_{\text{on}}|} \frac{\partial \nabla\varphi(\rho_i)}{\partial \rho_i}(\rho_i - \rho_i^*) \\ &= \sum_{i \in \mathcal{B}_{\text{on}}} P_i^{\text{BS,dyn}}(\rho_i - \rho_i^*) + \sum_{j \in \mathcal{A}} (P_j^{\text{AP,tx}} - P_j^{\text{AP,idle}})(\rho_j - \rho_j^*) - \eta\delta(\rho_0 - \rho_0^*). \end{aligned} \tag{2.10}$$

Let $\boldsymbol{u}^*(x)$ denote the optimal association vector for the optimal $\boldsymbol{\rho}^*$. Then, by (2.7), the optimal association probability is determined by

$$u_i^*(x) = \begin{cases} 1, \text{ if } i = \underset{j \in (\mathcal{A} \cup \mathcal{B}_{\text{on}})}{\text{argmin}}\ \pi_j(x) \\ 0, \text{ otherwise.} \end{cases} \tag{2.11}$$

By applying (2.8) and (2.11) to (2.9), we obtain

$$\begin{aligned} \nabla\varphi(\mathcal{A}, \mathcal{B}_{\text{on}}, \boldsymbol{\rho}) \cdot \triangle\boldsymbol{\rho}^* &= \sum_{i \in \mathcal{B}_{\text{on}}} P_i^{\text{BS,dyn}} \int_{\mathcal{K}} \frac{\gamma(x)}{C_i(x)} (u_i(x) - u_i^*(x))\, dx \\ &+ \sum_{j \in \mathcal{A}} (P_j^{\text{AP,tx}} - P_j^{\text{AP,idle}}) \int_{\mathcal{K}} \frac{\gamma(x)}{C_j(x)} (u_j(x) - u_j^*(x))\, dx \\ &- \eta\delta \sum_{i \in \mathcal{B}_{\text{on}}} \int_{\mathcal{K}_{\mathcal{A}}} \gamma(x)(u_i(x) - u_i^*(x))\, dx \\ &= \int_{\mathcal{K}_{\mathcal{A}}} \gamma(x) \sum_{i \in \mathcal{B}_{\text{on}}} \left(\frac{P_i^{\text{BS,dyn}}}{C_i(x)} - \eta\delta \right) (u_i(x) - u_i^*(x))\, dx \\ &+ \int_{\mathcal{K}_{\mathcal{A}}} \gamma(x) \sum_{j \in \mathcal{A}} \frac{(P_j^{\text{AP,tx}} - P_j^{\text{AP,idle}})}{C_j(x)} (u_j(x) - u_j^*(x))\, dx \\ &+ \int_{\mathcal{K} \setminus \mathcal{K}_{\mathcal{A}}} \gamma(x) \sum_{i \in \mathcal{B}_{\text{on}}} \frac{P_i^{\text{BS,dyn}}}{C_i(x)} (u_i(x) - u_i^*(x))\, dx \geq 0. \end{aligned}$$

■

By (2.11), the optimal $u_i(x)$ is either 0 or 1, which means that the solution of the problem in (2.4) yields a deterministic UE association. When $\eta = 0$, the deterministic rule in Lemma 2.2 implies that each UE associates with a single BS or a single AP that minimizes the Joule per bit, i.e., the most energy-efficient. Even when $\eta > 0$, a UE not in $\mathcal{K}_{\mathcal{A}}$ associates with the most energy-efficient BS in $\mathcal{B}_{\text{on}}$ but a UE within $\mathcal{K}_{\mathcal{A}}$ considers the network revenue δ as well, thus choosing BS more frequently than AP as η increases. Especially when η is sufficiently large (to be specific, $\eta > \max_{j \in \mathcal{B}_{\text{on}}, x \in \mathcal{K}_{\mathcal{A}}} \left\{ P_j^{\text{BS,dyn}} / C_j(x) \right\} / \delta$), a UE within $\mathcal{K}_{\mathcal{A}}$ always communicates with the most energy-efficient BS instead of AP. Note that the optimal UE association policy is independent of $\gamma(x)$ despite its inhomogeneity. Therefore, as long as the load vector $\boldsymbol{\rho}$ is feasible, each UE simply selects the most appropriate single BS or AP based on energy-efficiency and network revenue, regardless of the amount of traffic load.

2.5 BS Switching-on/off Algorithm

We now determine the optimal set $\mathcal{B}_{\text{on}}$ to complete the BS energy saving algorithm. The problem in (2.4) is a convex optimization problem for a given set $\mathcal{B}_{\text{on}}$, but it becomes a combinatorial problem when $\mathcal{B}_{\text{on}}$ is regarded as a variable. In this case, an optimal solution may be found through exhaustive search among $\mathcal{O}(2^{|\mathcal{B}|})$ possible cases, and hence, the problem becomes intractable if $|\mathcal{B}|$ is large. Therefore, we devise two heuristic algorithms that enable to solve the problem in polynomial time by a greedy approach.

We set up a switching-on/off algorithm, assuming that the BSs near the over-loaded BSs are turned on, that BSs are turned off in the order of the largest turn-off benefit among all the active BSs. We formalize the procedure of this switching-on/off algorithm as Algorithm 2. We initialize $\mathcal{B}_{\text{on}}$ and define $\mathcal{B}_{\text{vic}\{i\}}$ as the set of BSs in the vicinity of BS i. If the optimal load vector obtained by Lemma 2.2 for a given $\mathcal{B}_{\text{on}}$ is not feasible, we add $\mathcal{B}_{\text{vic}\{i\}}$ of the over-loaded BS i (i.e., $\rho_i^* > 1 - \epsilon$)

Algorithm 2 BS switching-on/off algorithm

1: Initialize, $\mathcal{B}_{\text{on}} = \mathcal{B}_{\text{init}}$;
2: **while** $\boldsymbol{\rho}^* \notin \psi$ **do**
3:     **for** $i \in \mathcal{B}_{\text{on}}$ **do**
4:         **if** $\rho_i^* > 1 - \epsilon$ **then**
5:             $\mathcal{B}_{\text{on}} \leftarrow \mathcal{B}_{\text{on}} \cup \mathcal{B}_{\text{vic}\{i\}}$;
6:         **end if**
7:     **end for**
8: **end while**
9: Calculate $D(i)$, $\forall i \in \mathcal{B}_{\text{on}}$;
10: Find BS $i^* \leftarrow \text{argmax}_{j \in \mathcal{B}_{\text{on}}} D(j)$;
11: **if** $\boldsymbol{\rho}^* \in \psi$ **then**
12:     $\mathcal{B}_{\text{on}} \leftarrow \mathcal{B}_{\text{on}} - \{i^*\}$, go to Step 9;
13: **else**
14:     Stop the algorithm;
15: **end if**

to $\mathcal{B}_{\text{on}}$. We define a determinant function D and calculate $D(i)$ of each BS i in $\mathcal{B}_{\text{on}}$. Then, we remove the BS i, whose $D(i)$ is the largest, from $\mathcal{B}_{\text{on}}$ if the resulting load vector $\boldsymbol{\rho}^*$ is feasible and iterate this removal process for the BS with the next largest $D(i)$ until the resulting $\boldsymbol{\rho}^*$ becomes infeasible.

As to the determinant function $D(i)$, we consider two heuristic algorithms below: *Cost function based* (CFB) algorithm and *AP density based* (ADB) algorithm, with the metrics D_{CFB} and D_{ADB}, respectively. The computational complexity of

CFB algorithm is in the order of $\mathcal{O}(|\mathcal{B}|^2)$ and that of ADB algorithm is in the order of $\mathcal{O}(|\mathcal{B}|)$.

2.5.1 Cost Function Based (CFB) Algorithm

The CFB algorithm is designed to turn off the BS that yields the maximum cost gain if it is turned off. Then, the determinant function of the CFB algorithm takes the form

$$D_{\text{CFB}}(i) = E(\mathcal{A}, \mathcal{B}_{\text{on}}, \hat{\boldsymbol{\rho}}^*) - E(\mathcal{A}, \mathcal{B}_{\text{on}} \backslash i, \tilde{\boldsymbol{\rho}}^*) - \eta(R(\hat{\boldsymbol{\rho}}^*) - R(\tilde{\boldsymbol{\rho}}^*)), \tag{2.12}$$

where $\hat{\boldsymbol{\rho}}^*$ and $\tilde{\boldsymbol{\rho}}^*$ are the optimal load vectors for $\mathcal{B}_{\text{on}}$ and $\mathcal{B}_{\text{on}} \backslash i$, respectively. Note that the metric $D_{\text{CFB}}(i)$ can be determined by the objective function of the problem in (2.4).

2.5.2 AP Density Based (ADB) Algorithm

In case that APs are not uniformly distributed, the traffic of the BS, which contains the largest number of APs within its coverage, could be most effectively offloaded through the APs. Thus, the ADB algorithm is designed to turn off the BS with the largest number of UEs, who are associated with APs, within its coverage. If we denote by $\mathcal{K}_i$ the coverage area of BS i, or

$$\mathcal{K}_i = \left\{ x \in \mathcal{K} \middle| i = \underset{j \in \mathcal{B}_{\text{on}}}{\text{argmin}}\, \pi_j(x) \right\}, \; \forall i \in \mathcal{B}_{\text{on}}, \tag{2.13}$$

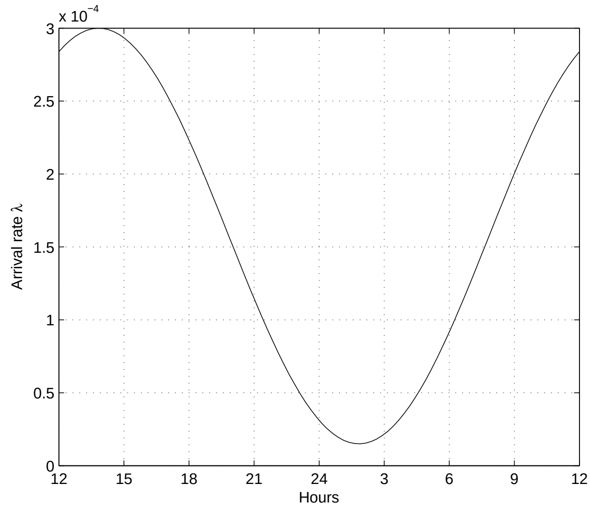


Figure 2.2: Daily traffic profile.

then the determinant function of the ADB algorithm takes the expression

$$D_{\text{ADB}}(i) = \frac{|\mathcal{K}_{\mathcal{A}} \cap \mathcal{K}_i|}{|\mathcal{K}_i|}. \tag{2.14}$$

2.6 Performance Evaluation

We have performed simulations to examine the performance of the proposed algorithms. We consider a network composed of 9 BSs[2] and 100 APs[3] deployed 1 km $\times$ 1 km area. The BSs are located at 400 m intervals and the APs are randomly distributed. We set SNR_{th} to 9 dB; δ to \$5 per gigabyte. For the traffic model, we

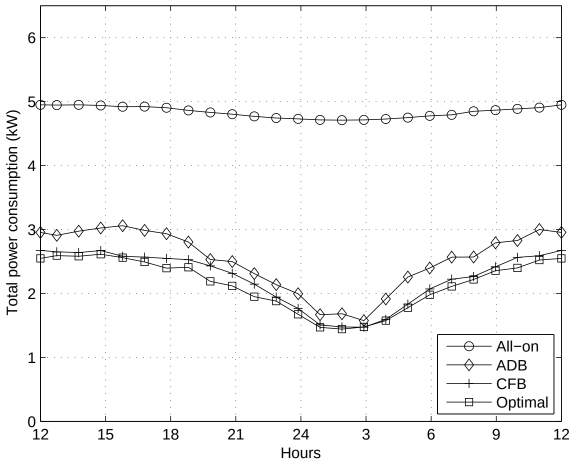
[2]http://www.sitefinder.ofcom.org.uk/search

[3]http://www.optimum.net/WiFi/Find

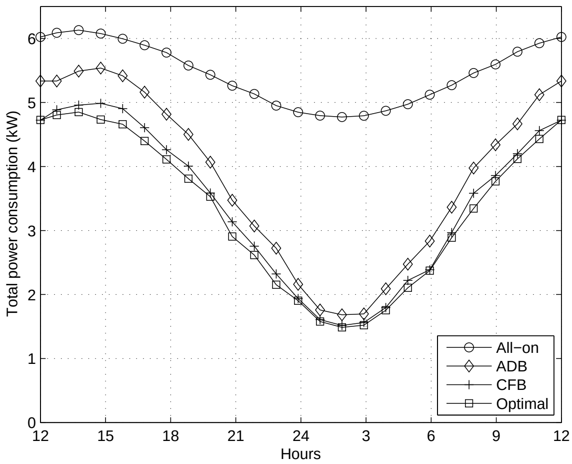
assumed that the arrival rate $\lambda(x)$ of file transfer requests is homogeneous for all $x \in \mathcal{K}$ and varies only with time. We also assumed that each file transfer request has just one file that is log-normal distributed with a mean $1/\mu(x) = 50$ kbytes for all $x \in \mathcal{K}$. We set the variables related to power consumption as follows: $P_i^{\text{out}} = 20$ W; $P_i^{\text{BS,dyn}} = 420$ W; $P_i^{\text{BS,sta}} = 460$ W; $P_j^{\text{out}} = 100$ mW; $P_j^{\text{AP,tx}} = 10.1$ W; $P_j^{\text{AP,idle}} = 9.2$ W, for $i \in \mathcal{B}$ and $j \in \mathcal{A}$ [19], [20].

We conducted simulations based on dynamic traffic load distribution in time domain. In order to obtain time-dependent results, we assumed that the daily traffic profile repeats periodically, i.e., we neglected the effect of weekend on traffic load. We modeled the daily behavior as a simple sinusoidal curve shown in Fig. 2.2, which is not far from the observation reported in [21]. As discussed in Algorithm 1, we performed Algorithm 2 at every T_h hours and optimal UE association at every T_m minutes. Only when the optimal load vector is not feasible, i.e., $\boldsymbol{\rho}^* \notin \psi$, we additionally performed BS switching-on/off algorithm. We set T_h to 1 and T_m to 15.

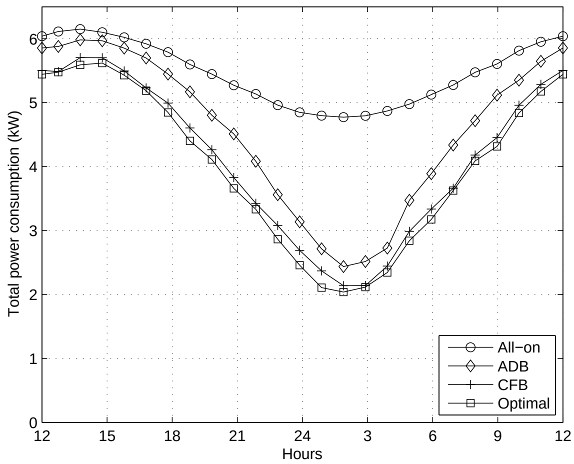
Fig. 2.3 depicts the resulting total power consumption of the all-on, ADB, CFB, and optimal schemes with respect to the flow of time during a day. The all-on scheme represents that all the BSs are switched on and UE association is performed in every T_m minutes. The optimal scheme represents the optimal solution of BS switching-on/off problem obtained through exhaustive search, instead of applying Algorithm 2. We observe that the optimal scheme outperforms all other schemes, and the ADB and CFB schemes perform close to the optimal scheme with the CFB scheme outperforming the ADB scheme.


Total power consumption (kW)
Hours
All−on
ADB
CFB
Optimal
0
1
2
3
4
5
6
12
15
18
21
24
3
6
9
12

(a)


Total power consumption (kW)
Hours
All−on
ADB
CFB
Optimal
0
1
2
3
4
5
6
12
15
18
21
24
3
6
9
12

(b)



(c)

Figure 2.3: Performance comparison of all-on, ADB, CFB, and the optimal schemes in terms of total power consumption: (a) $\eta = 0$, (b) $\eta = 10$, and (c) $\eta = 1000$.

We can check the effect of network revenue on power consumption by comparing Figs. 2.3(a)–2.3(c).[4] In the case of Fig. 2.3(a) with $\eta = 0$, which focuses on reducing the energy consumption by neglecting the revenue, we observe that the proposed schemes significantly curtail the energy consumption even in peak times. It happens because most of data generated in $\mathcal{K}_{\mathcal{A}}$ are offloaded through APs, which are more energy-efficient than BSs. As η increases, with more weight on network revenue, the number of UEs in $\mathcal{K}_{\mathcal{A}}$ communicating with BSs instead of APs increases. Hence, during the day time peak, the total power consumption of the proposed schemes approaches that of the all-on scheme. However, during the

[4]When $\eta > 50$, a UE within $\mathcal{K}_{\mathcal{A}}$ almost surely associates with the BS instead of the AP by Lemma 2.2.

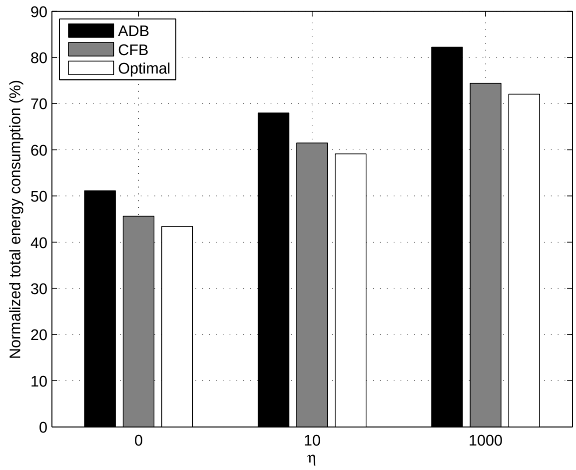


Figure 2.4: Normalized total energy consumption with respect to η.

night time, the performance gap between the all-on scheme and all others exists regardless of the values of η, which happens because most BSs are unutilized or under-utilized and so turned off during off-peak times.

Fig. 2.4 depicts the total amount of energy consumed for a whole day, normalized by that of the all-on scheme. We observe that as η increases the normalized total energy consumption also increases, which happens because in this case the BSs are more highly utilized to increase the network revenue. We observe that the CFB scheme performs close to the optimal scheme for all values of η but the performance gap between the ADB scheme and the optimal scheme increases as η increases. It happens because UEs in $\mathcal{K}_{\mathcal{A}}$ may not communicate with APs when η is large, and thus the density of APs is irrelevant with the BS switching-on/off problem.

