Green-energy Aware and Latency Aware User Associations in Heterogeneous Cellular Networks

Citation:
T. Han and N. Ansari, "Green-energy Aware and Latency Aware User Associations in Heterogeneous Cellular Networks," *Proc. IEEE Global Communications Conference* (*GLOBECOM 2013*), Atlanta, GA, Dec. 9-13, 2013, pp. 4946-4951.

URL:
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6855734

Green-energy Aware and Latency Aware User Associations in Heterogeneous Cellular Networks

Tao Han, *Student Member, IEEE*, and Nirwan Ansari, *Fellow, IEEE*
Advanced Networking Laboratory
Department of Electrical and Computer Engineering
New Jersey Institute of Technology, Newark, NJ, 07102, USA
Email: {th36, nirwan.ansari}@njit.edu

***Abstract*—Green communications has received much attention recently. For cellular networks, the base stations (BSs) account for more than 50 percent of the energy consumption of the networks. Therefore, reducing the power consumption of BSs is crucial to greening cellular networks. With the development of green energy technologies, BSs can be powered by green energy in order to reduce the on-grid power consumption, and subsequently reduce the carbon footprints. In this paper, we have proposed a distributed scheme to enable *G*reen-energy *A*ware and *L*atency *A*ware (GALA) user-BS associations in heterogeneous cellular networks whose BSs are powered by both on-grid power and green energy. The GALA algorithm considers both the traffic delivery latency and the green energy generation rate in BSs for determining the user-BS association. We have proved that the GALA algorithm converges to the optimal solution which minimizes the summation of the weighted latency ratios of BSs. The simulation results show that the GALA algorithm enables the trade-off between the on-grid power consumption and the average traffic delivery latency, and saves a significant amount of on-grid power in heterogeneous cellular networks that are powered by both on-grid power and green energy.**

I. Introduction

Owing to the direct impact of greenhouse gases on the earth environment and the climate change, the energy consumption of Information and Communications Technology (ICT) is becoming an environmental and thus social and economic issue. Cellular networks are among the major energy hoggers of communication networks, and their contributions to the global energy consumption increase rapidly. Therefore, greening cellular networks is crucial to reducing the carbon footprints of ICT.

Heterogeneous cellular networks, which optimize the network deployment by taking advantages of the heterogeneity of the next generation cellular networks, are promising in reducing the energy consumption of cellular networks. The network deployment featuring high density deployments of small, low power base stations (BSs) can achieve higher network energy efficiency than that of the sparse deployment of few high power BSs can. Heterogeneous network deployment improves the network efficiency since it employs high density and low power BSs. Etoh *et al.* [1] pointed out that heterogeneous network deployment will bring up to 50 percent reduction of the total power consumption of BSs.

This work was supported in part by NSF under grant no. CNS-1218181.

As green energy technologies advance, green energy such as sustainable biofuels, solar and wind energy can be utilized to power BSs. Telecommunication companies such as Ericsson and Nokia Siemens have designed green energy powered BSs for cellular networks [2]. By adopting green energy powered BSs, mobile service providers may further save on-grid power consumption and thus reduce their CO_2 emissions. However, since the green energy generation is not stable, green energy may not be a reliable energy source for heterogeneous cellular networks. Therefore, future cellular networks are likely to adopt hybrid energy supplies: on-grid power and green energy. Green energy is utilized to reduce the on-grid power consumption and thus reduce the CO_2 emissions while on-grid power is utilized as a backup power source.

In heterogeneous cellular networks with hybrid energy supplies, one of the most important issues is to properly associate mobile users with the serving BSs, referred to as the user-BS association problem. On the one hand, the transmission power of pico BSs (PBSs) are significantly lower than that of macro BSs (MBSs) in heterogeneous cellular networks. Thus, even if mobile users are much closer to PBSs than to the MBS, the users may still receive a stronger downlink pilot signal from the MBS. If mobile users are associated with the BS from which they receive the strongest downlink pilot signal, most of the users will be associated with the MBS. As a result, PBSs may be lightly loaded, and do not contribute much on enhancing the energy efficiency of cellular networks.

On the other hand, in heterogeneous cellular networks, the power consumption of MBSs is significantly larger than that of PBSs. Therefore, associating the users with PBSs may lead to lower power consumption. As a result, the user-BS association schemes that only consider BSs' power consumption [3] may tend to associate as many users with PBSs as possible and lead to heavy traffic congestion in PBSs and degrade the network quality of service (QoS). In this paper, we propose a distributed user-BS association scheme referred to as *G*reen-energy *A*ware and *L*atency *A*ware (GALA) user-BS associations for downlink traffic in heterogeneous cellular networks. GALA not only optimizes the utilization of green energy in order to reduce the on-grid power consumption of the heterogeneous cellular network, but also optimizes the traffic delivery latency of the network to enhance the network QoS.

