QoS-Aware Dynamic Cell Reconfiguration for Energy Conservation in Cellular Networks

Kyuho Son†, Santosh Nagaraj‡, Mahasweta Sarkar‡ and Sujit Dey*
†T-Mobile USA, Bellevue, WA 98006
‡San Diego State University, San Diego, CA 92182
*University of California, San Diego, La Jolla, CA 92093
E-mail: kyuho.son@gmail.com, {snagaraj,msarkar2}@mail.sdsu.edu, dey@ece.ucsd.edu

***Abstract*—Given the significant energy consumption in operating base stations (BSs), improving their energy efficiency is an important problem in cellular networks. To this end, this paper proposes a novel framework, called DCR (dynamic cell reconfiguration) that dynamically adjust the set of active BSs and user association according to user traffic demand for energy conservation. In order to overcome prohibitive computational complexity in finding an optimal solution, we take an approach to design simple yet effective algorithms. We demonstrate that the proposed framework is not only computationally efficient but also can achieve the performance close to the optimum solution from an exhaustive search. Through simulations based on a real dataset of BS topology and utilization, we show that DCR can yield about a 30-40% reduction compared to the conventional static scheme where all BSs are always turned on.**

I. Introduction

With the explosive growth in wireless communication usage and infrastructure, energy use of cellular wireless networks has lately become a critical issue [1]. Designers of communication and networking algorithms and protocols have traditionally ignored the complexity and power consumption at base stations (BSs) and instead focused on improving energy efficiency to prolong battery life-time of mobile terminals (MTs) only. Today, however, the situation has changed. Pushed by ever increasing energy costs and environmental concerns, all industries are seeking ways to reduce their energy consumption. Therefore, improving the energy efficiency of BSs, which have been identified to be the most power consuming part in current cellular networks [2], [3], is as important as in MTs.

In recent years, there have been many researches on green cellular networks, which includes different techniques presented in [2]–[14]. For example, the authors in [4]–[6] investigate the micro BSs or relay deployment strategy in terms of energy, capacity and cost. In [2], [8]–[10], several load-aware BS switching on/off algorithms were proposed. The greening effect of interference management with combinations of spatial and temporal power budget sharing is investigated in [11]. In [12], a computation unit deceleration technique has been proposed, which can conserve dynamic power effectively without turning off BSs. With cell zooming [13], the coverage of each active BS is dynamically varied so that the overall power consumption in the network can be minimized. There has also been work on modeling the component-level power consumption in BSs [14].

In this paper, we consider a problem of minimizing the total power consumption in BSs while satisfying the quality of service (QoS) requirements for all users in the network. To this end, we develop a novel framework for energy conservation, called *DCR (dynamic cell reconfiguration)*, which includes the active BS selection and user association algorithms. Because solving the problem is computationally intensive, we take an approach to design simple yet effective algorithms. Through analytical and simulation studies, we demonstrate that the proposed DCR framework is not only computationally efficient but also can achieve the performance close to the optimal solution that could be obtained from an exhaustive search.

The rest of this paper is organized as follows. Section II formally describes our system model and general problem. In Section III, we propose a dynamic cell reconfiguration framework. In Section IV, we present simulation results for the proposed algorithm. In Section V, we conclude the paper with our notes and observations.

II. System Model

A. Network and Channel Model

Let us consider a cellular wireless network with a set of BSs $\mathcal{B}$. Let $x \in \mathcal{L}$ denote a user location and $i \in \mathcal{B}$ be the index of the i-th BS. In this paper, although we concentrate on downlink communication that is a primary usage mode for mobile Internet, i.e., from BSs to MTs.

The time-averaged transmission rate of a user at location x served by BS i is denoted by $c_i(x)$ [bits/sec]. It depends on the set of active BSs $\mathcal{B}_{\text{on}}$ and their transmit power $p^{\text{tx}} = (p_1^{\text{tx}}, \cdots, p_{|\mathcal{B}|}^{\text{tx}})$. Further, note that $c_i(x)$ can capture the effect of shadowing.