2.7 Summary

In this chapter, we have presented a joint algorithm for BS switching and UE association in a heterogeneous network which is composed of cellular networks and WLANs. In devising the algorithm, we illuminated the tradeoff relation between energy consumption and network revenue. We formulated the problem of minimizing the total cost function such that the tradeoff relation is incorporated by a balancing parameter η. In order to make the problem tractable, we decomposed it into two subproblems: UE association problem and BS switching-on/off problem. First, for the UE association algorithm, we derived an optimal policy that each UE associates with a single BS or a single AP based on energy efficiency and network revenue. In particular, when $\eta = 0$ (i.e., when network revenue issue is not considered), each UE communicates with the most energy-efficient AP or BS. Second, for the BS switching-on/off problem, we proposed a couple of greedy algorithms, CFB and ADB: The CFB algorithm is designed to switch off the BS that generates the maximum cost gain while the ADB algorithm is designed to switch off the BS within which coverage the largest number of UEs are associated with APs.

According to the simulations conducted by applying the daily traffic profile, the proposed algorithms can reduce energy consumption by up to about 50% when compared with the all-on scheme. In particular, the CFB scheme turned out simple and efficient with its performance approaching the optimal solution by about 8%. We observed, as expected, the tendency that as η increases the data offloaded through APs decreases. As a result, during peak time energy consumption of the proposed schemes increases as η increases. However, during off-peak times, it decreases significantly without regard to η, which happens because most BSs are unutilized or under-utilized. Therefore, we may conclude that the proposed algorithms are effective in reducing energy consumption and keeping balance between energy consumption and network revenue at the same time.

Chapter 3

Energy-per-Bit Minimized Radio Resource Allocation in Heterogeneous Networks

3.1 Introduction

Today, the number of global mobile phone subscribers approaches 6 billion [22], and wireless devices and equipments consume about 9% of the total energy of information technology (i.e., as much as 6.1 TWh/year) [23]. According to recent surveys, around 80% of the total energy required for the operation of cellular networks, including *base stations* (BSs), *user equipments* (UEs), and the core network, is consumed at BS sites [7]. Furthermore, the number of mobile device UEs keeps increasing globally and the demanded per-user capacity keeps increasing as well. Therefore, as a natural consequence, the study for next-generation mobile network design has been focused on green radio communications with energy-efficient radio resource allocation playing a key role.

Another important issue of next-generation wireless network design is the integration of heterogeneous wireless networks [24], including *wireless local area networks* (WLANs) and cellular networks such as 3GPP-LTE and WiMAX, both utilizing *orthogonal frequency division multiple access* (OFDMA) technology. In dealing with heterogeneous wireless networks, there are two different approaches, namely network selection and multi-homing [25]: Network selection chooses the most appropriate access network among all available alternatives [26–28], whereas multi-homing simultaneously accesses to multiple wireless network interfaces [29]. In general, multi-homing is more beneficial in that it enables UEs to exploit network diversity using multiple interfaces.

Multi-homing capability allows each UE to obtain its required *quality of service* (QoS) from all available wireless access networks. This capability has the following advantages [30]: First, available resources from different wireless access networks can be aggregated to support applications with high data rate. Second, it can support mobility since at least one of the used interfaces will remain active during service provision. Third, the multi-homing concept balances the traffic load across different wireless access networks.

A large amount of work has been reported on resource allocation in downlink OFDMA systems [31–33]: Ref. [31] dealt with the joint optimal subcarrier and power allocation problem for weighted sum rate maximization. Ref. [32] proposed a low-complexity resource allocation algorithm that balances the tradeoff between spectral efficiency and energy efficiency. Ref. [33] approached the energy efficiency optimization problem as a fractional programming. For resource allocation in heterogeneous networks consisting of OFDMA based cellular network and WLAN access technologies, there are several radio resource management algorithms adopting the multi-homing approach [25,30,34–37]: Ref. [25] dealt with sum rate maximization problem under the proportional UE rate constraint and Ref. [34] proposed a max-min fairness based resource management strategy. Refs. [30] and [35] dealt with maximization of utility function while maintaining QoS. Refs. [36] and [37] presented a resource allocation algorithm for sum rate maximization in heterogeneous WLAN and femto-cell networks. However, to the best of the authors' knowledge, no work has yet been reported on energy efficiency of heterogeneous wireless networks.

In this chapter, we study an optimal radio resource management under minimum data rate requirements in heterogeneous networks consisting of an OFDMA based cellular network and multiple WLANs. The objectives of the chapter are three-fold — energy-per-bit minimization with multi-homing capability, determining optimal solution based on double-loop iteration method, and developing simple suboptimal algorithms. First, we consider an optimization problem that minimizes the ratio of the required energy and the transmitted bits under the heterogeneous network with multi-homing access capability. Second, allowing for traffic sharing and interworking between cellular network and WLAN, we formulate the resource allocation problem as a nonlinear fractional programming and then determine its optimal solution by using a double-loop iteration method. Third, we develop two simple suboptimal algorithms, namely time-fraction allocation first (TAF) and normalized time-fraction allocation (NTA), by carefully looking into the properties of the optimal solution.

The rest part of this chapter is organized as follows. In Section 3.2, we introduce the system model of the considered heterogeneous networks and formulate an energy-per-bit minimized radio resource allocation problem as a fractional programming. In Section 3.3, we derive a parametric programming out of the fractional programming and solve the problem by using a double-loop iteration method. Then, in Section 3.4, we discuss how to determine an optimal allocation of subcarriers, power, and time fraction at each iteration and, in Section 3.5, we develop two suboptimal algorithms. Finally, in Section 3.6, we evaluate the performance of the proposed algorithms.

3.2 System Model and Problem Formulation

We consider the downlink transmission in a single macro-cell network (e.g., 3GPP-LTE or WiMAX) in which M WLAN APs ($M \geq 1$) are overlaid, as shown in Fig. 3.1. The BS of the macro-cell network is located at the center of the cell and APs are scattered around the BS. We assume that the BS and each AP work on non-

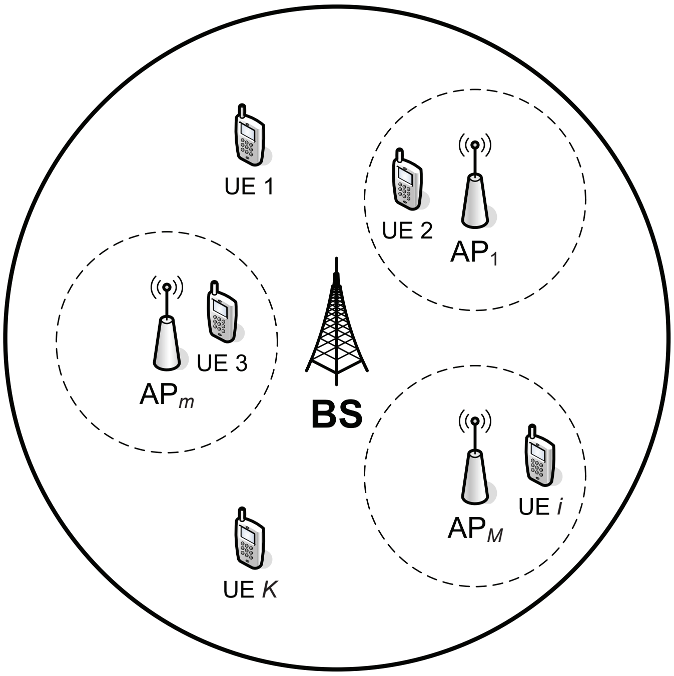


Figure 3.1: A heterogeneous network consisting of a macrocell network and multiple WLANs.

overlapping channels, so that there is no interference among the BS and APs.[1] We assume that all networks in the cell are operated by the same service provider and an intelligent centralized controller exists to manage the radio resource allocation of both the BS and all APs in the network. In order to synchronize two different access technologies operating on different frequency bands, we also assume a coordination-based *time division duplex* (TDD) scheme that predetermines the durations of uplink and downlink commonly for the macro-cell network and WLAN by dividing time into periodic super-frames, each consisting of two phases respectively for uplink and downlink. There are total K UEs which can communicate with either the BS or APs or both simultaneously. If a UE is in the coverage of both the BS and the AP, the UE can connect to both networks, which is called multi-homing access. We consider fully backlogged buffer model and non-real time traffic such as file transfer and online video with minimum rate requirement [39].

We consider an OFDMA based macro-cell network in which traffic bandwidth B is equally divided into N subcarriers, each with a bandwidth of $W(= B/N)$. We assume that each subcarrier may be shared by multiple UEs in time-division manner, and denote by $\alpha_{nk} \geq 0$ the time-sharing factor of UE k on subcarrier n. We also denote by p_{nk} and g_{nk} the average transmit power and the channel gain, respectively, of UE k on subcarrier n, assuming that g_{nk} is accurately known at the transmitter[2] and contains path loss and Rayleigh fading. Then, the maximum

[1]Since 13 channels and 24 channels are available in 2.4 GHz and 5 GHz industrial, scientific, and medical (ISM) bands, respectively [38], it is possible to assign non-overlapping channels to neighboring APs.

[2]In a TDD system, the BS obtains an estimate of the channel state information by using the

achievable data rate of UE k on subcarrier n is

$$r_{nk} = \begin{cases} \alpha_{nk} W \log_2 \left(1 + \dfrac{p_{nk} g_{nk}}{\alpha_{nk} W \sigma^2}\right), & \text{if } \alpha_{nk} > 0 \\ 0 & , \text{if } \alpha_{nk} = 0, \end{cases} \tag{3.1}$$

where σ^2 is the noise spectral density. Note that p_{nk}/α_{nk} in the equation implies the actual transmit power of UE k on subcarrier n. We define by $\mathbb{A}$ and $\mathbb{P}$ the set of the feasible subcarrier assignment matrices and the set of the power allocation matrices, respectively, i.e.,

$$\mathbb{A} = \left\{ [\alpha_{nk}]_{N \times K} \middle| \sum_{k=1}^{K} \alpha_{nk} \leq 1, \forall n; 0 \leq \alpha_{nk} \leq 1, \forall n, \forall k \right\}, \tag{3.2}$$

$$\mathbb{P} = \left\{ [p_{nk}]_{N \times K} \middle| p_{nk} \geq 0, \forall n, \forall k \right\}. \tag{3.3}$$

The total throughput of UE k served by the BS is

$$R_k^{\text{BS}} = \sum_{n=1}^{N} r_{nk}. \tag{3.4}$$

Note that r_{nk} is a concave function with respect to (p_{nk},α_{nk}) [40].

When the BS is in the transmit mode, total power consumption of the BS may be divided into two parts, i.e., static and dynamic power consumption [19]. Static power consumption represents the power consumed by the baseband signal processing and additional circuit blocks such as analog-to-digital conversion, modulation, channel coding, and signal detection [41,42], whereas dynamic power consumption represents the power consumed by power amplifier which changes dynamically in proportion to the transmit power. Then, total power consumption at the BS may be

uplink pilot signals transmitted by UEs and using uplink-downlink reciprocity.

expressed by

$$P_{\text{tot}}^{\text{BS}} = P_{\text{static}}^{\text{BS}} + \xi^{\text{BS}} \sum_{k=1}^{K} \sum_{n=1}^{N} p_{nk}, \tag{3.5}$$

where $P_{\text{static}}^{\text{BS}}$ indicates the static power consumption and $1/\xi^{\text{BS}}$ the efficiency of power amplifier at the BS, which is defined by the ratio of transmit power to input DC power.

In the case of IEEE 802.11 WLAN, we consider an improved version of *distributed coordination function* (DCF) with a reservation-based *medium access control* (MAC) protocol, namely, early backoff announcement (EBA) [43]. In EBA, the UEs can completely avoid collision by announcing the future backoff information in the MAC header. Thus, we may consider that WLAN operates in *time division multiple access* (TDMA) manner with each UE occupying the whole bandwidth in its allocated time fraction. We denote by t_{mk} the time fraction of UE k in the mth AP. We assume that an AP consumes a constant power $P_{\text{idle}}^{\text{AP}}$ in idle state and $P_{\text{tx}}^{\text{AP}}$ in transmit state. Note that $P_{\text{tx}}^{\text{AP}}$ denotes the total power consumed in transmit mode including the output power ($P_{\text{out}}^{\text{AP}}$) radiated at transmit antenna. Then, the total power consumption at the mth AP may be expressed by

$$P_m^{\text{AP}} = \sum_{k=1}^{K} t_{mk} P_{\text{tx}}^{\text{AP}} + \left(1 - \sum_{k=1}^{K} t_{mk}\right) P_{\text{idle}}^{\text{AP}}. \tag{3.6}$$

The first term represents the transmit power consumed at the mth AP while it is in active state and the second term indicates the power consumed in inactive state. We denote by $\tilde{r}_{mk}$ the achievable data rate of UE k through the mth AP, which is determined by the instantaneous *signal-to-noise ratio* (SNR) between the UE k and the mth AP as follows.

$$\tilde{r}_{mk} = c(P_{\text{out}}^{\text{AP}} \tilde{g}_{mk}/\tilde{\sigma}^2) \tag{3.7}$$

where $\tilde{g}_{mk}$ indicates the channel gain of UE k in the mth AP, $\tilde{\sigma}^2$ is noise variance of WLAN channel, and $c(\cdot)$ decides the achievable data rate based on SNR threshold [1]. Note that $P_{\text{out}}^{\text{AP}}$ is constant and same for all associated UEs. If we assume that each UE communicates with only one AP that provides the highest data rate among all APs, then the total throughput of UE k served by WLAN

$$R_k^{\text{AP}} = \sum_{m=1}^{M} t_{mk} \tilde{r}_{mk}. \tag{3.8}$$

Therefore, the total power consumption and the total throughput of the overall network are respectively given by

$$f_{\text{pow}}(\boldsymbol{P}, \boldsymbol{\alpha}, \boldsymbol{T}) = P_{\text{tot}}^{\text{BS}} + \sum_{m=1}^{M} P_m^{\text{AP}} \tag{3.9}$$

$$f_{\text{thr}}(\boldsymbol{P}, \boldsymbol{\alpha}, \boldsymbol{T}) = \sum_{k=1}^{K} \left(R_k^{\text{BS}} + R_k^{\text{AP}} \right), \tag{3.10}$$

where $\boldsymbol{P} = [p_{nk}]_{N \times K}$, $\boldsymbol{\alpha} = [\alpha_{nk}]_{N \times K}$, and $\boldsymbol{T} = [t_{mk}]_{M \times K}$.

Based on the above system model, we now formulate an energy-efficient resource allocation problem in the heterogeneous network. For the formulation, we adopt the concept of *energy-per-bit*, or the amount of energy needed to convey a bit [44, 45].[3] Since the energy-per-bit of the entire network can be determined by

[3]Note that the energy-per-bit is the inverse of the energy-efficiency defined as bit-per-Joule, in general [46].

the ratio of the total power consumption to the total network throughput (Joule/bit), it takes the relation $f_{\text{pow}}(\boldsymbol{P}, \boldsymbol{\alpha}, \boldsymbol{T})/f_{\text{thr}}(\boldsymbol{P}, \boldsymbol{\alpha}, \boldsymbol{T})$.

As the energy efficiency of the entire system increases as the energy-per-bit decreases, the optimization problem that maximizes the energy efficiency of the entire system is equivalent to that for minimizing the energy-per-bit under minimum rate constraints. Therefore, in the case of the heterogeneous network consisting of a BS and M APs, we can formulate the optimization problem as follows:

$$\mathcal{P}1 : \underset{\boldsymbol{P}\in\mathbb{P},\boldsymbol{\alpha}\in\mathbb{A},\boldsymbol{T}}{\text{minimize}} \; \frac{f_{\text{pow}}(\boldsymbol{P}, \boldsymbol{\alpha}, \boldsymbol{T})}{f_{\text{thr}}(\boldsymbol{P}, \boldsymbol{\alpha}, \boldsymbol{T})} \tag{3.11a}$$

$$\text{subject to} \; \sum_{n=1}^{N} r_{nk} + \sum_{m=1}^{M} t_{mk}\tilde{r}_{mk} \geq R_k^{\text{min}}, \; \forall k \tag{3.11b}$$

$$\sum_{k=1}^{K} t_{mk} \leq 1, \; \forall m \tag{3.11c}$$

$$t_{mk} \geq 0, \; \forall m, \; \forall k \tag{3.11d}$$

where R_k^{min} denotes the minimum rate requirement of UE k. Note that the optimization problem $\mathcal{P}1$ is a fractional programming [33], which is non-convex.

3.3 Parametric Approach to Fractional Programming

The objective function of a fractional programming, as we observe in $\mathcal{P}1$, takes the form of a ratio of two functions which are non-linear in general [47]. Since it is a very challenging task to solve a fractional programming directly, we derive a parametric convex optimization problem out of the fractional programming by introducing a parameter. Then, the parametric problem can help us to solve the fractional programming by using a double-loop iteration method.

3.3.1 Parametric Approach

We adopt a non-negative parameter θ to formulate a parametric optimization problem $\mathcal{P}2$ which is closely related with $\mathcal{P}1$.

$$\mathcal{P}2: \underset{\boldsymbol{P}\in\mathbb{P},\boldsymbol{\alpha}\in\mathbb{A},\boldsymbol{T}}{\text{minimize}} \quad f_{\text{pow}}(\boldsymbol{P},\boldsymbol{\alpha},\boldsymbol{T}) - \theta f_{\text{thr}}(\boldsymbol{P},\boldsymbol{\alpha},\boldsymbol{T}) \tag{3.12a}$$

$$\text{subject to} \quad \sum_{n=1}^{N} r_{nk} + \sum_{m=1}^{M} t_{mk}\tilde{r}_{mk} \geq R_k^{\text{min}},\ \forall k \tag{3.12b}$$

$$\sum_{k=1}^{K} t_{mk} \leq 1,\ \forall m \tag{3.12c}$$

$$t_{mk} \geq 0,\ \forall m,\ \forall k. \tag{3.12d}$$

Then, $\mathcal{P}2$ is a convex optimization problem for a given θ since the objective function is formulated as the difference between a convex function and a concave function and the constraint functions (3.12b)–(3.12d) are convex due to the concavity of r_{nk}.

The set I of feasible resource allocation matrices for problems $\mathcal{P}1$ and $\mathcal{P}2$ is convex by the property of the convex optimization problem [18]. We define the minimum value of the objective function of $\mathcal{P}2$ as follows

$$z(\theta) = \min_{(\boldsymbol{P},\boldsymbol{\alpha},\boldsymbol{T})\in I} \ f_{\text{pow}}(\boldsymbol{P},\boldsymbol{\alpha},\boldsymbol{T}) - \theta f_{\text{thr}}(\boldsymbol{P},\boldsymbol{\alpha},\boldsymbol{T}). \tag{3.13}$$

Then, the two problems $\mathcal{P}1$ and $\mathcal{P}2$ are related as described by the following theorem [47].