The rest of the paper is organized as follows. In Section II, we briefly review related works. In Section III, we define the system model. In Section IV, we formulate the user-BS association problem and analyze its properties. Section V presents the proposed user-BS association scheme. Section VI shows the simulation results, and concluding remarks are presented in Section VII.

II. Related Works

In this section, we briefly review related works on designing user-BS association schemes in heterogeneous cellular networks and on greening cellular networks.

A. Designing user-BS association scheme

To balance traffic load among BSs and enhance network QoS, Kim *et al.* [4] proposed a framework for the user-BS association in cellular networks to achieve flow level load balancing under spatially heterogeneous traffic distribution. Jo *et al.* [5] proposed cell biasing algorithms to balance traffic loads among MBSs and PBSs. The cell biasing algorithms perform user-BS association according to the biased measured pilot signal strength, and enables the traffic to be offloaded from MBSs to PBSs. Corroy *et al.* [6] proposed a dynamic user-BS association algorithm to maximize the sum rate of the network and adopted cell biasing to balance the traffic load among BSs. However, most of the existing solutions for user-BS associations in heterogeneous cellular networks do not consider the optimization of the green energy utilization.

B. Greening cellular networks

Green cellular networks have attracted tremendous research efforts from both academia and industry. Correia *et al.* [7] discussed the energy saving potentials at the component level, the link level and the network level for energy aware cellular networks, respectively. Hasan *et al.* [8] provided a comprehensive survey on the solutions that enable energy savings in BSs, and discussed how the heterogeneous network deployments can reduce the energy consumption of cellular networks. Han and Ansari [9] investigated the multicell cooperation solutions for improving the energy efficiency of cellular networks. They discussed traffic-intensity-aware multicell cooperation, which adapts the network layout of cellular networks according to user traffic demands in order to reduce the number of active BSs, and energy-aware multicell cooperation, which offloads traffic from on-grid BSs to off-grid BSs powered by renewable energy, thereby reducing the on-grid power consumption.

As green energy technologies advance, powering BSs with green energy is a promising solution to save on-grid power and reduce the carbon footprints [10]. Hassan *et al.* [11] classified the scenarios and the objectives on utilizing renewable energy in cellular networks. Zhou *et al.* [12] proposed the handover parameter tuning algorithm for target cell selection, and the power control algorithm for coverage optimization to guide mobile users to access the BSs with renewable energy supply. Considering a network with multiple energy supplies, Han and Ansari [13] proposed to optimize the utilization of green energy for cellular networks by the cell size optimization. The proposed algorithm achieves significant on-grid power savings by scheduling the green energy consumption along the time domain for individual BSs, and balancing the green energy consumption among BSs for the cellular network.

III. System Model

In this paper, we consider a heterogeneous cellular network with one MBS and multiple PBSs as shown in Fig. 1. Both the MBS and PBSs are powered by on-grid power and green energy. We consider solar energy as the green energy source. We focus on the downlink user-BS association scenario.

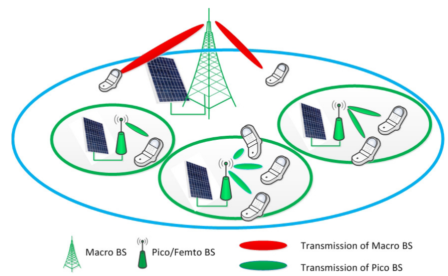

Fig. 1. The green energy powered heterogeneous cellular network.

A. Traffic Model

We consider the MBS and PBSs are deployed to provide data communications to an area. Denote $\mathcal{B}$ as the set of BSs including both the MBS and PBSs. We assume that the traffic arrives according to a Poisson process with the arrival rate per unit area at location x equaling to $\lambda(x)$, and the traffic loads have a general distribution with average traffic load of $\nu(x)$. Assuming a mobile user at location x is associated with BS j, then the user's data rate $r_j(x)$ can be generally expressed as a logarithmic function of the perceived signal to interference plus noise ratio, $SINR_j(x)$, according to the Shannon Hartley theorem [4],

$$r_j(x) = log_2(1 + SINR_j(x)). \tag{1}$$

Here,

$$SINR_j(x) = \frac{P_j g_j(x)}{\sigma^2 + \sum_{k \in \mathcal{B}, k \neq j} P_k g_k(x)}. \tag{2}$$

Here, P_j is the transmission power of BS j, and σ^2 denotes the noise power level. Then, the average traffic load density at location x in BS j is

$$\varrho_j(x) = \frac{\lambda(x)\nu(x)\eta_j(x)}{r_j(x)} \tag{3}$$

Here, $\eta_j(x)$ is an indicator function. If $\eta_j(x) = 1$, then the user at location x is associated with BS j; otherwise, the user is not associated with BS j. Assuming mobile users are uniformly distributed in the area and denoting $\mathcal{A}$ as the coverage area of all the BSs, the traffic load in BS j can be expressed as

$$\rho_j = \int_{x \in \mathcal{A}} \varrho_j(x) dx. \tag{4}$$

This value of ρ_j indicates the fraction of time during which BS j is busy.