B. Traffic Demand and BS Utilization

We assume that a user at location x have the required rate $\gamma(x)$ [bits/sec]. Then, in order to guarantee the QoS level of the user, the fraction of resource blocks (i.e., time/frequency) allocated by BS i would be $\gamma(x)/c_i(x)$.

We now define a routing probability $\pi_i(x)$ which dictates the user at location x is routed to BS i. As can be seen later in

978-1-4673-5939-9/13/$31.00 ©2013 IEEE

Section III-A, the optimal $\pi_i(x)$ would be either two extremes 0 or 1. The BS utilization, the average occupied percentage of the BS resource blocks, can be defined as follows:

$$\rho_i \doteq \int_{\mathcal{L}} \frac{\gamma(x)}{c_i(x)} \pi_i(x) dx. \tag{1}$$

Definition 2.1 (Feasible set): When the set of active BSs $\mathcal{B}_{\text{on}}$ and their transmit power p^{tx} are given, the set $\mathcal{F}(\mathcal{B}_{\text{on}}, p^{\text{tx}})$ of *feasible* utilization ρ can be defined as follows:

$$\begin{aligned} \mathcal{F}(\mathcal{B}_{\text{on}}, p^{\text{tx}}) \doteq \Big\{ & \rho = (\rho_1, \cdots, \rho_{|\mathcal{B}|}) \ \mid \ 0 \preceq \rho \preceq 1, \\ & \forall x \in \mathcal{L}, \quad 0 \preceq \pi(x) \preceq 1, \\ & \forall x \in \mathcal{L}, \quad \textstyle\sum_{i \in \mathcal{B}} \pi_i(x) = 1, \\ & \forall x \in \mathcal{L}, \ \forall i \in \mathcal{B} \backslash \mathcal{B}_{\text{on}}, \quad \pi_i(x) = 0 \Big\}. \end{aligned} \tag{2}$$

where we use "$\preceq$" to denote element-wise inequality for the vectors.

C. Power Consumption Model

We consider the BS power consumption model that can capture both *dynamic* power and *static* power as follows. The former is proportional to BS's utilization. On the other hand, the latter is the fixed amount of power that BSs dissipate while even inactive.

$$P_i \ = \ \underbrace{(1-q_i)\rho_i P_i^{\max}}_{dynamic} \ + \ \underbrace{q_i P_i^{\max}}_{static}, \tag{3}$$

where $q_i \in [0,1]$ is the portion of the static power for BS i, and $P_i^{\max}$ is the maximum power consumption when it is fully utilized. According to [14], $P_i^{\max}$ is again a function of the transmit power

$$P_i^{\max} \ = \ a_i p_i^{\text{tx}} \ + \ b_i, \tag{4}$$

where the coefficient a_i accounts for the power consumption that scales with the average transmit power due to amplifier, cooling, feeder losses, etc.

We would like to emphasize that our model given in (3) is general enough to grasp a variety of BS power consumption.

- *Energy-proportional BS* with $q_i = 0$: Assuming ideally equipped with energy-proportional devices, the BS does not consume any power when idle, and proportionally consume more power as its utilization increases.
- *Non-energy-proportional BS* with $q_i > 0$: In practice, several hardware devices inside a BS dissipate standby power even though the BS does not serve any traffic. As an extreme case of $q_i = 1$, the model becomes a constant consumption, which has been widely used in many works in literature [2], [3], [9], [10].

D. General Problem Statement

We consider a general problem that minimizes the total BS power consumption while guaranteeing QoS requirements for all users in a sense that all of their traffics are guaranteed to be served.

$$\begin{aligned} \text{[GP]:} \qquad \min \qquad & \sum_{i \in \mathcal{B}_{\text{on}}} P_i & (5) \\ \text{subject to} \qquad & \mathcal{B}_{\text{on}} \subseteq \mathcal{B}, & (6) \\ & p^{\text{tx}} \preceq p^{\text{tx,max}}, & (7) \\ & \rho \in \mathcal{F}(\mathcal{B}_{\text{on}}, p^{\text{tx}}). & (8) \end{aligned}$$

Our ultimate goal is to develop a framework for BS energy conservation that encompasses *(i)* active BS selection, *(ii)* user association and *(iii)* transmit power control. As a first step towards this goal, in this paper, we focus on building solutions for the first two subproblems assuming all BSs are operating at the maximum transmit power, i.e., $p^{\text{tx}} = p^{\text{tx,max}}$ instead of the constraint (7). We plan to integrate the transmit power control into a single unified framework in future work. Note that we drop the transmit power p^{tx} to keep our notations simple throughout the paper.