Algorithm 3 Dinkelbach's method

1: Initialize θ;

2: **do**

3: Determine $z(\theta)$ and $(\boldsymbol{\alpha}^*, \boldsymbol{P}^*, \boldsymbol{T}^*)$ (by applying Algorithm 4);

4: $\theta \leftarrow \dfrac{f_{\text{pow}}(\boldsymbol{\alpha}^*, \boldsymbol{P}^*, \boldsymbol{T}^*)}{f_{\text{thr}}(\boldsymbol{\alpha}^*, \boldsymbol{P}^*, \boldsymbol{T}^*)}$;

5: **while** $(|z(\theta)| > \epsilon)$

6: **return** $\theta, (\boldsymbol{\alpha}^*, \boldsymbol{P}^*, \boldsymbol{T}^*)$;

Theorem 3.1 *The optimal solution set* $(\bar{\boldsymbol{P}}^*, \bar{\boldsymbol{\alpha}}^*, \bar{\boldsymbol{T}}^*)$ *of* $\mathcal{P}1$ *is the same as that of* $\mathcal{P}2$ *for* $\theta = \theta^*$*, where* θ^* *is the root of* $z(\theta)$*. In addition,* θ^* *is the optimal energy-per-bit, i.e.,*

$$\theta^* = \frac{f_{\text{pow}}(\bar{\boldsymbol{P}}^*, \bar{\boldsymbol{\alpha}}^*, \bar{\boldsymbol{T}}^*)}{f_{\text{thr}}(\bar{\boldsymbol{P}}^*, \bar{\boldsymbol{\alpha}}^*, \bar{\boldsymbol{T}}^*)}. \tag{3.14}$$

Theorem 1 implies that for the fractional program $\mathcal{P}1$, there exists an equivalent problem which yields the same optimal solution and whose objective function takes the subtractive form $f_{\text{pow}}(\boldsymbol{P}, \boldsymbol{\alpha}, \boldsymbol{T}) - \theta^* f_{\text{thr}}(\boldsymbol{P}, \boldsymbol{\alpha}, \boldsymbol{T})$. In other words, solving $\mathcal{P}1$ is essentially equivalent to determining θ^* with $z(\theta^*) = 0$. Therefore, we may focus on solving the parametric problem $\mathcal{P}2$ to determine θ^*.

3.3.2 Double-Loop Iteration to Determine Optimal θ

In order to determine the optimal θ, we adopt the Dinkelbach's method [47], as described in Algorithm 3, which applies Newton's method.[4] This algorithm is

[4]Various iterative algorithms are presented in [48] for fractional programming, including Newton's method, binary search method, and their modifications. One may use another numerical algo-

proved to converge to the optimal point for any θ that satisfies $z(\theta) \leq 0$ [47]. We set θ to a value that meets the condition $z(\theta) \leq 0$ [Line 1]. In each iteration, we solve $\mathcal{P}2$ for a given θ using Algorithm 4 to be described in the next section [Line 3]. If the resulting $z(\theta)$ is sufficiently small (i.e., $|z(\theta)| \leq \epsilon$), then the determined $\boldsymbol{\alpha}^*, \boldsymbol{P}^*, \boldsymbol{T}^*$ are the optimal variables and θ is the optimal energy-per-bit. Otherwise, we calculate new θ using the determined $\boldsymbol{\alpha}^*, \boldsymbol{P}^*, \boldsymbol{T}^*$ [Line 4], and then start the next iteration. We may call this approach a double-loop iteration method as two loops of iterations are involved — one in Algorithm 3 and the other in Algorithm 4.

3.4 Optimal Resource Allocation Algorithm

Since $\mathcal{P}2$ is a standard convex optimization problem for a given θ, we solve the optimization problem $\mathcal{P}2$ using the Lagrangian dual approach [18]. Note that if $\theta = 0$, $\mathcal{P}2$ is the same as the energy-minimized resource allocation problem under minimum rate constrains in our earlier work [49]. This implies that the problem in [49] is a special case of $\mathcal{P}2$.[5]

rithm to obtain the root of $z(\theta)$.

[5]In this sense, we may call the allocation algorithm proposed in [49] the *energy consumption minimization* (ECM) algorithm.

If we take the Lagrangian of $\mathcal{P}2$ and incorporate (3.4)–(3.10), we get

$$\begin{aligned} L(\boldsymbol{P}, \boldsymbol{\alpha}, \boldsymbol{T}, \boldsymbol{\nu}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = P_{\text{static}}^{\text{BS}} + M P_{\text{idle}}^{\text{AP}} + \sum_{k=1}^{K} \lambda_k R_k^{\min} \\ - \sum_{m=1}^{M} \nu_m + \sum_{k=1}^{K} \sum_{n=1}^{N} \varphi_{nk}(p_{nk}, \alpha_{nk}, \lambda_k) \\ + \sum_{m=1}^{M} \sum_{k=1}^{K} \psi_{mk}(t_{mk}, \nu_m, \lambda_k, \mu_{mk}), \end{aligned} \tag{3.15}$$

where $\nu_m \geq 0$, $\lambda_k \geq 0$, and $\mu_{mk} \geq 0$ are the Lagrangian multipliers and

$$\varphi_{nk}(p_{nk}, \alpha_{nk}, \lambda_k) \triangleq \xi^{\text{BS}} p_{nk} - (\lambda_k + \theta) r_{nk}, \tag{3.16}$$

$$\psi_{mk}(t_{mk}, \nu_m, \lambda_k, \mu_{mk}) \triangleq \Big\{ P_{\text{tx}}^{\text{AP}} - P_{\text{idle}}^{\text{AP}} - (\lambda_k + \theta) \tilde{r}_{mk} + \nu_m - \mu_{mk} \Big\} t_{mk}. \tag{3.17}$$

Then, the dual problem of $\mathcal{P}2$ is

$$\underset{\boldsymbol{\nu} \geq 0, \boldsymbol{\lambda} \geq 0, \boldsymbol{\mu} \geq 0}{\text{maximize}} \; D(\boldsymbol{\nu}, \boldsymbol{\lambda}, \boldsymbol{\mu}), \tag{3.18}$$

for the dual function,

$$D(\boldsymbol{\nu}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \inf_{\boldsymbol{P} \in \mathbb{P}, \boldsymbol{\alpha} \in \mathbb{A}, \boldsymbol{T}} L(\boldsymbol{P}, \boldsymbol{\alpha}, \boldsymbol{T}, \boldsymbol{\nu}, \boldsymbol{\lambda}, \boldsymbol{\mu}). \tag{3.19}$$

Since $\mathcal{P}2$ is a convex optimization problem and satisfies the Slater's condition, the strong duality holds, that is, the duality gap is zero. Therefore, we can get the optimal solution of $\mathcal{P}2$ by solving (3.18). Let ($\boldsymbol{\nu}^*$, $\boldsymbol{\lambda}^*$, $\boldsymbol{\mu}^*$) denote the dual optimal solution. Then, the optimal $\boldsymbol{P}(\boldsymbol{\nu}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$, $\boldsymbol{\alpha}(\boldsymbol{\nu}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$, and $\boldsymbol{T}(\boldsymbol{\nu}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ that minimize $L(\boldsymbol{P}, \boldsymbol{\alpha}, \boldsymbol{T}, \boldsymbol{\nu}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ in (3.19) are the optimal variables of $\mathcal{P}2$, provided that the complementary slackness condition is satisfied [18].

The dual function $D(\boldsymbol{\nu}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ for given $\boldsymbol{\nu}$, $\boldsymbol{\lambda}$, and $\boldsymbol{\mu}$ takes the expression

$$\begin{aligned} D(\boldsymbol{\nu}, \boldsymbol{\lambda}, \boldsymbol{\mu}) &= \inf_{\boldsymbol{P}\in\mathbb{P}, \boldsymbol{\alpha}\in\mathbb{A}, \boldsymbol{T}} L(\boldsymbol{P}, \boldsymbol{\alpha}, \boldsymbol{T}, \boldsymbol{\nu}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \\ &= P_{\text{static}}^{\text{BS}} + MP_{\text{idle}}^{\text{AP}} + \sum_{k=1}^{K} \lambda_k R_k^{\text{min}} - \sum_{m=1}^{M} \nu_m \\ &+ \inf_{\boldsymbol{P}\in\mathbb{P}, \boldsymbol{\alpha}\in\mathbb{A}} \sum_{k=1}^{K} \sum_{n=1}^{N} \varphi_{nk}(p_{nk}, \alpha_{nk}, \lambda_k) \\ &+ \inf_{\boldsymbol{T}} \sum_{m=1}^{M} \sum_{k=1}^{K} \psi_{mk}(t_{mk}, \nu_m, \lambda_k, \mu_{mk}). \end{aligned} \tag{3.20}$$

We decompose the dual function into two parts — one for minimizing the φ_{nk} term by allocating power and subcarriers, and the other for minimizing the ψ_{mk} term by adjusting time fraction. The dual function cannot be exclusively divided into two parts because the minimum rate constraint (3.12b) should be met by both parts. However, it is possible to get the optimal solution even if we deal with the two parts separately, as will be proved in the next two subsections.

3.4.1 Optimal Allocation of Subcarrier and Power

We first derive the optimal subcarrier assignment and the optimal power allocation by solving $\inf_{\boldsymbol{P}\in\mathbb{A}, \boldsymbol{\alpha}\in\mathbb{A}} \sum_{k=1}^{K} \sum_{n=1}^{N} \varphi_{nk}(p_{nk}, \alpha_{nk}, \lambda_k)$ as similarly done in [31]. Since $\varphi_{nk}(p_{nk}, \alpha_{nk}, \lambda_k)$ is convex in p_{nk}, the optimal transmit power p_{nk}^* that minimizes φ_{nk} can be easily derived by differentiating φ_{nk} with respect to p_{nk}, i.e.,

$$\left.\frac{\partial \varphi_{nk}}{\partial p_{nk}}\right|_{p_{nk}=p_{nk}^*} = \xi^{\text{BS}} - (\lambda_k + \theta) \left.\frac{\partial r_{nk}}{\partial p_{nk}}\right|_{p_{nk}=p_{nk}^*} = 0. \tag{3.21}$$

By the well-known water-filling method, we get

$$p_{nk}^{*} = \alpha_{nk} W \left[\frac{\lambda_k + \theta}{\xi^{\mathrm{BS}} \ln 2} - \frac{\sigma^2}{g_{nk}} \right]^{+}, \tag{3.22}$$

where $[x]^{+} = \max(0, x)$. By substituting (3.22) into (3.16), we obtain

$$\varphi_{nk}^{*}(\alpha_{nk}, \lambda_k) = \alpha_{nk} J_{nk}(\lambda_k), \tag{3.23}$$

where

$$J_{nk}(\lambda_k) \triangleq \begin{cases} -(\lambda_k + \theta) W \log_2 \left(\dfrac{(\lambda_k + \theta) g_{nk}}{\sigma^2 \xi^{\mathrm{BS}} \ln 2} \right) & \\ +W \left(\dfrac{\lambda_k + \theta}{\ln 2} - \dfrac{\sigma^2 \xi^{\mathrm{BS}}}{g_{nk}} \right) & , \text{ if } \lambda_k > \dfrac{\sigma^2 \xi^{\mathrm{BS}} \ln 2}{g_{nk}} - \theta \\ 0 & , \text{ otherwise.} \end{cases} \tag{3.24}$$

Proposition 3.2 *For any* $\lambda_k \geq 0$, $J_{nk}(\lambda_k) \leq 0$*; and* $J_{nk}(\lambda_k) < 0$*, if* $\lambda_k > \sigma^2 \xi^{BS} \ln 2 / g_{nk} - \theta$.

Proof : By differentiating (3.24), we get

$$\frac{\partial J_{nk}(\lambda_k)}{\partial \lambda_k} = \begin{cases} -W \log_2 \left(\dfrac{(\lambda_k + \theta) g_{nk}}{\sigma^2 \xi^{\mathrm{BS}} \ln 2} \right), \text{ if } \lambda_k > \dfrac{\sigma^2 \xi^{\mathrm{BS}} \ln 2}{g_{nk}} - \theta \\ 0 \qquad , \text{ otherwise.} \end{cases} \tag{3.25}$$

Since $\partial J_{nk}(\lambda_k)/\partial \lambda_k \leq 0$, $J_{nk}(\lambda_k)$ is a non-increasing function of λ_k. In addition, $J_{nk}(\lambda_k) = 0$, for all $\lambda_k \leq \sigma^2 \xi^{\mathrm{BS}} \ln 2 / g_{nk} - \theta$. Consequently, we get $J_{nk}(\lambda_k) \leq 0$ for any $\lambda_k \geq 0$, and get $J_{nk}(\lambda_k) < 0$ if $\lambda_k > \sigma^2 \xi^{\mathrm{BS}} \ln 2 / g_{nk} - \theta$. ■

Proposition 3.3 *For ergodic fading channels with a continuous* cumulative distribution function *(CDF) and a given $\boldsymbol{\lambda}$, the almost surely unique solution of (3.19) yields the following optimal time-sharing factors and power*

$$\begin{cases} \alpha^*_{nk^*_n} = 1, p^*_{nk^*_n} = W\left[\dfrac{\lambda_{k^*_n} + \theta}{\xi^{BS} \ln 2} - \dfrac{\sigma^2}{g_{nk^*_n}}\right]^+ \\ \alpha^*_{nk} = 0,\ p^*_{nk} = 0,\ \forall k \neq k^*_n \end{cases} \tag{3.26}$$

$$k^*_n = \arg\min_k J_{nk}(\lambda_k). \tag{3.27}$$

Proof : For the UE k^*_n determined by (3.27) and for a given $\boldsymbol{\lambda}$, it holds

$$\begin{aligned} &\sum_{k=1}^{K}\sum_{n=1}^{N} \varphi_{nk}(p_{nk}, \alpha_{nk}, \lambda_k) \geq \sum_{k=1}^{K}\sum_{n=1}^{N} \alpha_{nk} J_{nk}(\lambda_k) \\ \geq &\sum_{n=1}^{N}\left(J_{nk^*_n}(\lambda_{k^*_n})\sum_{k=1}^{K}\alpha_{nk}\right) \geq \sum_{n=1}^{N} J_{nk^*_n}(\lambda_{k^*_n}), \end{aligned} \tag{3.28}$$

where the first inequality holds due to the optimal φ^*_{nk} in (3.23); the second one holds due to the definition of k^*_n; and the third one follows from $J_{nk}(\lambda_k) \leq 0$ in Proposition 3.2 and the condition $\sum_{k=1}^{K} \alpha_{nk} \leq 1$ in (3.2). The equality holds for the optimal allocation $(\boldsymbol{\alpha}^*, \boldsymbol{P}^*)$ specified in (3.26), which is optimal for $\inf\limits_{\boldsymbol{P}\in\mathbb{P}, \boldsymbol{\alpha}\in\mathbb{A}} \sum_{k=1}^{K} \sum_{n=1}^{N} \varphi_{nk}(p_{nk}, \alpha_{nk}, \lambda_k)$. The almost sure uniqueness of $(\boldsymbol{\alpha}^*, \boldsymbol{P}^*)$ can be similarly proved as in [31] by using the fact that the event of two UEs having the same $J_{nk}(\lambda_k)$ has Lebesgue-measure zero when the fading process has a continuous CDF. ■

Proposition 3.3 has established two important facts: First, the optimal value of α_{nk} is either 0 to 1, which means that each subcarrier is dedicated to a single UE, though it is assumed to be shared among multiple UEs. Second, the optimal single

UE to which a subcarrier is entirely allocated is that with the smallest $J_{nk}(\lambda_k)$. In this sense, $J_{nk}(\lambda_k)$ plays a key role in determining the subcarrier assignment.

As to power allocation, we can determine an optimal power allocation jointly with subcarrier allocation, similar to [31], which leads to water-filling and "winner-takes-all" strategies. That is, the allocated power is determined by the water-filling algorithm and a subcarrier is exclusively allocated to a single UE, if the channel fading has a continuous CDF. However, differently from [31], we minimize the energy-per-bit coupled with the power and subcarrier allocation of the macro-cell network as well as the time fraction allocation of the WLAN in the heterogeneous networks.[6]

In the case of heterogeneous network we are dealing with, it happens that the BS limits the amount of allocatable resource by the upper bound of the water level of each UE and then APs allocate the time fraction for the UEs that require additional resource to minimize the energy-per-bit, as will be discussed in the next subsection.

3.4.2 Optimal Allocation of Time Fraction

Next, we determine the optimal time fraction by solving $\inf_{\boldsymbol{T}} \sum_{m=1}^{M} \sum_{k=1}^{K} \psi_{mk}(t_{mk}, \nu_m, \lambda_k, \mu_{mk})$. Note that $t_{mk} = 0$ for all UEs that do not communicate with any APs,

[6] If we compare the difference between Proposition 3.3 and the algorithm in [31] in terms of the double-loop iteration method discussed in Section 3.3, the latter only deals with the power and subcarrier allocation which is a part of the inner loop (pertaining to Algorithm 4). Note that the inner loop deals with the time fraction allocation as well in combination with power and subcarrier allocation.

so that they get the resource only from the BS. Thus, in allocating time fraction, we consider only the UEs that are within the coverage of APs and capable of communicating with one or more APs. Recalling the assumption that each UE communicates with only one AP that provides the highest data rate among all APs, we denote by m_k^* the AP that communicates with the UE k.