B. Energy Model

In the network, both the MBS and PBSs have their own solar panels for generating green energy. Therefore, the BSs are powered by hybrid energy sources: on-grid power and green energy. If green energy generated by the solar panel is not sufficient, the BSs consume on-grid power. Since the MBS usually consumes more energy than PBSs, we assume the MBS is equipped with a larger solar panel that has a higher energy generation capacity than that of a PBS. Owing to the disadvantages of "banking" green energy [14], we do not assume that green energy can be stored. In other words, if green energy is not consumed when it is generated, it is wasted.

The BS's power consumption consists of two parts: the static power consumption and the dynamic power consumption [15]. The static power consumption is the power consumption of a BS without any traffic load. The dynamic power consumption refers to the additional power consumption caused by traffic load in the BS, which can be well approximated by a linear function of the traffic load [15]. Denote p_j^s as the static power consumption of BS j. Then, BS j's power consumption can be expressed as

$$p_j = \beta_j \rho_j + p_j^s. \tag{5}$$

Here, β_j is the linear coefficient which reflects the relationship between the traffic load and the dynamic power consumption in BS j. Denote e_j as the green energy generation rate in BS j. The on-grid power consumption in BS j is

$$p_j^o = \max(p_j - e_j, 0). \tag{6}$$

IV. Problem Formulation

In determining the user-BS associations, the network aims to 1) reduce the on-grid power consumption by optimizing the utilization of green energy, and 2) enhance the network QoS by reducing the traffic delivery latency in the BSs.

Denoting $\mathcal{A}_j$ as BS j's coverage area. We assume that the users associated with BS j are uniformly distributed in its coverage area, and the traffic arrival processes are independent. Since the traffic arrival at a location is a Poisson process, the traffic arrival in BS j, which is the sum of the traffic arrivals from its coverage area, is also a Poisson process. Since the service rate follows a general distribution, the BS realizes an M/G/1 queuing system. Assuming mobile users are served based on the round robin fashion, the traffic delivery in the BS can be modeled as an $M/G/1-PS$ (processor sharing) queue [16].

A mobile user, who is located at x and associated with BS j, is assumed to have the traffic load $\nu(x)$. To fulfill the user's traffic demand, the required service time is

$$\gamma(x) = \frac{\nu(x)}{r_j(x)}. \tag{7}$$

According to [16], the average traffic delivery time for the user in BS j is

$$T_j(x) = \frac{\nu(x)}{r_j(x)(1-\rho_j)}. \tag{8}$$

In BS j, the average waiting time for traffic load $\nu(x)$ is

$$W_j(x) = \frac{\rho_j \nu(x)}{r_j(x)(1-\rho_j)}. \tag{9}$$

Denote $\mu_j(x)$ as the latency ratio that measures how much time a user at location x must be sacrificed in waiting for per unit service time.

$$\mu_j(x) = \frac{W_j(x)}{T_j(x)} = \frac{\rho_j}{1-\rho_j}. \tag{10}$$

According to Eq. (10), $\mu_j(x)$ only depends on the traffic load in BS j. Therefore, all the users associated with BS j have the same latency ratio. Thus, we define

$$\mu_j = \frac{\rho_j}{1-\rho_j} \tag{11}$$

as the latency ratio of BS j. A smaller μ_j indicates that BS j introduces less latency to its associated users. The network aims to minimize the traffic delivery latency. Therefore, one goal of the network is to minimizes the summation of the latency ratio of all the BSs.