Proposition 3.4 *For the AP m_k^* that communicates with the UE k,*

$$\lambda_k = \left(P_{\text{tx}}^{\text{AP}} - P_{\text{idle}}^{\text{AP}} + \nu_{m_k^*} - \mu_{m_k^* k}\right)/\tilde{r}_{m_k^* k} - \theta. \tag{3.29}$$

Proof: For the optimal time fraction of APs, we get by (3.17)

$$\begin{aligned}
&\inf_{\boldsymbol{T}} \sum_{m=1}^{M} \sum_{k=1}^{K} \psi_{mk}(t_{mk}, \nu_m, \lambda_k, \mu_{mk}) \\
&= \inf_{\boldsymbol{T}} \sum_{m=1}^{M} \sum_{k=1}^{K} \left(P_{\text{tx}}^{\text{AP}} - P_{\text{idle}}^{\text{AP}} + \nu_m - (\lambda_k + \theta)\tilde{r}_{mk} - \mu_{mk}\right) t_{mk} \\
&= \begin{cases} 0 & , \text{ if } \left(P_{\text{tx}}^{\text{AP}} - P_{\text{idle}}^{\text{AP}} + \nu_{m_k^*} - (\lambda_k + \theta)\tilde{r}_{m_k^* k} - \mu_{m_k^* k}\right) = 0 \\ -\infty, & \text{otherwise.} \end{cases}
\end{aligned} \tag{3.30}$$

Note that $\tilde{r}_{m_k^* k} > 0$ because the achievable rate $\tilde{r}_{m_k^* k}$ of UE k from AP m_k^* should be positive. Therefore, we get (3.29). ■

Proposition 3.4 claims that each UE has its own upper bound of the water level which decreases as $\tilde{r}_{m_k^* k}$ increases. Since $\mu_{mk} \geq 0$, the Lagrangian multiplier λ_k has the upper bound $\check{\lambda}_k$ (i.e., $\check{\lambda}_k = (P_{\text{tx}}^{\text{AP}} - P_{\text{idle}}^{\text{AP}} + \nu_{m_k^*})/\tilde{r}_{m_k^* k} - \theta$). Hence, $\check{\lambda}_k + \theta$ is inversely proportional to $\tilde{r}_{m_k^* k}$. The transmit power allocation specified in (3.26) shows that the water level increases as λ_k increases for a given θ. Thus, UE k

getting a higher data rate $\tilde{r}_{m_k^* k}$ from the AP is allocated with small power from the BS because $\check{\lambda}_k$ is lower for a given θ. In contrast, if UE k gets a lower data rate from AP, the UE takes a higher priority of power allocation from the BS as its $\check{\lambda}_k$ is higher. Therefore, λ_k plays a key role in balancing wireless resources, i.e., the amount of power of the BS and the time fraction of the AP. Proposition 3.4 established that λ_k determines the water level in the BS power allocation. On the other hand, λ_k also controls the time fraction allocation, as the following proposition describes.

PROPOSITION 3.5 *For $(\boldsymbol{\alpha}^*, \boldsymbol{P}^*)$ obtained from Proposition 3.3 and for $(\boldsymbol{\nu}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ satisfying Proposition 3.4, the solution of (3.19) yields the optimal time fraction*

$$\begin{cases} t_{m_k^* k} = \dfrac{1}{\tilde{r}_{m_k^* k}} \left[R_k^{\min} - \displaystyle\sum_{n=1}^{N} r_{nk}(\alpha_{nk}^*, p_{nk}^*) \right]^+, & \text{if } \lambda_k = \check{\lambda}_k \quad \text{(3.31a)} \\ t_{m_k^* k} = 0 & , \text{if } \lambda_k < \check{\lambda}_k \quad \text{(3.31b)} \\ t_{mk} = 0 & , \forall m \neq m_k^*. \quad \text{(3.31c)} \end{cases}$$

If $\sum_{k=1}^{K} t_{mk} < 1$ for any mth AP, the residual time $t_{mk^*} = 1 - \sum_{k=1}^{K} t_{mk}$ is allocated to

$$k^* = \underset{t_{mk}>0, \lambda_k = \check{\lambda}_k, P_{\text{tx}}^{\text{AP}} - P_{\text{idle}}^{\text{AP}} < \theta \tilde{r}_{mk}}{\arg\max} \tilde{r}_{mk}. \tag{3.31d}$$

Proof : From Proposition 3.4, $\lambda_k = \left(P_{\text{tx}}^{\text{AP}} - P_{\text{idle}}^{\text{AP}} + \nu_{m_k^*} - \mu_{m_k^* k}\right) / \tilde{r}_{m_k^* k} - \theta$. By the complementary slackness condition, the optimal $\mu_{m_k^* k}$ should satisfy either $i)$ $\mu_{m_k^* k} > 0$ and $t_{m_k^* k} = 0$ or $ii)$ $\mu_{m_k^* k} = 0$ and $t_{m_k^* k} \geq 0$. Thus, if $\mu_{m_k^* k} > 0$, then $\lambda_k < \check{\lambda}_k$ and $t_{m_k^* k} = 0$, and thus (3.31b) holds. Otherwise, $\lambda_k = \check{\lambda}_k$, then the UE k is allocated with the amount of time fraction which satisfies in (3.12b) with equality, and thus (3.31a) holds. Since we assume that each UE is connected to only one AP, the time fraction allocated to the other AP's is zero, and thus (3.31c) holds. Regardless of t_{mk}, ψ_{mk} has the infimum for the Lagrangian multipliers satisfying (3.29). However, t_{mk} is determined by the minimum data rate constraint in (3.12b). Once the equality condition of (3.12b) is met through (3.31a)–(3.31c), we additionally allocate the residual time fraction to the UE that will get the largest benefit if the time fraction is allocated. The benefit is determined by the objective function of $\mathcal{P}2$ in (3.12a).

$$\begin{aligned} & f_{\text{pow}}(\boldsymbol{P}^*, \boldsymbol{\alpha}^*, \boldsymbol{T}) - \theta f_{\text{thr}}(\boldsymbol{P}^*, \boldsymbol{\alpha}^*, \boldsymbol{T}) \\ &= \sum_{m=1}^{M} \sum_{k=1}^{K} t_{mk} \left(P_{\text{tx}}^{\text{AP}} - P_{\text{idle}}^{\text{AP}} - \theta \tilde{r}_{mk} \right) + \Omega(\boldsymbol{P}^*, \boldsymbol{\alpha}^*), \end{aligned} \tag{3.32}$$

where $\Omega(\cdot)$ indicates the remaining part of (3.12a) which is irrelevant to time fraction. Therefore, if $P_{\text{tx}}^{\text{AP}} - P_{\text{idle}}^{\text{AP}} < \theta \tilde{r}_{mk}$, the UE k that will take the largest data rate from mth AP takes all the residual time fraction of the mth AP. ■

The time fraction allocation in Proposition 3.5 implies that λ_k is an indicator that determines whether or not the aid of APs is needed by UE k. If λ_k is strictly less than $\check{\lambda}_k$, UE k can meet the minimum data rate requirement without the help of APs. However, if $\lambda_k = \check{\lambda}_k$, it needs support of APs. Hence, each AP allocates time fraction to the UEs that cannot meet the required QoS without the support of APs due to a high data rate from the corresponding APs and a low water level. In addition, if there is any time fraction of an AP left after the allocation of (3.31a)–(3.31c), the residual time fraction is allocated to the UE getting the highest data rate from the AP when $P_{\text{tx}}^{\text{AP}} - P_{\text{idle}}^{\text{AP}} < \theta \tilde{r}_{mk}$. Note that if $\theta = 0$, the residual time fraction is not allocated to any UEs since $\left(P_{\text{tx}}^{\text{AP}} - P_{\text{idle}}^{\text{AP}}\right) > 0$ by (3.31d) [49]. This happens because in the case of considering only the minimization of the total power consumption, any additional allocation of time fraction would result in an increase of the total power consumption.

As shown in Proposition 3.5, t_{mk} is determined by means of the condition in (3.12b) in case $\lambda_k = \check{\lambda}_k$. Therefore, despite the decomposition of the dual function $D(\boldsymbol{\nu}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ into two parts, we can determine the optimal time fraction by using the constraint (3.12b), commonly valid for the two parts of $D(\boldsymbol{\nu}, \boldsymbol{\lambda}, \boldsymbol{\mu})$.

3.4.3 Lagrangian Multipliers Update Algorithm

By optimizing the allocation of power and subcarrier as given in Proposition 3.3 and by optimizing the allocation of time fraction as in Proposition 3.5, for the given ($\boldsymbol{\nu}$, $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$) satisfying Proposition 3.4, we can minimize the energy-per-bit of the whole network. The resulting terms α_{nk}^*, p_{nk}^*, and t_{mk}^* then form an optimal solution to the optimization problem. However, the individual rate constraint or the time fraction constraints in $\mathcal{P}2$ may not be satisfied.

To determine the optimal values of $\boldsymbol{\nu}^*$, $\boldsymbol{\lambda}^*$, and $\boldsymbol{\mu}^*$, we adopt an iterative updating algorithm such that the constraints in $\mathcal{P}2$ are satisfied. For given $\boldsymbol{\nu}^*$ and $\boldsymbol{\lambda}^*$, it is easy to determine $\boldsymbol{\mu}^*$ by Proposition 3.4, which yields

$$\mu_{mk}^* = P_{\text{tx}}^{\text{AP}} - P_{\text{idle}}^{\text{AP}} + \nu_m^* - (\lambda_k^* + \theta)\tilde{r}_{mk}. \tag{3.33a}$$

We update $\boldsymbol{\nu}^*$ and $\boldsymbol{\lambda}^*$ using the subgradient method [50] such that

$$\lambda_k[i+1] = \left[\lambda_k[i] + \beta[i] \left(R_k^{\min} - \sum_{n=1}^{N} r_{nk}[i] - \sum_{m=1}^{M} t_{mk}[i]\tilde{r}_{mk}\right)\right]^+, \tag{3.33b}$$

$$\nu_m[i+1] = \left[\nu_m[i] + \gamma[i] \left(\sum_{k=1}^{K} t_{mk}[i] - 1\right)\right]^+, \tag{3.33c}$$

where $\beta[i]$ and $\gamma[i]$ are sufficiently small positive values that can be tuned by using different procedures [50].[7]

Based on the above discussions, we may formalize an optimal resource allocation algorithm as described in Algorithm 4: First, we initialize the Lagrangian multipliers. Then, for a given set of multipliers, we conduct the process of optimal subcarrier assignment, power allocation, and time fraction allocation. We iterate this process by updating the multipliers using the subgradient method until it reaches convergence. According to [51], this iteration converges to $\boldsymbol{\nu}^*$ and $\boldsymbol{\lambda}^*$ from any initial $\boldsymbol{\nu}[0]$ and $\boldsymbol{\lambda}[0]$ as long as $\beta[i]$ and $\gamma[i]$ are chosen to be sufficiently small.

If the total time fraction allocated to an AP is larger than 1, all the time fraction values of that particular AP are set to 0 [Lines 7–9]. In practice, it is not possible that an AP is connected to its UEs for more than one time fraction. We can avoid such situation by arranging the BS to allocate more resource to the UEs coupled with the AP. If the time fraction of each UE communicating with the AP is set to 0, λ_k of each UE increases due to (3.33b). Hence, the amount of power allocated by BS to the UEs increases as described in Proposition 3.3. The number of iterations required to achieve the $\epsilon-$optimality (i.e., $D^* - D < \epsilon$) is in the order of $\mathcal{O}(1/\epsilon^2)$ [18], and the resulting computational complexity is in the order of $\mathcal{O}\left((M+N)K(1/\epsilon^2)\right)$ [52].

[7] For example, $\beta[i] = c/\sqrt{i}$ where c is the initial step size and i is the iteration number.

Algorithm 4 Optimal resource allocation algorithm

Initialize $(\boldsymbol{\nu}[0], \boldsymbol{\lambda}[0])$, $q[0] \leftarrow 0, i \leftarrow 0$;

do

  $i \leftarrow i + 1$;

  Determine $(\boldsymbol{\alpha}, \boldsymbol{P})$ by applying Proposition 3.3;

  Determine $\boldsymbol{T}$ by applying Proposition 3.5;

  Update $\boldsymbol{\nu}[i]$ from (3.33c);

  for $m = 1$ to M **do**

    if $\sum_{k=1}^{K} t_{mk} > 1$ **then** $t_{mk} \leftarrow 0, \forall k$;

  end for

  Update $\boldsymbol{\lambda}[i]$ by using (3.33b);

  Update $\boldsymbol{\mu}[i]$ by using (3.33a);

  $q[i] \leftarrow D(\boldsymbol{\nu}[i], \boldsymbol{\lambda}[i], \boldsymbol{\mu}[i])$;

while $(|q[i] - q[i-1]| > \epsilon)$

return $(\boldsymbol{\alpha}, \boldsymbol{P}, \boldsymbol{T})$;

3.5 Design of Suboptimal Algorithms

Note that a UE that gets a higher data rate service from APs tends to take a lower priority in getting resource allocation from the BS due to a lower water level. This property motivates to develop two simple and suboptimal algorithms, namely, *time-fraction allocation first* (TAF) algorithm and *normalized time-fraction allocation* (NTA) algorithm. They are designed to allocate the time fraction of UEs whose data rates served by APs are high enough to meet the minimum data rate requirements without the BS, and then assign power and subcarriers to the remaining UEs to be served by the BS. Although a UE is guaranteed with a minimum data rate only from time fraction allocation, the BS may allocate additional power and subcarrier resource to the UE to minimize the energy-per-bit of the entire network and allow multi-homing access to all UEs. Their computational complexities are in the order of $\mathcal{O}\left(MK + NK(1/\epsilon^2)\right)$.

3.5.1 Time-Fraction Allocation First (TAF) Algorithm

As discussed above, the key idea of suboptimal algorithm is to determine the time fraction allocation at APs first and then to allocate power and subcarriers at the BS. We name it *time-fraction allocation first* (TAF) algorithm.

Algorithm 5 describes the procedures of the TAF algorithm. The term S_m represents the set of UEs that can communicate with the mth AP [Line 3] and the UE with the highest data rate in S_m is denoted as k^* [Line 5]. At first, each AP allocates to its UEs the time fraction required for guaranteeing a minimum rate, in the

Algorithm 5 TAF algorithm

$\langle \boldsymbol{T}$ *allocation*$\rangle$

for $m = 1$ to M **do**

  $S_m \leftarrow \{k | \tilde{r}_{mk} > 0\}$;

  while $S_m \neq \phi$ **do**

    $k^* \leftarrow \underset{k \in S_m}{\text{argmax}}\ \tilde{r}_{mk}$, $t_{mk^*} \leftarrow R_{k^*}^{\min} / \tilde{r}_{mk^*}$;

    if $\sum_{k=1}^{K} t_{mk} \leq 1$ **then** $R_{k^*}^{\min} \leftarrow 0$;

    else $t_{mk^*} \leftarrow \left[1 - \sum_{k=1, k \neq k^*}^{K} t_{mk}\right]^+$;

      $R_{k^*}^{\min} \leftarrow R_{k^*}^{\min} - t_{mk^*} \tilde{r}_{mk^*}$;

    end if

    $S_m \leftarrow S_m - \{k^*\}$;

  end while

end for

$\langle (\boldsymbol{\alpha}, \boldsymbol{P})$ *allocation*$\rangle$

Initialize $\boldsymbol{\lambda}[0]$, $q[0] \leftarrow 0, i \leftarrow 0$;

do

  $i \leftarrow i + 1$;

  Determine $(\boldsymbol{\alpha}, \boldsymbol{P})$ based on Proposition 3.3;

  Update $\boldsymbol{\lambda}[i]$ by using (3.33b);

  $q[i] \leftarrow \sum_{k=1}^{K} \sum_{n=1}^{N} \varphi_{nk}(p_{nk}, \alpha_{nk}, \lambda_k[i])$;

while $(|q[i] - q[i-1]| > \epsilon)$

return $(\boldsymbol{\alpha}, \boldsymbol{P}, \boldsymbol{T})$;

descending order of the data rate of the UEs. This time fraction allocation of the mth AP continues as long as the sum of the allocated time fractions does not exceed 1. Once time fraction is allocated to a UE, then its minimum rate requirement gets satisfied. Thus, $R_k^{\min}$ should be set to 0 [Line 6]. However, if the residual time fraction is not large enough to meet the minimum data rate requirement of UE k^*, then all the residual time fraction is allocated to that UE and the BS additionally allocates resource to that UE to meet the remaining data rate requirement [Line 7]. After completing time fraction allocation, the BS performs power and subcarrier allocation to all UEs such that the energy-per-bit gets minimized. The procedure of power and subcarrier allocation of the BS is similar to that in Algorithm 4, except that optimization is not needed for the time fraction matrix $\boldsymbol{T}$ and the Lagrange multipliers $\boldsymbol{\nu}$ and $\boldsymbol{\mu}$.

3.5.2 Normalized Time-Fraction Allocation (NTA) Algorithm

Normalized time-fraction allocation (NTA) algorithm allocates the time fraction required to guarantee a minimum rate to all the UEs communicating with APs and normalizes the time fractions such that they sum up to one. Then, the BS performs power and subcarrier allocation to all UEs including the UEs that have not received sufficient data rate to meet a minimum requirement from APs.

Algorithm 6 describes the NTA algorithm. The mth AP allocates time fraction to satisfy the minimum data rate of all the UEs communicating with itself [Lines 3–6]. It normalizes all the time fraction values to the sum of the allocated time

Algorithm 6 Normalized time-fraction algorithm ($\boldsymbol{T}$ allocation)

1: **for** $m = 1$ to M **do**
2: $\quad S_m = \{k|\tilde{r}_{mk} > 0\}$;
3: $\quad$ **while** $S_m \neq \phi$ **do**
4: $\quad\quad k^* \leftarrow \arg\max\limits_{k \in S_m} \tilde{r}_{mk}$, $t_{mk^*} \leftarrow R_{k^*}^{\text{min}} / \tilde{r}_{mk^*}$;
5: $\quad\quad S_m \leftarrow S_m - \{k^*\}$;
6: $\quad$ **end while**
7: $\quad$ **if** $\sum_{k=1}^{K} t_{mk} > 1$ **then**
8: $\quad\quad t_{mk} \leftarrow t_{mk} / \sum_{i=1}^{K} t_{mi}, \forall k$;
9: $\quad\quad R_k^{\text{min}} \leftarrow R_k^{\text{min}} - t_{mk}\tilde{r}_{mk}, \forall k$;
10: $\quad$ **end if**
11: **end for**

fraction values when the sum is greater than 1 [Lines 7 and 8]. The procedure of optimal power and subcarrier allocation is omitted because it is the same as in the TAF algorithm [Lines 13–21].

3.6 Performance Evaluation

We have performed simulations to evaluate the performance of the proposed algorithms for the heterogeneous network model shown in Fig. 3.2. We use a Monte Carlo simulation method and get the average values from hundreds of trials. We assume that the BS is located at the center of a cell of radius 500 m, all APs ($M = 4$) are symmetrically located at equal distance of 350 m, from the BS, and UEs are

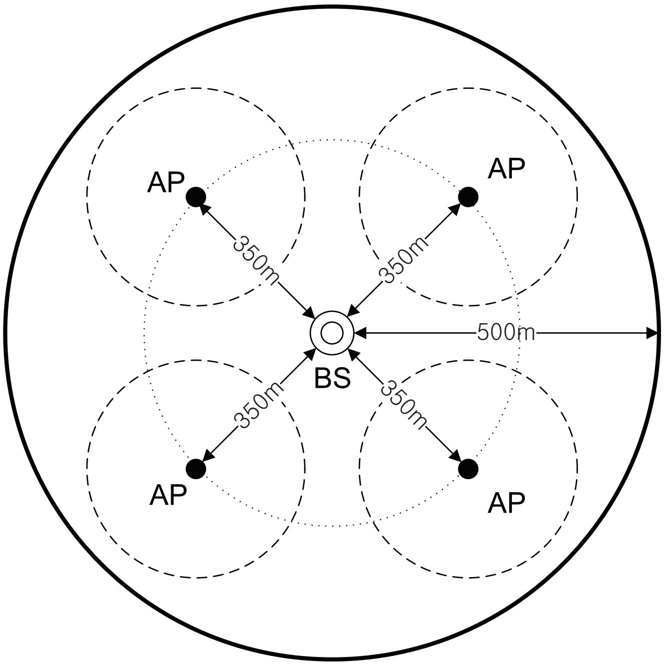


Figure 3.2: Heterogeneous network topology for simulations.

Table 3.1: SNR versus Rate [1].

SNR range (dB)	Rate (Mbps)
>24.56	54
[24.05, 24.56)	48
[18.8, 24.05)	36
[17.04, 18.8)	24
[10.79, 17.04)	18
[9.03, 10.79)	12
[7.78, 9.03)	9
[6.02, 7.78)	6
<6.02	0

uniformly distributed within the cell. We determine the achievable data rates of the UEs served by APs by applying the rate adaptation scheme based on SNR threshold, as shown in Table 3.1 [1]. We assume that the OFDMA system under consideration has 1,024 traffic subcarriers with the subcarrier spacing of 15 kHz. We set the minimum rate requirement for each UE to 3.5 Mbps (i.e., $R_k^{\text{min}} = 3.5$ Mbps, $\forall k$) and take the drain efficiency of 35$\%$ for the power amplifier in the BS (i.e., $1/\xi^{\text{BS}} = 0.35$). We set the parameters related to power consumption as follows: $P_{\text{tx}}^{\text{AP}} = 10.1$ W; $P_{\text{idle}}^{\text{AP}} = 9.2$ W; and $P_{\text{static}}^{\text{BS}} = 77$ W [19, 20].

We compare the performances of six different schemes, which are the optimal, TAF, NTA, the *power and subcarrier allocation only* (PSAO) algorithms, the ECM algorithm proposed in [49],[8] and the *sum rate maximization* (SRM) algorithm. The

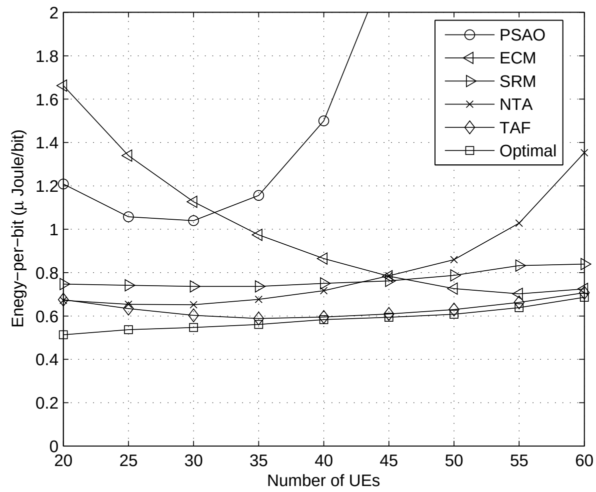
[8]The problem presented in [49] is the same as the optimization problem $\mathcal{P}2$ with $\theta = 0$. Thus,

optimal scheme represents the optimal solution of the energy-per-bit minimization problem $\mathcal{P}1$. The PSAO algorithm assumes that WLAN AP does not exist and thus the BS performs only power and subcarrier allocation without time fraction allocation. In the SRM scheme, we limit total transmit power to 40 W [53] (i.e., $\sum_{k=1}^{K}\sum_{n=1}^{N} p_{nk} \leq 40$). The SRM scheme maximizes total throughput by utilizing full transmit power under minimum rate requirements.

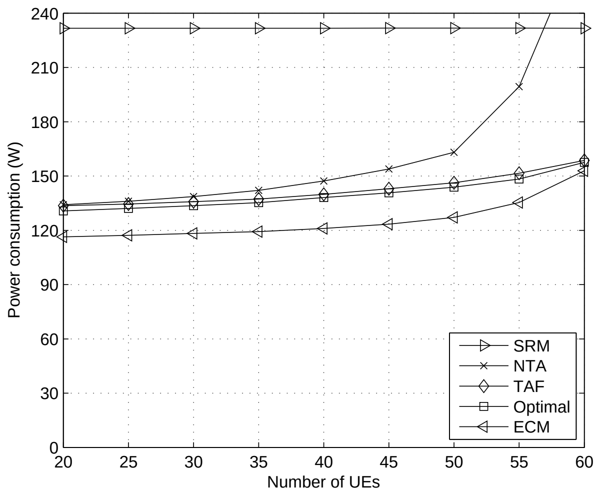
Fig. 3.3(a) depicts the energy-per-bit of resource allocation schemes with respect to the number of UEs. We observe that the energy-per-bit of the PSAO scheme is quasiconvex as proved in [32]. We also observe that the optimal algorithm outperforms the NTA and TAF algorithms, and the TAF scheme is more energy-efficient than the NTA scheme. This implies that assigning time fraction first to the UEs getting a high data rate from APs is more beneficial than assigning time fraction to all the UEs in the vicinity of APs. As the number of UEs increases, energy-per-bit may decrease due to multi-user diversity. However, the energy-per-bits of the SRM and optimal schemes increase although the number of UEs increases. This happens because total QoS requirement becomes more strict with the increasing number of UEs, which dominates the multi-user diversity effect.

Fig. 3.3(b) depicts the total power consumption of various resource allocation schemes. In the SRM scheme, since the maximum transmit (or radiated) power of the BS is limited to 40 W, the total power consumption of the overall network is constant at about 231 W, which can be calculated by (3.9). We observe that the

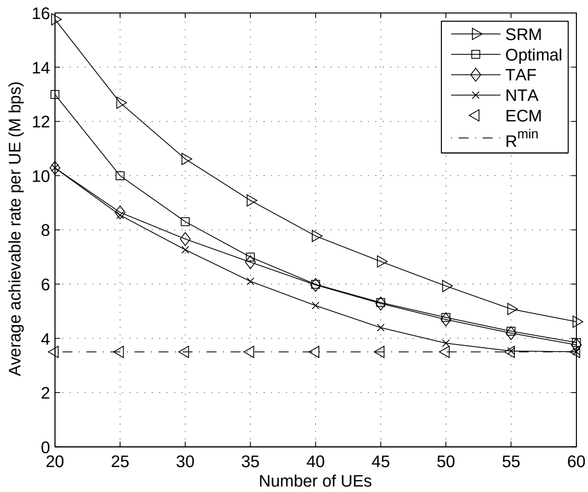
the aim of the ECM algorithm is not at minimizing the energy-per-bit but at minimizing the total power consumption.


PSAO
ECM
SRM
NTA
TAF
Optimal
Enegy−per−bit (μ Joule/bit)
Number of UEs

(a)


SRM
NTA
TAF
Optimal
ECM
Power consumption (W)
Number of UEs

(b)



(c)

Figure 3.3: Performance comparison with respect to the number of UEs: (a) Energy-per-bit, (b) total power consumption, and (c) average achievable rate per UE.

optimal scheme outperforms the NTA and TAF schemes not only in energy-per-bit but also in total power consumption. We also observe that the power consumption of the NTA scheme increases rapidly with the number of UEs. It happens because APs allocate time fraction to the UEs taking lower rates (e.g., 6 Mbps or 9 Mbps) when the number of UEs is sufficiently large, which results in a less energy-efficient usage of time fraction.

Fig. 3.3(c) depicts the average per UE data rate. We may compare the optimal scheme with the ECM scheme. Since the ECM scheme aims at minimizing the total energy consumption, the BS does not perform power and subcarrier allocation to the UEs that are allocated with sufficient time fraction from APs to meet the minimum rate requirement. As a result, most UEs tend to opt for single network allocation than multi-homing access. Hence, in the ECM scheme, the total data rate served by the BS and APs to all UEs is exactly equal to the minimum rate requirement. However, in the case of the optimal scheme considering the minimization of the energy-per-bit, the BS allocates more resource even to the UEs that get sufficient time fraction from APs, as long as the energy-per-bit decreases. Thus, the optimal scheme tends to allocate a data rate higher than the minimum rate requirement, with additional power consumed for reducing the energy-per-bit, as demonstrated in Fig. 3.3(a). In addition, we may compare the optimal scheme with the SRM scheme. We observe that the SRM scheme achieves a data rate higher than the optimal scheme by 25% on the average sense. However, based on Fig. 3.3(a), the optimal scheme improves energy-per-bit by about 32% compared with the SRM scheme which consumes full transmit power without considering energy efficiency. Therefore, the optimal scheme significantly improves energy-per-bit at the cost of small additional power consumption compared with the ECM scheme and relatively little decrease of the sum rate compared with the SRM scheme.

Fig. 3.4 depicts the energy-per-bit of the proposed resource allocation schemes with respect to the minimum rate requirement $R^{\min}$, which is identical for all 40 UEs. The energy-per-bit of the optimal scheme increases with the increasing number of UEs because total minimum rate requirement of all UEs also increases. We also observe that the performance of the TAF scheme approaches that of the NTA scheme as $R^{\min}$ decreases, and approaches the optimal scheme as $R^{\min}$ increases. At $R^{\min} = 1$ Mbps, the system operates in an under-utilized regime such that the

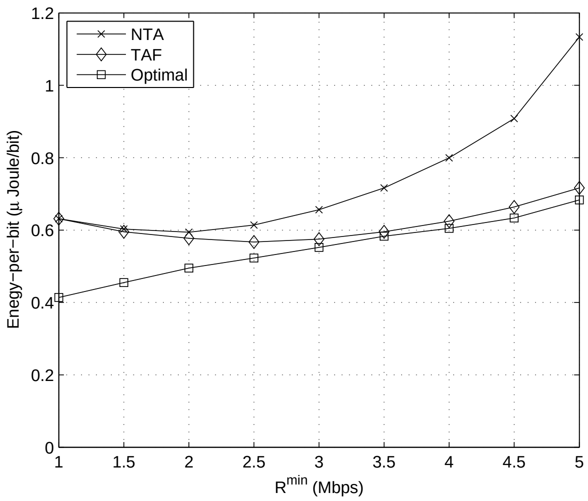
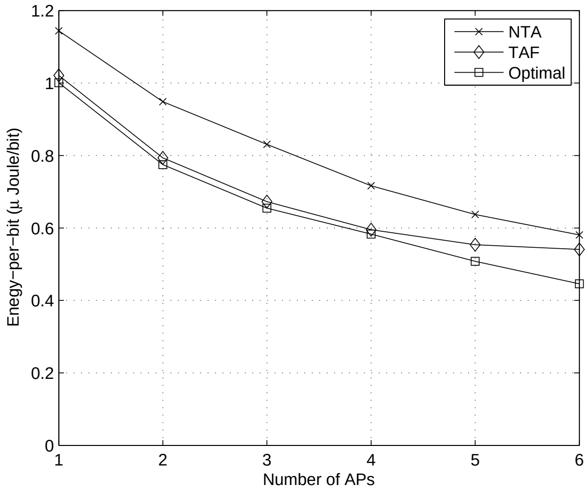


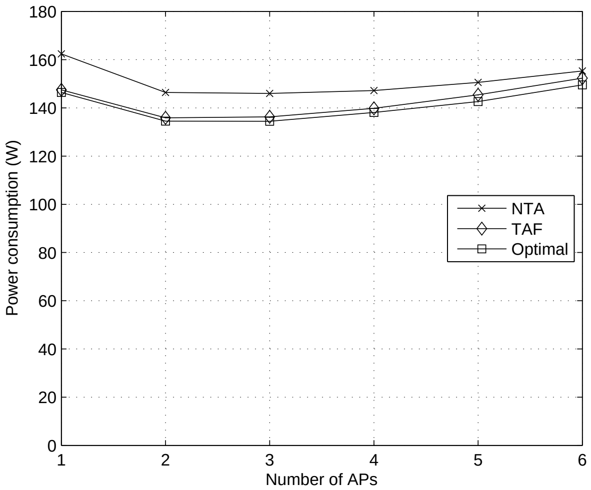
Figure 3.4: Energy-per-bit with respect to $R^{\min}$.

total amount of time fraction required to satisfy the minimum data rate of all UEs associated with APs is less than 1. In this case, the result of time fraction allocation of the TAF algorithm is exactly the same as that of the NTA algorithm, so that the additional power and subcarrier allocations of them are also identical. However, as $R^{\min}$ increases, it is impossible to allocate time fraction to all UEs communicating with APs until their minimum rate requirements are met. Thus, in the case of high $R^{\min}$, it is more energy-efficient to allocate time fraction first to the UEs taking higher data rates from APs.

Fig. 3.5 compares the proposed resource allocation schemes with respect to the number of APs when the number of UEs is fixed at 40. We observe that the optimal scheme outperforms the other two schemes in terms of the energy-per-bit as well as



(a)



(b)

Figure 3.5: Performance comparison with respect to the number of APs: (a) Energy-per-bit and (b) total power consumption.

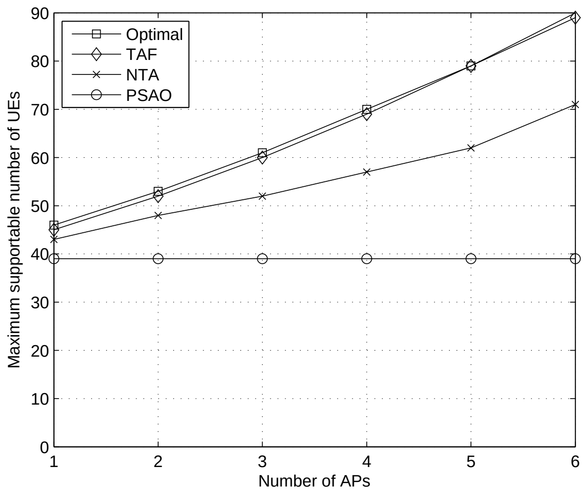


Figure 3.6: Maximum supportable number of UEs with respect to the number of APs.

the total power consumption, irrespective of the number of APs. In Fig. 3.5(a), the energy-per-bit decreases as the number of APs increases. This is the contribution of the WLANs which enable to achieve a high data rate at small power consumption. In Fig. 3.5(b), we observe that the total power consumption is convex with respect to the number of APs. When the number of APs is small, the BS consumes a considerable amount of power to support the minimum rate requirement of 40 UEs and thus, WLAN offloading strongly affects the total power consumption. However, when the number of APs is large, the total power consumption increases as the number of APs increases, which happens because WLAN offloading affects little on total power consumption.

Fig. 3.6 depicts the maximum supportable number of UEs with respect to the number of APs. In practical systems, the maximum transmit power of the BS is regulated due to interference mitigation. In order to reflect this practical limitation, we assume that $\mathcal{P}1$ and $\mathcal{P}2$ are infeasible if $\sum_{k=1}^{K}\sum_{n=1}^{N} p_{nk} > 40$. Thus, the maximum supportable number of UEs means the maximum number of UEs that are guaranteed with the minimum rate requirement by consuming transmit power of 40 W or less at the BS side. The maximum supportable number of UEs of the PSAO scheme is constant because the PSAO scheme is irrelevant to the number of APs. The maximum supportable number of UEs increases as the number of APs increases due to WLAN offloading and network diversity. In addition, the TAF scheme performs nearly close to the optimal scheme because the performance of the TAF scheme approaches that of the optimal scheme as the number of UEs increases, as appears in Fig. 3.3.

Fig. 3.7 depicts the energy-per-bit with respect to the number of UEs in multi-cell network, assuming that the center BS is surrounded by six BSs and the center BS considers the average amount of intercell interference as well as noise. We observe that the energy-per-bit of each scheme in Fig. 3.3(a) is lower than its counterpart in Fig. 3.7 due to intercell interference. When the number of UEs exceeds 40, the SRM and NTA schemes become infeasible, which means that the QoS requirements of all UEs are not guaranteed even with full power consumption. We observe that the performance gap between the optimal scheme and the TAF scheme is wider than that shown in Fig. 3.3(a). It happens because the TAF scheme does not differentiate cell center UEs with cell edge UEs when it performs time fraction allocation.

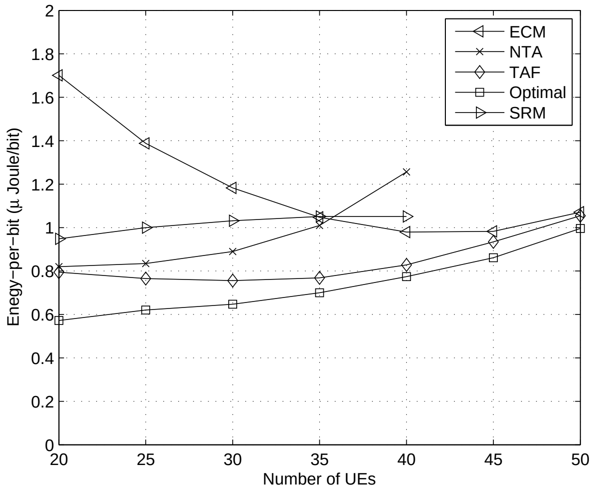


Figure 3.7: Energy-per-bit with respect to the number of UEs in multi-cell environment.

3.7 Summary

In this chapter, we have presented an energy-per-bit minimized radio resource allocation scheme in the heterogeneous networks composed of an OFDMA-based macro-cell network and multiple TDMA-based WLANs. Specifically, we have investigated the energy-per-bit minimization problem in multi-homing environment while guaranteeing minimum data rate requirements. As the resulting optimization problem is a fractional programming, we have derived a parametric programming out of the original problem and solved the original problem by using a double-loop iteration method. Resorting to the Lagrangian dual approach, we have determined the optimal resource allocation policies as follows:

- For subcarrier allocation, we allocate each subcarrier exclusively to the UE that has the smallest $J_{nk}(\lambda_k)$.

• For power allocation, we determine the optimal power by adopting the water-filling algorithm with the water level of each UE decreasing as the data rate served by an AP increases for a given θ. As a result, we found that the BS allocates a smaller power to the UE getting a higher data rate from the corresponding AP.

• For time fraction allocation, we arrange each AP to allocate time fraction to the UEs that cannot get sufficient resource from the BS due to low water level. We allocate the residual time fraction of an AP to the UE getting the highest data rate from the AP if this additional allocation decreases the energy-per-bit.