On the other hand, since the BSs are powered by both green energy and on-grid power, the network also aims to minimize the utilization of on-grid power by optimizing the utilization of green energy. According to Eq. (6), on-grid power is only consumed when green energy is not sufficient in the BS. When $p_j > e_j$, to avoid on-grid power consumption, BS j has to reduce its traffic load. We define the green traffic capacity as the maximum traffic load that can be supported by green energy. Denote $\hat{\rho}_j$ as the green traffic capacity of BS j. Then,

$$\hat{\rho}_j = \min(\frac{e_j - p_j^s}{\beta_j}, 1-\epsilon). \tag{12}$$

Here, ϵ is an arbitrary small positive constant to guarantee $\hat{\rho}_j < 1$. To reduce traffic loads from ρ_j to $\hat{\rho}_j$, BS j has to shrink its coverage area. As a result, the traffic loads are offloaded to its neighboring BSs and may lead to traffic congestion in the neighboring BSs. The traffic congestion increases the traffic delivery latency of the network. To achieve a trade-off between the traffic delivery latency and the on-grid power consumption, we define the energy-latency coefficient in BS j as θ_j. We further define the virtual traffic load in BS j after the energy-latency trade-off as

$$\bar{\rho}_j = (1-\theta_j)\rho_j + \theta_j \hat{\rho}_j. \tag{13}$$

Here, $0 \le \theta_j < 1$. If the energy-latency coefficient θ_j is set to zero, BS j is latency-sensitive; otherwise, if θ_j approaches one, BS j is energy-sensitive. Given the available green energy and the energy-latency coefficient in BS j, the virtual traffic load is the desired traffic load in BS j.

Considering both the on-grid power consumption and the traffic delivery latency, the goal of the network is

$$\min_{\boldsymbol{\rho}} \quad \sum_{j\in\mathcal{B}} \frac{w_j \rho_j}{1-\rho_j} \tag{14}$$

$$subject\ to: \quad 0 \le \rho_j \le 1-\epsilon. \tag{15}$$

Here, $\boldsymbol{\rho} = (\rho_1, \rho_2, \cdots, \rho_{|\mathcal{B}|})$, and

$$w_j = \frac{\rho_j}{\bar{\rho}_j}. \tag{16}$$

In the objective function, w_j indicates the weight of BS j's latency ratio. If BS j has sufficient green energy, $0 < w_j \leq 1$; otherwise, $w_j > 1$. When the amount of available green energy in BS j is not sufficient, the green traffic capacity, $\hat{\rho}_j$, is smaller than ρ_j. Then, $\bar{\rho}_j < \rho_j$ and $w_j > 1$. With a large weight, the latency ratio of BS j has a high priority while minimizing Eq. (14) as compared with those of the BSs having a small weight. Therefore, as compared with $w_j \leq 1$, BS j with $w_j > 1$ achieves a smaller latency ratio. Since $\frac{d\mu_j}{d\rho_j} > 0$, a smaller latency ratio indicates a smaller traffic load in BS j, which is desirable for saving on-grid power in BS j. Thus, introducing the weights for BSs' latency ratios in the objective function enables the green energy aware and traffic delivery latency aware user-BS associations.

V. GALA: A Distributed User-BS Association Scheme

In this section, we present the GALA algorithm: a distributed user-association scheme which consists of the user side algorithm and the BS side algorithm. The BS side algorithm measures the traffic load in the BS and updates the advertising traffic load. Based on the advertised traffic load, the BS's energy-latency coefficient, and the BS's green traffic capacity, the user side algorithm selects the optimal BS to minimize $\psi(\boldsymbol{\rho}) = \sum_{j\in\mathcal{B}} \frac{w_j \rho_j}{1-\rho_j}$. In order to guarantee convergence of the distributed user-BS association scheme, we assume that the time scale of the traffic arrival and departure process is faster relative to that of BSs in advertising their traffic loads. In other words, BSs broadcast their traffic loads after the system exhibits the stationary performance. We assume that all the BSs are synchronized and advertise their traffic loads simultaneously. We define the time interval between two consecutive traffic load advertisements as a time slot. We assume that the energy-latency coefficient is a constant and the green energy generation rate is consistent during the time period of establishing a stable user-BS association [1].

1) The User Side Algorithm: At the beginning of the kth time slot, BSs broadcast their traffic load $\rho_j(k)$, energy-latency coefficient θ_j, and the green traffic capacity $\hat{\rho}_j$, $\forall j \in \mathcal{B}$ to mobile users.

Let

$$\phi_j(k) = \frac{\bar{\rho}_j(k)^2}{(1-\theta_j)\rho_j(k)^2 + \theta_j \rho_j(k)\hat{\rho}_j(2-\rho_j(k))} \tag{17}$$

The BS selection rule for a user at location x can be expressed as

$$b^k(x) = \arg\max_{j\in\mathcal{B}} r_j(x)(1-\rho_j(k))^2 \phi_j(k) \tag{18}$$

Here, $b^k(x)$ is the index of the BS selected by the user.

2) The BS Side Algorithm: Upon receiving BSs' broadcasting messages including the traffic load, the energy-latency coefficient, and the green traffic capacity, mobile users select BSs according to the user side BS selection algorithm. Then, the coverage area of BS j is updated as

$$\hat{\mathcal{A}}_j(k) = \{x | j = b^k(x),\ \forall x \in \mathcal{A}\} \tag{19}$$

Then, given $\boldsymbol{\rho}(k) = (\rho_1(k), \rho_2(k), \cdots, \rho_{|\mathcal{B}|}(k))$, $\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_{|\mathcal{B}|})$, and $\hat{\boldsymbol{\rho}} = (\hat{\rho}_1, \hat{\rho}_2, \cdots, \hat{\rho}_{|\mathcal{B}|})$, BS j's perceived traffic load in the kth time slot is

$$M_j(\boldsymbol{\rho}(k), \boldsymbol{\theta}, \hat{\boldsymbol{\rho}}) = \min\Big(\int_{x\in\hat{\mathcal{A}}_j(k)} \varrho_j(x)dx, 1-\epsilon\Big). \tag{20}$$

Since the green energy generation rate is consistent before the algorithm converges and the energy-latency coefficient is a constant, $M_j(\boldsymbol{\rho}(k), \boldsymbol{\theta}, \hat{\boldsymbol{\rho}})$ evolves based only on $\boldsymbol{\rho}(k)$. Thus, we use $M_j(\boldsymbol{\rho}(k))$ instead of $M_j(\boldsymbol{\rho}(k), \boldsymbol{\theta}, \hat{\boldsymbol{\rho}})$ for simplicity in the following analysis. The perceived traffic load of BS j evolves as follows: after BSs have broadcast $\boldsymbol{\rho}(k)$, users select their associating BSs according to the user side algorithm; based on the user-BS associations, BSs calculate their perceived traffic load $M_j(\boldsymbol{\rho}(k))$.

After having derived the perceived traffic load, BSs update their next advertising traffic load.

$$\boldsymbol{\rho}(k+1) = \delta\boldsymbol{\rho}(k) + (1-\delta)\boldsymbol{M}(\boldsymbol{\rho}(k)). \tag{21}$$

Here, $\boldsymbol{M}(\boldsymbol{\rho}(k)) = (M_1(\boldsymbol{\rho}(k)), M_2(\boldsymbol{\rho}(k)), \cdots, M_{|\mathcal{B}|}(\boldsymbol{\rho}(k)))$, and $0 < \delta < 1$ is an exponential averaging parameter.

Since both $\boldsymbol{M}(\boldsymbol{\rho}(k))$ and $\boldsymbol{\rho}(k)$ are defined on $[0, 1-\epsilon]$, $\boldsymbol{M}(\boldsymbol{\rho}(k))$ is a continuous mapping to itself. According to the Brouwer's fixed-point theorem, there exists a solution $\boldsymbol{\rho}^* = \boldsymbol{M}(\boldsymbol{\rho}^*)$.

Lemma 1. *$\psi(\boldsymbol{\rho})$ is a convex function of $\boldsymbol{\rho}$ when ρ_j is defined in $[0, 1-\epsilon]$, $\forall j \in \mathcal{B}$.*

Proof: The lemma can be proved by showing $\bigtriangledown^2\psi(\boldsymbol{\rho}) > 0$ when $\rho_j \in [0, 1-\epsilon]$, $\forall j \in \mathcal{B}$. ∎

Lemma 2. *When $\boldsymbol{\rho}(k) \neq \boldsymbol{\rho}^*$, $\boldsymbol{M}(\boldsymbol{\rho}(k))$ provides a descent direction of $\psi(\boldsymbol{\rho})$ at $\boldsymbol{\rho}(k)$.*

Proof: The lemma is proved by showing $\langle \bigtriangledown\psi(\boldsymbol{\rho})|_{\boldsymbol{\rho}=\boldsymbol{\rho}(k)}, \boldsymbol{M}(\boldsymbol{\rho}(k)) - \boldsymbol{\rho}(k)\rangle < 0$. Let $\eta_j^m(x)$ and $\eta_j(x)$ be the user association indication of BS j that result in the traffic load $M_j(\boldsymbol{\rho}(k))$ and $\rho_j(k)$, respectively.