Based on the above optimal subcarrier, power, and time fraction allocation policies, we have developed two suboptimal algorithms, TAF and NTA, which first allocate the time fraction of UEs getting a high data rate from APs and then allocate power and subcarriers to all UEs. In the TAF algorithm, each AP determines the time fraction allocation in the descending order of data rate but, in the NTA algorithm, each AP allocates time fraction to all UEs with the time fractions normalized to sum up to 1. According to simulations, the TAF scheme has turned out to outperform the NTA scheme. This means that assigning time fraction first to the UEs getting high rates from APs is more energy-efficient than assigning time fraction to all the UEs located near to APs. Also the proposed optimal algorithm has turned out to outperform not only the PSAO scheme which does not use WLANs but also the TAF and NTA schemes in terms of power consumption and energy efficiency. By comparing the optimal scheme with the SRM scheme, we may confirm the trade-off relation between spectral efficiency and energy efficiency. Whereas the ECM scheme allocates a data rate equal to the minimum rate requirement to reduce the total power consumption, the optimal algorithm tends to allocate a data rate higher than the minimum rate requirement with the exceeding power consumption, which contributes to the reduction of the energy-per-bit. Therefore, we may conclude that the proposed energy-per-bit minimized resource allocation scheme is suitable for an energy-efficient design of heterogeneous networks.

Chapter 4

Energy Efficient Scheduling for Carrier Aggregation in OFDMA Based Wireless Networks

4.1 Introduction

Carrier aggregation (CA) is perceived as one of the most promising techniques to provide a higher data rate. According to the LTE-Advanced standard [8], CA supports up to 100 MHz system bandwidth by aggregating up to five *component carriers* (CCs) of 20 MHz and allows a *user equipment* (UE) to use one or multiple CCs simultaneously. Based on how to configure multiple CCs, there are three types of CA [54]: i) Intra-band contiguous aggregation which combines adjacent

CCs within the same band, ii) intra-band non-contiguous aggregation which merges multiple CCs within the same band but in a non-contiguous manner, and iii) inter-band non-contiguous aggregation which connects CCs separated across multiple bands. Among them, we concentrate on the inter-band non-contiguous CA.

Energy efficient communication is a vital aspect of next generation system design, especially for battery constrained UEs because UEs operate based on battery power in the most practical cases. A significant portion of the battery is consumed by network-related part [55]. With CA, it may not be a good idea to always allocate multiple CCs to UEs since the concurrent transmission using CA forces UEs to turn on additional *radio frequency* (RF) elements, which leads to considerable increase of UE power consumption. Hence, it is meaningful to study energy efficient scheduling for CA considering UE power consumption.

There are several researches reported on scheduling algorithm for CA in *orthogonal frequency multiple access* (OFDMA) systems [56–58]. In OFDMA based wireless networks with CA, the BS allocates CCs and *resource blocks* (RBs) to appropriate UEs at each time slot. In [56], the authors proposed a CC allocation and RB allocation scheme considering load balancing and cross-CC proportional fairness. In [57] and [58], the authors presented a suboptimal CC/RB allocation algorithm which maximizes utility function.

Some studies are reported on energy efficient scheduling with the consideration of UE power consumption [59,60]. Ref. [59] dealt with a low-complexity energy efficient scheduling for uplink OFDMA transmission without CA. Ref. [60] proposed a dynamic CC allocation algorithm to improve a new energy efficiency metric without the RB scheduling. However, no study has yet been reported on enhancing the energy efficiency considering UE power consumption in OFDMA based wireless networks supporting CA.

In this chapter, we study an energy efficient scheduling for downlink OFDMA systems with inter-band non-contiguous CA. Specifically, we focus on the power consumption of UEs utilizing CA and define the energy efficiency of each UE as the ratio of the downlink data rate and the UE power consumption. Then, we formulate a CC and RB scheduling problem to achieve the proportional fairness of the energy efficiency of all UEs. To get over the high computational complexity of determining the optimal solution, we develop a low complexity scheduling algorithm, namely *energy efficiency proportional fairness* (EEPF) algorithm.

The rest part of the chapter is as follows. In Section 4.2, we describe the system model. In Section 4.3, we present an energy efficient scheduling algorithm for OFDMA based CA system. Then, in Section 4.4, we evaluate the performance of the proposed algorithm.

4.2 System Model

We consider a downlink OFDMA system in which traffic bandwidth BW is equally divided into N RBs, each with a bandwidth of W. The BS supports CA technology in which up to C CCs can be aggregated. We assume that each CC belongs to different frequency band (inter-band non-contiguous CA) and has the same amount of BW and also the same number of RBs. There are total K UEs that can communi-

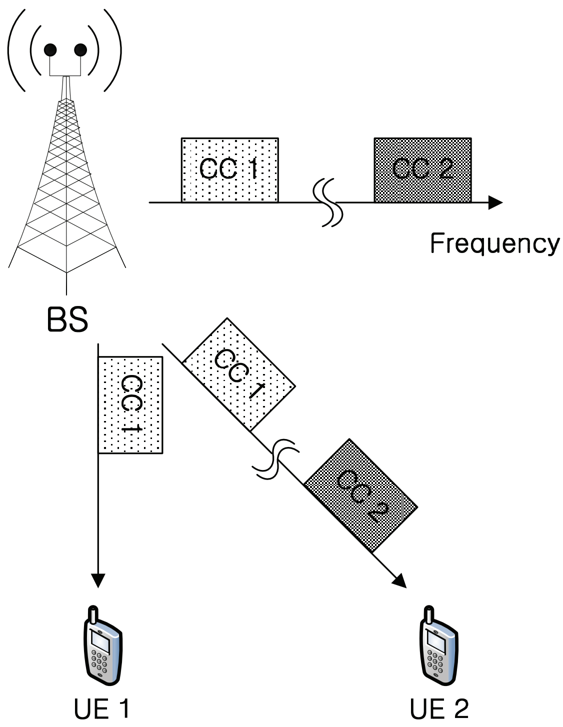


Figure 4.1: Network architecture supporting CA.

cate with either single CC or multiple CCs at the same time, as shown in Fig. 4.1.

We denote by $P_{\max}$ and $g_{k,c,n}$ the maximum transmission power available on each CC and the channel gain of UE k on RB n of CC c, respectively. Then, the maximum achievable data rate of UE k on RB n of CC c at time slot t is

$$r_{k,c,n}[t] = W \log_2 \left(1 + \frac{P_{\max} g_{k,c,n}[t]}{N\sigma^2}\right), \tag{4.1}$$

where σ^2 is the noise variance. Since we assume equal power allocation on each RB,[1] the allocated power on each RB is equal to $P_{\max}/N$. We denote by $\mathcal{C}_k$ and $\mathcal{N}_k^c$

[1]Equal power allocation leads to a throughput degradation compared with the water-filing power

the index set of CCs allocated to UE k and the index set of RBs allocated to UE k on CC c, respectively. Each RB is assigned to only one UE, so that $\mathcal{N}_i^c \bigcap \mathcal{N}_j^c \neq \phi, \forall i \neq j, \forall c$. The total throughput of UE k served by the BS at time slot t is

$$\tilde{r}_k[t] = \sum_{c \in \mathcal{C}_k} \sum_{n \in \mathcal{N}_k^c} r_{k,c,n}[t]. \tag{4.2}$$

We model the power consumption of downlink CA UEs by applying the model designed in [62], [63]. The network-related power of a downlink UE k is consumed at receive RF (RxRF), analog-to-digital converter (ADC), and receive baseband (RxBB) as shown in Fig. 4.2. The RxRF power consumption, P_{RxRF}, is dependent on the amount of received power, the RxBB power consumption, P_{RxBB}, is dependent on the downlink data rate, and the ADC power consumption, P_{ADC}, is dependent only on the bandwidth. Thus, the total UE power consumption at time slot t is expressed by

$$\begin{aligned}\tilde{p}_k[t] = \sum_{c \in \mathcal{C}_k} \Bigg\{ & P_{\text{Rx}} + P_{\text{RxRF}}\left(\frac{P_{\max}}{N}\sum_{n=1}^{N} g_{k,c,n}[t]\right) \\ & + P_{\text{RxBB}}\left(\sum_{n \in \mathcal{N}_k^c} r_{k,c,n}[t]\right) + P_{\text{ADC}}(BW)\Bigg\} + P_{\text{con}},\end{aligned} \tag{4.3}$$

where P_{Rx} is the base power when a UE is in Rx mode and P_{con} is the average power while a UE is in active mode. Note that we consider inter-band CA, and thus there are additional RFs, narrowband ADCs, and BBs. The terms P_{RxRF}, P_{RxBB}, and P_{ADC}

allocation but the degradation negligible if the power is poured on the RBs with good channel gains. Fortunately, in a multi-user system with an adaptive resource allocation scheme, the transmission power is usually allocated to the UEs with good channel gains [61].

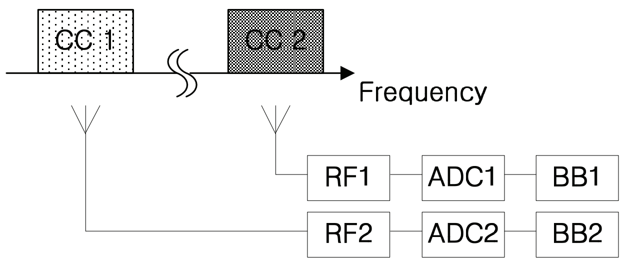


Figure 4.2: UE power consumption model for inter-band CA.

are linearly dependent on variables, i.e., $P_{\text{RxRF}}(x) = \alpha_0 x + \alpha_1$, $P_{\text{RxBB}}(x) = \beta_0 x + \beta_1$, and $P_{\text{ADC}}(x) = \gamma_0 x + \gamma_1$.

If we denote by $R_k[t]$ and $P_k[t]$ the average throughput and average power consumption of UE k at time slot t respectively, they are given by

$$R_k[t] = \left(1 - \frac{1}{T}\right) R_k[t-1] + \frac{1}{T}\tilde{r}_k[t] \tag{4.4}$$

$$P_k[t] = \left(1 - \frac{1}{T}\right) P_k[t-1] + \frac{1}{T}\tilde{p}_k[t], \tag{4.5}$$

where T indicates the effective window size during which the throughput or power consumption is averaged. Then, we define the energy efficiency of UE k as the ratio of the average data rate to the average power consumption of UE k as follows

$$\eta_k[t] = \frac{R_k[t]}{P_k[t]}. \tag{4.6}$$

Note that although we consider a downlink transmission, the denominator of the energy efficiency is the power consumption of UE k, not that of the BS. Since we allocate the transmission power equally to each RB, the BS always consumes a constant amount of power regardless of how the CCs/RBs are allocated and to which

UEs they are allocated. However, at the UE side, the more CCs are allocated, the much more power is consumed, because it has to use additional RF, ADC, and BB. Therefore, we consider energy efficiency in conjunction with UE power consumption.

4.3 Energy Efficiency Proportional Fairness (EEPF) Scheduling

In this section, we discuss how to perform energy efficient scheduling for CA in consideration of proportional fairness. To achieve the proportional fairness in terms of the energy efficiency, it is necessary to allocate CCs/RBs with the objective of maximizing the product of the energy efficiency of all UEs [64]. Thus, we can formulate an *energy efficiency proportional fairness* (EEPF) scheduling for CA system that determines the optimal allocation sets $\mathcal{C}_k$ and $\mathcal{N}_k^c$ as follows

$$\max_{\mathcal{C}_k, \mathcal{N}_k^c} U[t] = \sum_{k=1}^{K} \log(\eta_k[t]), \tag{4.7}$$

where $U[t]$ is the utility function for CA system at time slot t. Note that the proportional fairness criterion enables the scheduler to balance the tradeoff between maximizing the overall energy efficiency and preserving some degree of fairness [65]. Since $U[t-1]$ is given at time slot t, maximizing $U[t]$ is equivalent to maximizing the difference ΔU of the utility function. That is,

$$\max_{\mathcal{C}_k,\mathcal{N}_k^c} \Delta U = \max_{\mathcal{C}_k,\mathcal{N}_k^c} U[t] - U[t-1] \tag{4.8}$$

$$= \max_{\mathcal{C}_k,\mathcal{N}_k^c} \sum_{k=1}^{K} \left\{ \log\left(\frac{R_k[t]}{P_k[t]}\right) - \log\left(\frac{R_k[t-1]}{P_k[t-1]}\right) \right\} \tag{4.9}$$

$$= \max_{\mathcal{C}_k,\mathcal{N}_k^c} \sum_{k=1}^{k} \left\{ \log\left(\frac{R_k[t]}{R_k[t-1]}\right) - \log\left(\frac{P_k[t]}{P_k[t-1]}\right) \right\} \tag{4.10}$$

By using the Taylor expansion with the assumption that $T \gg 1$, we get

$$\max_{\mathcal{C}_k,\mathcal{N}_k^c} \Delta U \tag{4.11}$$

$$\approx \max_{\mathcal{C}_k,\mathcal{N}_k^c} \sum_{k=1}^{K} \left\{ \frac{\tilde{r}_k[t]}{R_k[t-1]} - \frac{\tilde{p}_k[t]}{P_k[t-1]} \right\} \tag{4.12}$$

$$= \max_{\mathcal{C}_k,\mathcal{N}_k^c} \sum_{k=1}^{K} \sum_{c\in\mathcal{C}_k} J(c,k,\mathcal{N}_k^c), \tag{4.13}$$

where

$$J(c,k,\mathcal{N}_k^c) = \sum_{n\in\mathcal{N}_k^c} \left(\frac{r_{k,c,n}[t]}{R_k[t-1]} - \beta_0 \frac{r_{k,c,n}[t-1]}{P_k[t-1]} \right) - \left(\frac{P_{\text{Rx}} + \alpha_0 \frac{P_{\max}}{N} \sum_{n=1}^{N} g_{k,c,n}[t] + \gamma_0 BW + \alpha_1 + \beta_1 + \gamma_1}{P_k[t-1]} \right). \tag{4.14}$$

It is very difficult to determine the optimal solution for (4.7) because the computational complexity of the optimal scheduler is in the order of $\mathcal{O}(K^{CN})$, so that the complexity exponentially increases as the number of CCs or the number of RBs increases. Therefore, we propose a low-complexity EEPF scheduling algorithm that maximizes $\sum_{k=1}^{K} \sum_{c\in\mathcal{C}_k} J(c,k,\mathcal{N}_k^c)$.

We define $D(n, c, k)$ as the RB allocation metric which takes the expression

$$D(n, c, k) = \frac{r_{k,c,n}[t]}{R_k[t-1]} - \beta_0 \frac{r_{k,c,n}[t]}{P_k[t-1]}, \tag{4.15}$$

which is a part of J in (4.14). Since only the metric D is concerned with the index set $\mathcal{N}_k^c$ of allocated RBs, we refer to D as the RB allocation metric. When β_0 is extremely small, the second term of D becomes negligible. Then, the RB allocation metric D is given by

$$D(n, c, k) \approx \frac{r_{k,c,n}[t]}{R_k[t-1]}. \tag{4.16}$$

The scheduler that allocates RBs to the UEs having the largest D in (4.16) is equivalent to the traditional proportional fairness scheduler for CA system [56]. We denote by $J(c, k, \mathcal{N}_k^c)$ the CC allocation metric because J is the whole part related to the index set $\mathcal{C}_k$ of allocated CCs. The metric J may be divided into two parts; the first part which is the sum of D and the residual second part which decreases as the number of allocated CCs increases. If we only consider the first part, it is beneficial to allocate RBs to the UEs having the largest D. However, although a UE is allocated with several RBs which belong to multiple CCs by considering only the first part, this allocation may not be the solution that maximizes the sum of the metric J because the second part decreases. Therefore, the basic principle of the proposed algorithm is that each RB is allocated to the UE having the largest D and the UE having the smallest J is prevented from being allocated with the corresponding CC. We formalize the procedure of this scheduling algorithm as Algorithm 7.

We define B_c as the set of UEs blocked from being allocated with CC c and initialize B_c for all CCs. In each iteration, RB n in CC c is assigned to UE k having

Algorithm 7 EEPF scheduling

Initialize $B_c \leftarrow \phi$ for all c, $m \leftarrow 0$;

while $m < K$ **do**

    for $c = 1$ **to** C **do**

        for $n = 1$ **to** N **do**

            $k^* \leftarrow \underset{k \notin B_c}{\text{argmax}} D(n, c, k)$;

            $\mathcal{C}_{k^*}[m] \leftarrow \mathcal{C}_{k^*}[m] \bigcup \{c\}$;

            $\mathcal{N}_{k^*}^c[m] \leftarrow \mathcal{N}_{k^*}^c[m] \bigcup \{n\}$;

        end for

    end for

    $(c', k') \leftarrow \underset{k \notin \bigcup_{c=1}^{C} B_c, \mathcal{N}_k^c[m] \neq \phi}{\text{argmin}} J(c, k, \mathcal{N}_k^c[m])$;

    $B_{c'} \leftarrow B_{c'} \bigcup \{k'\}$;

    $m \leftarrow m + 1$;

end while

$m^* \leftarrow \underset{m}{\text{argmax}} \sum_{k=1}^{K} \sum_{c \in \mathcal{C}_k[m]} J(c, k, \mathcal{N}_k^c[m])$;

$\tilde{\mathcal{C}}_k \leftarrow \mathcal{C}_k[m^*]$;

$\tilde{\mathcal{N}}_k^c \leftarrow \mathcal{N}_k^c[m^*]$;

the largest value of the RB allocation metric $D(n, c, k)$ [Lines 3–9]. However, if UE k is included in B_c, RB n is assigned to the other UE having the second largest D. After the RB and CC allocation, we calculate the CC allocation metric J using (4.14). Then, we determine the pair (c', k') that minimizes the CC allocation metric J and add UE k' to the CC allocation blocking set $B_{c'}$ [Lines 10–12]. Note that UE k' is prohibited from being allocated only with CC c'. In other words, UE k' is permitted to be allocated with all other CCs except for CC c'. Since we conduct the iteration K times, each UE can be allocated with all CCs except at most one CC. When the iteration ends, we compare the utility of each iteration and determine the optimal allocation sets that maximize $\sum_{k=1}^{K} \sum_{c \in \mathcal{C}_k} J(c, k, \mathcal{N}_k^c)$ [Lines 14–16]. The computational complexity of the EEPF algorithm is in the order of $\mathcal{O}(K^2CN)$.

4.4 Performance Evaluation

We have performed simulations to evaluate the performance of the proposed algorithm for CA system. We assume that the OFDMA based CA system under consideration has 2 CCs at 800 MHz band and 2.1 GHz band, and each CC has 10 MHz bandwidth with 50 RBs and 180 kHz spacing. Each channel model of 800 MHz band CC and 2.1 GHz band CC is based on Winner II [66] and 3GPP LTE [67], respectively. We limit the total transmit power of BS in each CC to 10 W. We set the parameters related to power consumption as follows: $P_{\text{Rx}} = 0.42$ W; $P_{\text{con}} = 1.53$ W; $\alpha_0 = -1.11$ mW/dBm; $\alpha_1 = -60.7$ mW; $\beta_0 = 2.89$ mW/Mbps; $\beta_1 = -26.6$ mW; $\gamma_0 = 11.6$ mW/MHz; $\gamma_1 = -229$ mW [63].