$$\begin{aligned} &\langle \bigtriangledown\psi(\boldsymbol{\rho})|_{\boldsymbol{\rho}=\boldsymbol{\rho}(k)}, \boldsymbol{M}(\boldsymbol{\rho}(k)) - \boldsymbol{\rho}(k)\rangle \\ &= \sum_{j\in\mathcal{B}} \frac{(M_j(\boldsymbol{\rho}(k)) - \rho_j(k))}{(1-\rho_j(k))^2\phi_j(k)} \\ &= \sum_{j\in\mathcal{B}} \frac{\int_{x\in\mathcal{A}} \lambda(x)\nu(x)(\eta_j^m(x) - \eta_j(x))dx}{r_j(x)(1-\rho_j(k))^2\phi_j(k)} \\ &= \int_{x\in\mathcal{A}} \lambda(x)\nu(x) \sum_{j\in\mathcal{B}} \frac{\eta_j^m(x) - \eta_j(x)}{r_j(x)(1-\rho_j(k))^2\phi_j(k)} dx. \end{aligned} \tag{22}$$

Since

$$\eta_j^m(x) = \begin{cases} 1, \text{for } j = b^k(x) \\ 0, \text{for } otherwise, \end{cases} \tag{23}$$

$$\sum_{j\in\mathcal{B}} \frac{\eta_j^m(x) - \eta_j(x)}{r_j(x)(1-\rho_j(k))^2\phi_j(k)} \leq 0. \tag{24}$$

[1]Since the time scale of the traffic arrival and departure process is typically less than several minutes, the user-BS association process is at a time scale of several minutes. The solar power generation is usually modeled at a time scale of a hour. Thus, this assumption is reasonable.

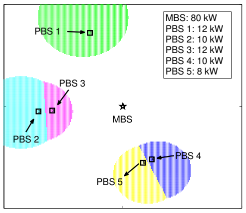

Fig. 2. The max. rate algorithm.

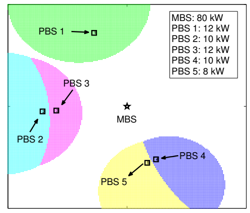

Fig. 3. The α-distributed algorithm [4].

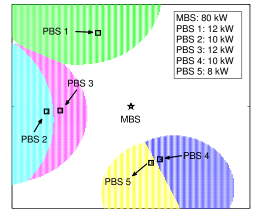

Fig. 4. The GALA algorithm ($\theta = 0.6$).

Because $\rho_j(k) \neq \rho_j^*$,

$$\sum_{j \in \mathcal{B}} \frac{\eta_j^m(x) - \eta_j(x)}{r_j(x)(1-\rho_j(k))^2 \phi_j(k)} < 0. \quad (25)$$

Thus, $\langle \bigtriangledown \psi(\boldsymbol{\rho})|_{\boldsymbol{\rho}=\boldsymbol{\rho}(k)}, \boldsymbol{M}(\boldsymbol{\rho}(k)) - \boldsymbol{\rho}(k) \rangle < 0$. ∎

Theorem 1. *The traffic load vector $\boldsymbol{\rho}$ converges to the optimal traffic load vector $\boldsymbol{\rho}^*$ that minimizes $\psi(\boldsymbol{\rho})$.*

Proof:

$$\begin{aligned} &\boldsymbol{\rho}(k+1) - \boldsymbol{\rho}(k) \\ &= \delta\boldsymbol{\rho}(k) + (1-\delta)\boldsymbol{M}(\boldsymbol{\rho}(k)) - \boldsymbol{\rho}(k) \\ &= (1-\delta)(\boldsymbol{M}(\boldsymbol{\rho}(k)) - \boldsymbol{\rho}(k)). \end{aligned} \quad (26)$$

Since $\boldsymbol{M}(\boldsymbol{\rho}(k))$ gives a descent direction of $\psi(\boldsymbol{\rho})$ at $\boldsymbol{\rho}(k)$ and $0 < \delta < 1$, $\boldsymbol{\rho}(k+1)$ also provides a descent direction of $\psi(\boldsymbol{\rho})$ at $\boldsymbol{\rho}(k)$. Then, $\psi(\boldsymbol{\rho}(k+1)) < \psi(\boldsymbol{\rho}(k))$ until $\boldsymbol{\rho}(k+1) = \boldsymbol{\rho}(k)$. Therefore, the traffic load vector $\boldsymbol{\rho}$ converges to the optimal traffic load vector $\boldsymbol{\rho}^*$ that minimizes $\psi(\boldsymbol{\rho})$. ∎

VI. Simulation Results

Simulations are set up to evaluate the performance of the proposed GALA user association algorithm in heterogeneous cellular networks. In the simulation, one MBS and five PBSs are deployed in a $1000m \times 1000m$ area. The MBS's transmission power is 20 W, and each PBS's transmission power is 2 W. We adopt COST 231 Walfisch-Ikegami [17] as the propagation model with 9 dB rayleigh fading and 5 dB shadowing fading. The carrier frequency is 2110 MHz, the bandwidth is 5 MHz, the antenna feeder loss is 3 dB, the transmitter gain is 1 dB, the noise power level is -104 dBm, and the receiver sensitivity is -97 dBm. The green energy generation rate is consistent during the simulation, and the BSs' energy-latency coefficients are set to be the same. We compare our algorithm with the maximum rate algorithm and the α-distributed algorithm [4]. In the maximum rate algorithm, the users select BSs based on their data rates which are determined by the users' perceived SINRs. The α-distributed algorithm [4] consists of several optimization policies for user-BS associations to balance the flow level traffic among BSs. We compare our algorithm with the α-distributed algorithm implemented with the latency minimization policy.