We compare the performances of five different schemes, which are the optimal, cross-CC, *maximal rate* (MR), *round robin* (RR), and no CA scheduling schemes. In the optimal scheme, we determine the optimal allocation sets $\mathcal{C}_k$ and $\mathcal{N}_k^c$ through the exhaustive search. The cross-CC scheme is the conventional PF scheduling for CA system [56], so that with the cross-CC scheduler, each RB is assigned to the UE that maximizes D in (4.16) at each time slot t. Since β_0 is extremely small, the RB allocation metric D in (4.15) is almost the same as D in (4.16). Therefore, the procedure of the cross-CC scheme is the same as Lines 3–9 in Algorithm 7, which means that the cross-CC scheme is equivalent to the proposed EEPF scheme without performing the iteration. The MR scheduling aims at maximizing total throughput, thus each RB is assigned to the UE that maximizes $r_{k,c,n}[t]$ at each time slot t. The RR scheduling allocates RBs to UEs regardless of instantaneous UE channel gains. The no-CA scheme allows only single CC allocation to all UEs. We formalize the procedure of the no-CA algorithm as Algorithm 8.

Fig. 4.3 compares the proposed EEPF scheduling algorithm with the optimal scheme for the different size of time window, T. The utility of the EEPF scheme is normalized by that of the optimal scheme. We observe that the performance of the EEPF scheme approaches that of the optimal scheme as T increases. This happens because the approximation in (4.12) becomes more accurate as T increases. We also observe that when T is larger than 90, the normalized utility of the EEPF scheme converges for all N. Therefore, we set T to 150 in the following simulation results.

Fig. 4.4(a) depicts the average sum of the energy efficiency of all UEs, with respect to the number of UEs. We observe that except the RR scheduling case, the

Algorithm 8 No CA scheduling

Initialize $B_c \leftarrow \phi$ for all c;

for $c = 1$ **to** C **do**

  for $n = 1$ **to** N **do**

    $k^* \leftarrow \underset{k \notin B_c}{\mathrm{argmax}} D(n, c, k)$;

    $\mathcal{C}_{k^*} \leftarrow \mathcal{C}_{k^*} \bigcup \{c\}$;

    $\mathcal{N}_{k^*}^c \leftarrow \mathcal{N}_{k^*}^c \bigcup \{n\}$;

    $B_{c'} \leftarrow B_{c'} \bigcup \{k^*\}, \forall c', c' \neq c$;

  end for

end for

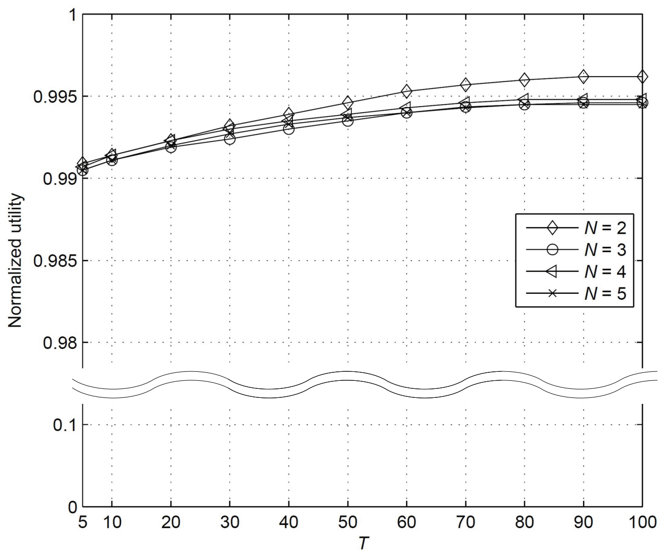
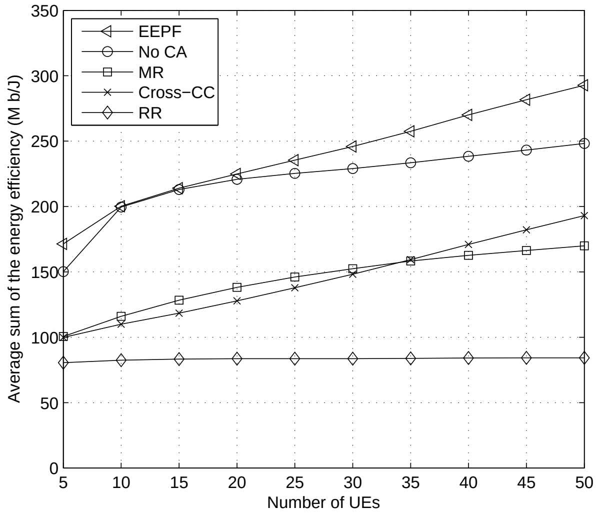
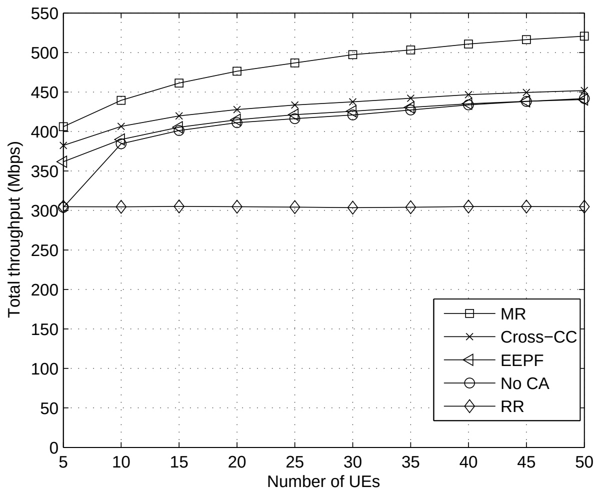


Figure 4.3: Normalized utility of the EEPF with respect to T.



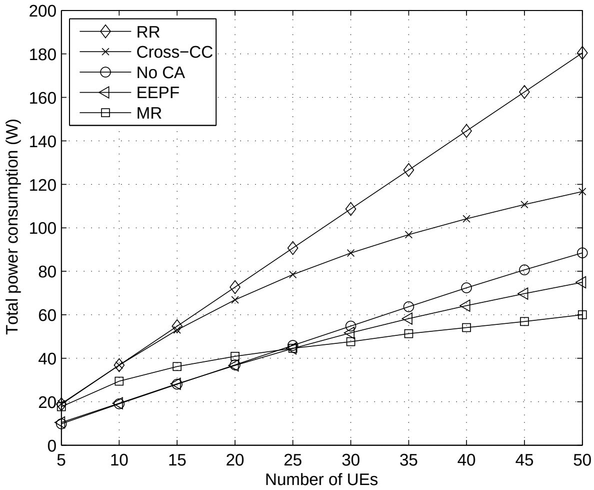
Average sum of the energy efficiency (M b/J)
Number of UEs
EEPF
No CA
MR
Cross−CC
RR
0
50
100
150
200
250
300
350
5
10
15
20
25
30
35
40
45
50

(a)



Total throughput (Mbps)
Number of UEs
MR
Cross−CC
EEPF
No CA
RR
0
50
100
150
200
250
300
350
400
450
500
550
5
10
15
20
25
30
35
40
45
50

(b)



(c)

Figure 4.4: Performance comparison with respect to the number of UEs: (a) Average sum of the energy efficiency, (b) total throughput, and (c) total power consumption.

performances of all other schemes increase as the number of UEs increases. We also observe that the EEPF scheme outperforms other schemes and the performance gap between the EEPF and no CA schemes increases as the number of UEs increases.

Fig. 4.4(b) depicts the total throughput with respect to the number of UEs. We observe that the performances of the MR, cross-CC, EEPF, no CA increase as the number of UEs, which results from the multi-user diversity. We also observe that the MR and cross-CC schemes outperform the EEPF scheme. In detail, the MR and cross-CC schemes achieve a total throughput higher than the EEPF scheme by 14% and 4%, respectively. However, based on Fig. 4.4(a), the EEPF scheme improves the energy-efficiency by about 64% and 66% over the MR and cross-CC schemes, respectively.

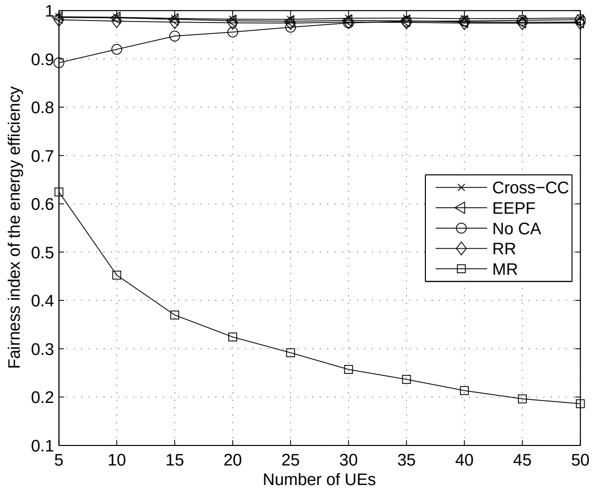
Fig. 4.4(c) depicts the total power consumption with respect to the number of UEs. We observe that when the number of UEs exceeds 25, the MR scheme outperforms other schemes. This occurs because the MR scheme allocates RBs to only a few UEs which are beneficial in the aspect of the total throughput, which leads to significant unfairness as will be discussed below based on the next figure. We also observe that with the exception of the MR scheme, the total power consumption of the EEPF scheme is superior to that of other schemes. This power consumption gain causes the EEPF scheme to enhance the energy efficiency at the cost of relatively little decrease of the total throughput compared with the cross-CC scheme.

Fig. 4.5 compares the fairness performance of the proposed schemes by using the Jain's fairness index [68]:

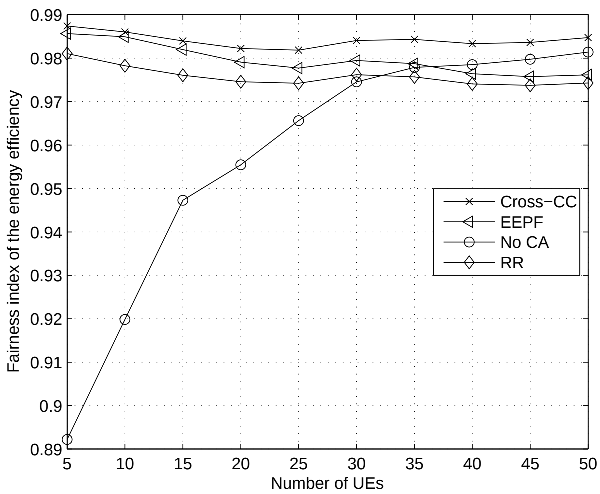
$$\frac{\left|\sum_{k=1}^{K} \eta_k\right|^2}{K \sum_{k=1}^{K} \eta_k^2}. \tag{4.17}$$

In Fig. 4.5(a), we observe that the fairness of the MR scheme is significantly worse than the other schemes as mentioned earlier. In Fig. 4.5(b), we observe that the EEPF scheme is quite fair since its fairness index is always more than 0.97 irrespective of the number of UEs.

Fig. 4.6 depicts the ratio of the single CC allocation with respect to the number of UEs. The ratio is concerned with how many CCs are allocated to UEs at each time slot. If the ratio is one, only single CC is allocated to all UEs at all time slots. However, in case that the ratio is zero, multiple CCs are simultaneously allocated



(a)



(b)

Figure 4.5: Fairness index of the energy efficiency: (a) With the MR scheduling and (b) without the MR scheduling.

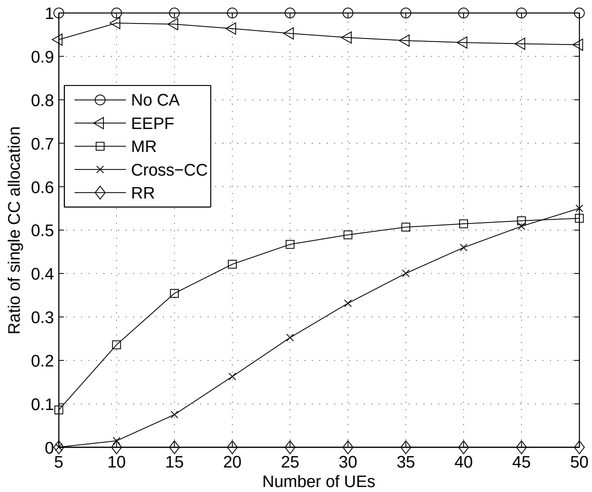


Figure 4.6: Ratio of single CC allocation.

to all UEs at all time slots. Thus, the ratios of the no CA and the RR scheme is equal to one and zero, respectively. We observe that the ratio of the EEPF scheme is always over 0.9, thus the EEPF scheme tends to allocate a single CC to each UE for reducing the UE power consumption, which makes the EEPF scheme more energy-efficient.

4.5 Summary

In this chapter, we have presented an energy efficient and proportional fair CC/RB allocation algorithm for inter-band non-contiguous CA in OFDMA based networks which is the most common modulation scheme for broadband wireless standards such as WiMAX and 3GPP LTE. We define the utility as the log-sum of the energy efficiency of all UEs, and the energy efficiency as the ratio of the time-averaged downlink data rate and the UE power consumption, respectively. Note that the BS power consumption is not considered in the definition of the energy efficiency due to the equal power allocation on each RB. To achieve the proportional fairness in terms of the energy efficiency, we developed a CC/RB allocation problem that maximizes the utility. With an approximation of the utility, we found that the utility increases as we allocate the RBs to the UEs having the largest RB allocation metric D and prevent the UEs having the smallest CC allocation metric J from being allocated with the corresponding CCs. Based on this property, we devised a low-complexity scheduling algorithm, called EEPF scheduling scheme.

According to the simulation results comparing the EEPF scheme with the optimal, no CA, cross-CC, MR, and RR scheduling schemes, we have observed that the proposed EEPF scheme performs close to the optimal scheme and it outperforms other schemes. In terms of the energy efficiency, the EEPF scheme achieves a higher performance than the cross-CC scheme dose by about 66% at the cost of relatively little (4%) decrease of the total throughput. We have also observed the probability that the scheduled UE uses a single CC is always higher than 90% for any number of UEs. Therefore, we may conclude that CA can be used to increase the data rate but we have to carefully exploit it not to deteriorate the energy efficiency.

Chapter 5

Conclusion

5.1 Research Contributions

The dissertation has studied novel methods of efficient BS control for green wireless communications. The BS control strategies may be classified based on various time scales, i.e., from several hours time scale to several milli-seconds time scale, which covers the BS switching-on/off, UE association, radio resource allocation, and UE scheduling. Then, those strategies are considered under two different system models – heterogeneous networks consisting of cellular networks and WLANs, and cellular networks adopting OFDMA with CA. Above system models are two typical ways of meeting the increasing demand for mobile data traffic. However, an abuse of the network diversity of heterogeneous network or CA incurs significant increase of energy consumption. Thus, in order to improve the system's energy efficiency, efficient BS control strategies have been developed by combining the system

Table 5.1: Characteristics of the proposed BS control approaches.

BS control	BS switching-on/off and UE association	Radio resource management	UE scheduling
System model	Heterogeneous networks (cellular + WLAN)		Inter-band CA
Objective	Balancing tradeoff between energy consumption and network revenue	Optimizing energy efficiency of entire network	Achieving PF of energy efficiency considering UE power consumption
Way to find solution	Time scale based problem decomposition	Double-loop iteration method	Low-complexity algorithm
Property	Network selection based on both energy efficiency and revenue	Allocate smaller power to UEs getting higher rate from APs	Tendency to allocate single CC

model and time-scale based operation in three aspects: 1) BS switching-on/off and UE association in heterogeneous networks, 2) optimal radio resource allocation in heterogeneous networks, and 3) energy efficient UE scheduling for CA in OFDMA based cellular networks. Table 5.1 summarizes the characteristics of the proposed BS control approaches.

The first part of the dissertation has presented a joint algorithm for BS switching-on/off and UE association for a heterogeneous network consisting of cellular networks and WLANs. To develop the algorithm, the tradeoff relationship between energy consumption and network revenue was identified. Then, a problem was formulated with the objective of minimizing the total cost function such that the tradeoff relation is incorporated by a balancing parameter η. In order to make the problem tractable, it was decomposed into two subproblems based on time scale, i.e., UE association problem and BS switching-on/off problem, and the corresponding algorithms were developed. Simulation results based on daily traffic profile demonstrated that the proposed algorithms are effective in reducing energy consumption while keeping balance between energy consumption and network revenue at the same time.

The second part of the dissertation has presented an energy-per-bit minimized radio resource allocation scheme for a heterogeneous network composed of an OFDMA-based cellular network and TDMA-based WLANs with multi-homing capability. Specifically, the energy-per-bit minimization problem was investigated while guaranteeing minimum data rate requirements. As the resulting optimization problem is a fractional programming, a parametric programming was derived out of the original problem and the original problem was solved by using a double-loop iteration method. Resorting to the Lagrangian dual approach, the optimal resource allocation policies were determined. Above optimal subcarrier, power, and time fraction allocation policies leaded to developing two suboptimal algorithms, TAF and NTA, which first allocate the time fraction of UEs getting a high data rate from APs and then allocate power and subcarriers to all UEs. Simulation results demonstrated that the proposed optimal algorithm outperforms not only the PSAO scheme which does not use WLANs but also the TAF and NTA schemes in terms of power consumption and energy efficiency.

The third part of the dissertation has presented an energy efficient CC/RB allocation algorithm for OFDMA based networks using CA. Specifically, it was focused on the UE power consumption for the inter-band CA in downlink transmission. The utility is defined as the log-sum of the energy efficiency which is the ratio of the time-averaged downlink data rate and the UE power consumption. To achieve the proportional fairness in terms of the energy efficiency, a CC/RB allocation problem was developed for maximizing the utility. With an approximation of the utility, the tendency was found that the utility increases as the RBs are allocated to the UEs having the largest RB allocation metric and the UEs having the smallest CC allocation metric are prevented from being allocated with the corresponding CCs. Based on this property, a low-complexity scheduling algorithm was developed, namely EEPF scheduling scheme. Simulation results demonstrated that the proposed EEPF scheme performs close to the optimal scheme and is the most energy efficient among the existing schemes for CA.

As such, the dissertation has presented several solutions for green wireless communications in consideration of specific deployment scenarios, operating conditions, and optimization time scale. The first part is adequate for heterogeneous networks on several minutes or hours time scale. The second part is suitable for a singe cellular network and WLANs with multi-homing capability on several milli-seconds time scale. The last part is for CA in an OFDMA based cellular network on several milli-seconds time scale. If they operate together in the context of green communication, this integrated solution will provide a substantial gain on the energy performance. Therefore, the integrated solution may present a useful future direction towards greener communications.