The coverage area of the algorithms are shown in Figs. 2, 3 and 4, respectively. In the figures, the blank area is the coverage area of the MBS. The maximum rate algorithm associates more users to the MBS because users usually receive the pilot signal with higher SINR from the MBS. As a result of such association, the MBS will be very congested. Considering the load balancing, the α-distributed algorithm offloads traffic from the MBS to PBSs to minimize the network latency. Therefore, the coverage areas of PBSs increase. Since PBSs usually consume less power than the MBS, the GALA algorithm further offloads traffic from the MBS to PBSs to reduce power consumption. During the traffic offloading, GALA is to optimize the utilization of green energy to reduce the on-grid power consumption. For example, in the simulation, PBS 4 has a larger green energy generation rate than PBS 5. To reduce the on-grid power consumption, GALA increases the coverage area of PBS 4 and reduces that of PBS 5.

Figs. 5 and 6 compare the on-grid power consumption and the average delivery latency of these algorithms, respectively. In Fig. 5, we can see that the on-grid power consumption of the maximum rate algorithm is around 60 kW, which is significantly higher than the α-distributed algorithm and the GALA algorithm. This is because the maximum rate algorithm leads to severe traffic congestion in the MBS. By balancing traffic loads among BSs, the α-distributed algorithm reduces the on-grid power consumption of the networks. The GALA algorithm considers the green energy generation rate in BSs and optimizes the utilization of green energy. As a result, the GALA algorithm further saves about 27% on-grid power as compared with the α-distributed algorithm. As shown in Fig. 6, the GALA algorithm does not significantly increase the traffic delivery latency. In the simulation, as compared with the α-distributed algorithm, the latency increased by the GALA algorithm is upto 5%.

Figs. 7 and 8 show the on-grid power consumption and the average traffic delivery latency of the GALA algorithm with different energy-latency coefficient θ, respectively. A larger energy-latency coefficient indicates that the BSs are more energy sensitive. As a result, the GALA algorithm achieves less on-grid power consumption. Meanwhile, the average traffic delivery latency is sacrificed in order to achieve on-grid power savings. However, as shown in Fig. 8, the increase of the traffic delivery latency is not very significant.

VII. Conclusion

In this paper, we have proposed a distributed user-BS association scheme referred to as GALA. During the procedure of establishing user-BS associations, the GALA algorithm not only considers the traffic delivery latency, but also considers

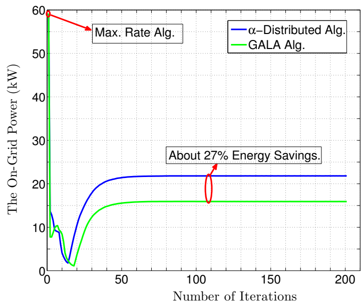

Fig. 5. The on-grid power consumption comparison ($\theta = 0.6$).

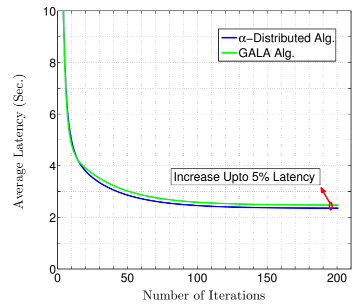

Fig. 6. The average traffic delivery latency comparison ($\theta = 0.6$).

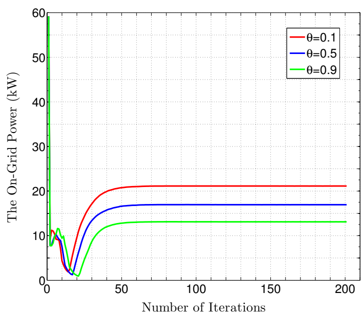

Fig. 7. The on-grid power consumption of GALA with different θ.

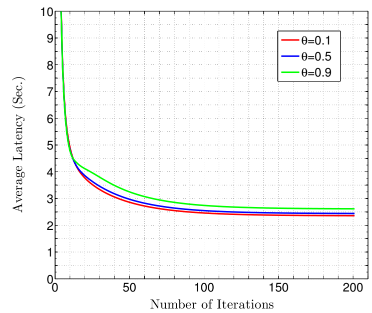

Fig. 8. The average traffic delivery latency of GALA with different θ.

the green energy generation rate in BSs. By optimizing the user-BS associations, the GALA algorithm reduces the on-grid power consumption and at the same time avoids congesting the BSs. Since the GALA algorithm involves an energy-latency trade-off, it increases the average traffic delivery latency of the network. However, the simulation results show that the increase of the traffic delivery latency by the GALA algorithm is not significant. In addition, the energy-latency trade-off of the GALA algorithm is controllable by adjusting the energy-latency coefficient. Therefore, the GALA algorithm enables the trade-off between the on-grid power consumption and the traffic delivery latency, and saves a significant amount of on-grid power in heterogeneous cellular networks powered by both on-grid power and green energy.

REFERENCES

[1] M. Etoh, T. Ohya, and Y. Nakayama, "Energy consumption issues on mobile network systems," in *International Symposium on Applications and the Internet*, Turku, Finland, Aug. 2008.

[2] Ericson Inc., "Sustainable energy use in mobile communications," Aug. 2007, white Paper.

[3] T. Han and N. Ansari, "Optimizing cell size for energy saving in cellular networks with hybrid energy supplies," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM'12)*, Anaheim, CA, USA, Dec. 2012.

[4] H. Kim, G. de Veciana, X. Yang, and M. Venkatachalam, "Distributed $\alpha-$optimal user association and cell load balancing in wireless networks," *IEEE/ACM Transactions on Networking*, vol. 20, no. 1, pp. 177–190, Feb. 2012.

[5] H.-S. Jo, Y. J. Sang, P. Xia, and J. Andrews, "Heterogeneous cellular networks with flexible cell association: A comprehensive downlink sinr analysis," *IEEE Transactions on Wireless Communications*, vol. 11, no. 10, pp. 3484–3495, Oct. 2012.

[6] S. Corroy, L. Falconetti, and R. Mathar, "Dynamic cell association for downlink sum rate maximization in multi-cell heterogeneous networks," in *Proceedings of IEEE International Conference on Communications (ICC'12)*, Ottawa, CA, Jun. 2012.

[7] L. Correia, D. Zeller, O. Blume, D. Ferling, Y. Jading, G. Auer, and L. Van Der Perre, "Challenges and enabling technologies for energy aware mobile radio networks," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 66 –72, Nov. 2010.

[8] Z. Hasan, H. Boostanimehr, and V. Bhargava, "Green cellular networks: A survey, some research issues and challenges," *IEEE Communications Surveys and Tutorials*, vol. 13, no. 4, pp. 524–540, 2011.

[9] T. Han and N. Ansari, "On greening cellular networks via multicell cooperation," *IEEE Wireless Communications Magazine*, vol. 20, no. 1, pp. 82–89, 2013.

[10] ——, "ICE: Intelligent cell breathing to optimize the utilization of green energy," *IEEE Communications Letters*, vol. 16, no. 6, pp. 866–869, 2012.

[11] H. A. H. Hassan, L. Nuaymi, and A. Pelov, "Classification of renewable energy scenarios and objectives for cellular networks," in *Proceedings of IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'13)*, London, UK, Sep. 2013.

[12] J. Zhou, M. Li, L. Liu, X. She, and L. Chen, "Energy source aware target cell selection and coverage optimization for power saving in cellular networks," in *Proceedings of the 2010 IEEE/ACM Int'l Conference on Green Computing and Communications*, Hangzhou, China, Dec. 2010.

[13] T. Han and N. Ansari, "On optimizing green energy utilization for cellular networks with hybrid energy supplies," *IEEE Transactions on Wireless Communications*, 2013, to appear.

[14] I. Goiri, K. Le, T. D. Nguyen, J. Guitart, J. Torres, and R. Bianchini, "Greenhadoop: leveraging green energy in data-processing frameworks," in *Proceedings of the 7th ACM european conference on Computer Systems*, Bern, Switzerland, Apr. 2012.

[15] O. Arnold, F. Richter, G. Fettweis, and O. Blume, "Power consumption modeling of different base station types in heterogeneous cellular networks," in *IEEE Future Network and Mobile Summit*, Florence, Italy, Jun. 2010.

[16] L. Kleinrock, *Queueing Systems: Computer applications*. Wiley-Interscience, 1976, ISBN: 978-0471491118.

[17] "Evolution of land mobile radio (including personal) ccommunications: Cost 231." [Online]. Available: http://www.awe-communications.com/Propagation/Urban/COST/