5.2 Future Research Directions

So far, the dissertation has investigated efficient BS control strategies depending on the time scale for green wireless communications, but the proposed schemes can be extended and some of the possible future directions are described below.

First, in the BS switching-on/off and UE association scheme, the revenue of the cellular operators is assumed to be proportional to the amount of data usage for the analytical simplicity. This pricing plan is referred as to usage based pricing. However, in reality, most operators offer "tiered" data plans where different flat-rate prices are set for different data usage caps and the extra volume exceeding the determined data cap is charged proportional to the usage. Therefore, it is worth studying how to perform BS operation and UE association under the practical pricing plans.

Second, in the radio resource allocation scheme, the energy-per-bit minimization is studied in a single-cell heterogeneous network. Thus, the study can be extended to multi-cell heterogeneous network scenario. In addition, it can be dealt with the inter-cell interference issues of cell edge UEs since cell edge UEs suffer from performance degradation due to inter-cell interference in practical cellular systems. However, it is very difficult to optimize the energy efficiency for multi-cell networks because the received interference from the adjacent cells causes in a non-convex and NP-hard problem. It may be possible to resolve the problem by applying cooperative or noncooperative game theory.

Third, in the UE scheduling scheme for CA, only the inter-band CA is considered out of three types of CA. In the inter-band non-contiguous CA case, each UE has to turn on additional RF elements to use multiple CCs at the same time. On the other hand, in the intra-band contiguous CA case, even if a UE is allocated with multiple CCs simultaneously, the UE can successfully receive desired signals by using a single RF element. Hence, it may be more energy-efficient to allocate multiple CCs to each UE in the intra-band contiguous CA. Therefore, it is necessary to extend the UE scheduling scheme to the intra-band CA case, furthermore, the hybrid CA case where the inter-band CA and intra-band CA coexist.

Bibliography

[1] J. Yee and H. Pezeshki-Esfahani, "Understanding wireless LAN performance trade-offs," *Commun. Syst. Design*, pp. 32–35, Nov. 2002.

[2] Cisco systems Inc., "Cisco visual networking index: Global mobile data traffic forecast update, 2012–2017," White paper, Feb. 2013.

[3] EARTH, "Energy aware radio and network technologies project," [Online]. Available: https://www.ict-earth.eu/default.html.

[4] eWIN, "eWIN: Energy-efficient wireless networking," [Online]. Available: http://wireless.kth.se/blog/projects/ewin/.

[5] X. Wang, A. V. Vasilakos, M. Chen, Y. Liu, and T. T. Kwon, "A survey of green mobile networks: Opportunities and challenges," *Mobile Netw. Appl.*, vol. 17, no. 1, pp. 4–20, 2012.

[6] D. Feng, C. Jiang, G. Lim, Jr. L. J. Cimini, G. Feng, and G. Y. Li, "A survey of energy-efficient wireless communications," *IEEE Commun. Surveys & Tutorials*, vol. 15, no. 1, pp. 167–178, 2013.

[7] A. Fehske, G. Fettweis, J. Malmodin, and G. Biczok, "The global footprint of mobile communications: The ecological and economic perspective," *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 55–62, Aug. 2011.

[8] 3GPP, "Evolved universal terrestrial radio access (E-UTRA); stage 2 description," TS 36.300, v9.3.0, Apr. 2010.

[9] E. Oh, B. Krishnamachari, X. Liu, and Z. Niu, "Towards dynamic energy-efficient operation of cellular network infrastructure," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 56–61, June 2011.

[10] A. Fehske, G. Fettweis, J. Malmodin, and G. Biczok, "The global footprint of mobile communications: The ecological and economic perspective," *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 55–62, Aug. 2011.

[11] M. A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Optimal energy savings in cellular access networks," in *Proc. IEEE ICC*, June 2009.

[12] Ofcom, "Sitefinder: Mobile phone base station databases," [Online]. Available: http://www.sitefinder.ofcom.org.uk/.

[13] E. Oh and B. Krishnamachari, "Energy savings through dynamic base station switching in ceilular wireless access networks," in *Proc. IEEE GLOBECOM*, Dec. 2010.

[14] M. A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Switch-off transients in cllular access networks with sleep modes," in *Proc. IEEE ICC*, June 2011.

[15] K. Son, H. Kim, Y. Yi, and B. Krishnamachari, "Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1525–1536, Sept. 2011.

[16] C. Peng, S.-B. Lee, S. Lu, H. Luo, and H. Li, "Traffic-driven power saving in operational 3G cellular networks," in *Proc. ACM MOBICOM*, Sept. 2011.

[17] Z. Niu, Y. Wu, J. Gong, and Z. Yang, "Cell zooming for cost-efficient green cellular networks," *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 74–79, Nov. 2011.

[18] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[19] O. Arnold, F. Richter, G. Fettweis, and O. Blume, "Power consumption modeling of different base station types in heterogeneous cellular networks," in *Proc. ICT-MobileSummit*, June 2010.

[20] Novarum Inc. (Apr. 2009), "Enterprise 801.11n wireless access point performance benchmark," [Online]. Available: http://www.novarum.com/documents/Enterprise802.11nSingleAPBenchmarkTestingv1.3.pdf.

[21] Sandvine. (2010), "Mobile internet phenomena report," [Online]. Available: http://www.sandvine.com/downloads/documents/Phenomena_1H_2012/Sandvine_Global_Internet_Phenomena_Report_1H_2012.pdf.

[22] ITU. (Nov. 2012), "Measuring the information society 2012," [Online]. Available: http://www.itu.int/dms_pub/itu-d/opb/ind/D-IND-ICTOI-2012-SUM-PDF-E.pdf.

[23] J. Mogul, "Improving energy efficiency for networked applications," in *Proc. ANCS Keynote Speech*, June 2007.

[24] 3GPP, "IP flow mobility and seamless wireless local area network (WLAN) offload," TS 36.211, v11.2.0, Mar. 2012.

[25] P. Xue, P. Gong, J. H. Park, D. Park, and D. K. Kim, "Radio resource management with proportional rate constraint in the heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1066–1075, Mar. 2012.

[26] X. Yan, Y. A. Sekercioglu, and S. Narayanan, "A survey of vertical handover decision algorithms in fourth generation heterogeneous wireless networks," *Comput. Netw.*, vol. 54, no. 11, pp. 1848–1863, Aug. 2010.

[27] J. Marquez-Barja, C. T. Calafate, J.-C. Cano, and P. Manzoni, "An overview of vertical handover techniques: Algorithms, protocols and tools," *Comput. Netw.*, vol. 34, no. 8, pp. 985–997, June 2011.

[28] S. Fernandes and A. Karmouch, "Vertical mobility management architectures in wireless networks: A comprehensive survey and future directions," *IEEE Commun. Surveys & Tutorials*, vol. 14, no. 1, pp. 45–63, 2012.

[29] P. Nikander, J. Ylitalo, and J. Wall, "Integrating security, mobility, and multi-homing in a HIP way," in *Proc. NDSS*, Feb. 2003.

[30] M. Ismail and W. Zhuang, "A distributed multi-service resource allocation algorithm in heterogeneous wireless access medium," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 2, pp. 425–432, Feb. 2012.

[31] X. Wang and G. B. Giannakis, "Resource allocation for wireless multiuser OFDM networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4359–4372, July 2011.

[32] C. Xiong, G. Y. Li, S. Zhang, Y. Chen, and S. Xu, "Energy- and spectral-efficiency tradeoff in downlink OFDMA networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3874–3886, Nov. 2011.

[33] C. Isheden, Z. Chong, E. Jorswieck, and G. Fettweis, "Framework for link-level energy efficiency optimization with informed transmitter," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2946–2957, Aug. 2012.

[34] P. Xue, P. Gong, J. H. Park, D. Park, and D. K. Kim, "Max-min fairness based radio resource management in fourth generation heterogeneous networks," in *Proc. ISCIT*, Sept. 2009.

[35] M. Ismail and W. Zhuang, "Decentralized radio resource allocation for single-network and multi-homing services in cooperative heterogeneous wireless access medium," *IEEE Trans. Wireless Commun.*, vol. 11, no. 11, pp. 4085–4095, Nov. 2012.

[36] A. T. Gamage and X. Shen, "Uplink resource allocation for interworking of WLAN and OFDMA-based femtocell systems," in *Proc. IEEE ICC*, June 2013.

[37] A. R. Elsherif, W.-P. Chen, A. Ito, and Z. Ding, “Adaptive small cell access of licensed and unlicensed bands,” in *Proc. IEEE ICC*, June 2013.

[38] B. G. Lee and S. Choi, *Broadband Wireless Access and Local Networks: Mobile WiMAX and WiFi*, Artech House, 2008.

[39] X. Wang, G. B. Giannakis, , and A. G. Marques, “A unified approach to QoS-guaranteed scheduling for channel-adaptive wireless networks,” *Proc. IEEE*, vol. 95, no. 12, pp. 2410–2431, Dec. 2007.

[40] W. Yu and J. M. Cioffi, “FDMA capacity of gaussian multiple-access channels with ISI,” *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 102–111, Jan. 2002.

[41] G. Auer *et al.*, “How much energy is needed to run a wireless network?,” *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 40–49, Oct. 2011.

[42] C. Desset *et al.*, “Flexible power modeling of LTE base stations,” in *Proc. IEEE WCNC*, Apr. 2012.

[43] J. Choi, J. Yoo, S. Choi, and C. Kim, “EBA: An enhancement of the IEEE 802.11 DCF via distributed reservation,” *IEEE Trans. Mobile Comput.*, vol. 4, no. 4, pp. 378–390, July/Aug. 2005.

[44] S. Cui, A. Goldsmith, and A. Bahai, “Energy-constrained modulation optimization,” *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, Sept. 2005.

[45] V. Rodoplu and T. H. Meng, “Energy-efficiency based resource allocation for the scalar broadcast channel,” in *Proc. IEEE WCNC*, Apr. 2012.

[46] V. Rodoplu and T. H. Meng, “Bits-per-joule capacity of energy-limited wireless networks,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 857–865, Mar. 2007.

[47] W. Dinkelbach, “On nonlinear fractional programming,” *Management Science*, vol. 13, no. 7, pp. 492–498, Mar. 1967.

[48] T. Ibaraki, "Parametric approaches to fractional programs," *Mathematical Programming*, vol. 26, no. 3, pp. 345–362, 1983.

[49] S. Kim, B. G. Lee, and D. Park, "Radio resource allocation for energy consumption minimization in multi-homed wireless networks," in *Proc. IEEE ICC*, June 2013.

[50] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, July 2006.

[51] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, 2 ed., 1995.

[52] N. Mokari, M. R. Javan, and K. Navaie, "Cross-layer resource allocation in OFDMA systems for heterogeneous traffic with imperfect CSI," *IEEE Trans. Veh. Technol.*, vol. 59, no. 2, pp. 1011–1017, Feb. 2010.

[53] 3GPP, "Base station (BS) radio transmission and reception," TS 36.104, v10.7.0, July 2012.

[54] M. Iwamura, K. Etemad, M.-H. Fong, R. Nory, and R. Love, "Carrier aggregation framework in 3GPP LTE-advanced," *IEEE Commun. Mag.*, vol. 48, no. 8, pp. 60–67, Aug. 2010.

[55] R. Mittal, A. Kansal, and R. Chandra, "Empowering developers to estimate app energy consumption," in *Proc. ACM MOBICOM*, Aug. 2012.

[56] Y. Wang, K. I. Pedersen, T. B. Sϕrensen, and P. E. Mogensen, "Carrier load balancing and packet scheduling for multi-carrier systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1780–1789, May 2010.

[57] F. Wu, Y. Mao, S. Leng, and X. Huang, "A carrier aggregation based resource allocation scheme for pervasive wireless networks," in *Proc. IEEE DASC*, Dec. 2011.

[58] Z. Huang, Y. Ji, and B. Zhao, "An efficient resource allocation algorithm with carrier aggregation in LTE advanced systems," in *Proc. IEEE WCSP*, Oct. 2012.

[59] G. Miao, N. Himayat, G. Y. Li, and S. Talwar, "Low-complexity energy-efficient scheduling for uplink OFDMA," *IEEE Trans. Commun.*, vol. 60, no. 1, pp. 112–120, Jan. 2013.

[60] F. Liu, K. Zheng, W. Xiang, and H. Zhao, "Design and performance analysis of an energy-efficient uplink carrier aggregation scheme," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 12, Dec. 2014.

[61] Y. J. Zhang and K. B. Letaief, "Multiuser adaptive subcarrier-and-bit allocatin with adaptive cell selection for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1566–1575, Sept. 2004.

[62] A. R. Jensen, M. Lauridsen, P. Mogensen, T. B. Sϕrensen, and P. Jensen, "LTE UE power consumption model: For system level energy and performance optimization," in *Proc. IEEE VTC*, Sept. 2012.

[63] M. Lauridsen, H. Wang, and P. Mogensen, "LTE UE energy saving by applying carrier aggregation in a HetNet scenario," in *Proc. IEEE VTC*, June 2013.

[64] A. Jalali, R. padovani, and R. Pankaj, "Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system," in *Proc. IEEE VTC*, May 2000.

[65] F. Kelly, "Charging and rate control for elastic traffic," *European Trans. Telecommun.*, vol. 8, pp. 33–37, 1997.

[66] IST-4-027756 WINNER II, "Winner II interim channel models," D1.1.1 v1.1, Feb. 2007.

[67] 3GPP, "Further advancements for E-UTRA physical layer aspects," TR 36.814 v9.0.0, Mar. 2010.

[68] R. Jain, D-M. Chiu, and W. R. Hawe, *A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Computer System*, Digital Equipment Corporation, 1984.

국문 초록

본 논문에서는 친환경 무선 통신을 위한 효율적인 기지국 제어 기법들에 관하여 논의한다. 제안하는 기지국 제어 기법들은 시간 척도에 따라 크게 세 종류로 나뉠 수 있다. 시간 단위의 기지국 제어는 기지국을 켜고 끄는 것과 관계되고, 분 단위의 기지국 제어는 기지국과 단말 사이의 접속과 관계되고, 그리고 밀리초 단위의 기지국 제어는 단말 스케쥴링 또는 무선 자원 할당과 관계된다. 본 논문에서는 두 가지 시스템을 고려하는데, 그 하나는 셀룰러 네트워크와 무선 LAN (local area network) 으로 구성된 이종 네트워크이고 다른 하나는 캐리어 집성(carrier aggregation)을 적용한 OFDMA (orthogonal frequency division multiple access; 직교 주파수 분할 다중 접속) 기반의 셀룰러 네트워크이다. 각각의 시스템 모델과 적절한 기지국 제어 기법을 결합하여, 본 논문에서는 친환경 무선 통신을 위한 세 가지의 새로운 기지국 제어 기법들을 제안한다. 구체적으로 1) 이종 네트워크에서 기지국 개폐 및 단말 접속 기법, 2) 이종 네트워크에서 최적의 무선 자원 할당 기법, 3) OFDMA 기반의 셀룰러 네트워크에서 캐리어 집성을 위한 에너지 효율적인 단말 스케쥴링 기법 등이다.

논문의 첫 번째 부분에서는 셀룰러 네트워크와 무선 LAN으로 구성된 이종 네트워크에서 기지국 개폐와 단말 접속을 동시에 수행하는 알고리즘을 제안한다. 우선 전체 네트워크의 에너지 소모와 네크워트 사업자의 수익 사이의 균형을 맞추도록 고안된 전체 비용 함수를 최소화하는 문제를 형성한다. 활성화된 기지국의 집합을 결정하는 시간 척도가 단말 접속 결정을 위한 시간 척도보다 상당히 크다고 가정하면, 형성된 문제는 단말 접속 문제와 기지국 개폐 문제로 분해할 수 있고, 이때 단말 접속 문제는 볼록 문제이므로 최적해를 구할 수 있다. 기지국 개폐 문제는 조합 최적화(combinatorial optimization) 문제이므로 풀기 어렵기 때문에, 전체 비용 함수를 고려하거나 각 기지국 커버리지 내의 AP (access point) 밀도를 고려하는 두 개의 휴리스틱 알고리즘을 제안한다. 시뮬레이션을 통해 제안한 두 개의 휴리스틱 알고리즘이 기지국을 항상 켜는 알고리즘에 비해 에너지 소모를 현저하게 감소시킨다는 것을 확인한다.

논문의 두 번째 부분에서는 동시에 다수의 무선 인터페이스에 접속할 수 있는 멀티호밍이 허용된 이종 네트워크 환경에서 비트 당 에너지를 최소화하는 무선 자원 할당 기법을 제안한다. 비트 당 에너지를 최소화하는 것을 목적으로 하는 최적화 문제는 비선형 분수 계획(fractional programming) 문제에 속한다. 따라서 형성된 분수 계획 문제를 매개 변수 최적화(parameter optimization) 문제로 변환한 후, 이중 루프 반복(double-loop iteration) 방법을 적용하고 각 반복마다 라그랑지안 쌍대(Lagrangian duality) 기법을 이용하여 최적해를 구한다. 또한 최적의 자원 할당 정책의 특성을 활용하여 준최적의 자원 할당 알고리즘을 제안한다. 시뮬레이션 결과를 통해 제안하는 최적의 자원 할당 기법이 동종 네트워크 환경에서 현존하는 기법들에 비해 에너지 효율을 상당히 향상시키고, 준최적 알고리즘에 비해 에너지 소모 측면 뿐 아니라 네트워크 전체의 에너지 효율 측면에서도 상당한 이득이 있음을 확인한다.

논문의 세 번째 부분에서는 OFDMA 기반의 무선 네트워크에서 캐리어 집성을 위한 에너지 효율적인 단말 스케쥴링 알고리즘을 제안한다. 배터리 기반으로 동작하는 단말을 고려하여, 에너지 효율을 시간 평균 하향링크 전송률과 단말 전력 소모의 비로 정의한다. 이때, 에너지 효율의 비례 공정(proportional fair)를 보장하기 위한 CC (component carrier; 컴포넌트 캐리어) 및 RB (resource block; 자원 블록) 할당 문제를 형성한다. 형성된 문제의 최적해를 찾는 것은 매우 어렵기 때문에, 낮은 복잡도의 에너지 효율적인 스케쥴링 알고리즘을 제안한다. 시뮬레이션을 통해 제안한 알고리즘이 최적해에 근접하고, 캐리어 집성을 위한 기존 스케쥴링 알고리즘들보다 우월한 성능을 보임을 확인한다.

주요어: 친환경 통신, 이종 네트워크, 캐리어 집성, 기지국 개폐, 단말 접속, 무선 자원 할당.
학 번 : 2007-20941