III. Dynamic Cell Reconfiguration Framework

In this section, we present details on our framework, called *dynamic cell reconfiguration (DCR)*, that includes the user association and active BS section algorithms.

A. User Association

We shall start by considering a given set of active BSs $\mathcal{B}_{\text{on}}$. In this case, the static power consumption term can be ignored. So the remaining subproblem in [GP] is to determine which BS each user should be associated to, or equivalently, to find an optimal BS utilization ρ.

$$\begin{aligned} \min \quad & \sum_{i \in \mathcal{B}_{\text{on}}} \left[(1-q_i) P_i^{\max} \rho_i + L_i(\rho_i) \right] \\ \text{subject to} \quad & \rho \in \mathcal{F}(\mathcal{B}_{\text{on}}), \end{aligned} \tag{9}$$

where $L_i(\rho_i)$ is a penalty function we intentionally introduce. By adding the penalty into the objective, we can allow the system to balance the traffic load among BSs and avoid a cell getting too congested. Although there may be other method of penalizing the congested cell for the purpose of load balancing, in this paper, we introduce the following penalty function with three configurable parameters.

$$L_i(\rho_i) = \begin{cases} 0, & \rho < \rho_{th}, \\ L_{\max} \cdot \left(\dfrac{\rho_i - \rho_{th}}{1 - \rho_{th}} \right)^{\beta}, & \rho \geq \rho_{th}, \end{cases} \tag{10}$$

where $L_{\max} \geq 0$ is the maximum penalty value and $\rho_{th} \in [0,1]$ is the BS utilization threshold we start penalizing the BS; $\beta > 0$ controls the sharpness of the penalty function. It is worthwhile mentioning that the modified problem given in (9) is asymptotically equivalent to the original subproblem without the penalty function L_i in any of the following conditions: as $L_{\max}$ goes to zero, ρ_{th} goes to one, or β goes to infinity.

Theorem 1: When the problem given in (9) is feasible, Then, the optimal policy is for user at location x to associate with BS $i^*(x)$, given by

$$i^*(x) = \operatorname*{argmax}_{i \in \mathcal{B}_{\text{on}}} c_i(x) \cdot [(1-q_i)P_i^{\max} + L_i'(\rho_i^*)]^{-1}, \quad (11)$$

where ρ^* is the optimal BS utilization.

But the subtlety is here that we have a chicken-and-egg problem because the policy in (11) needs requires the optimal utilization ρ^* in advance. However, the following iterative algorithm does not require such an assumption and converges to the global optimum without knowing ρ^*.

User Association Algorithm

MTs: At the k-th iteration, MTs measure the average transmission rate $c_i(x)$ and receive the BS utilization $\rho^{[k]}$. Then, the MTs select the BS $i^{[k]}(x)$ according to (11) by using the current $\rho^{[k]}$ instead of the optimal utilization ρ^*.

BSs: During k-th period, each BS estimates its average utilization $\rho_i^{[k+1]}$. Then, the BS broadcast this information to MTs for the next iteration.

With a slight modification of the technique used in [15], we can generalize the optimality and convergence proofs of the proposed user association algorithm, although they are omitted here due to space limitations.

B. Active BS Selection

We investigate another piece of subproblem in DCR, i.e., active BS selection. By solving this problem, we will be able to answer which BSs remain turning on to guarantee the QoS level of users and which BSs need to minimize energy consumption in the network.

$$\text{[P1]:} \quad \min_{\mathcal{B}_{\text{on}} \subseteq \mathcal{B}} \quad \text{UA}(\mathcal{B}_{\text{on}}) + \sum_{i \in \mathcal{B}_{\text{on}}} q_i P_i^{\max},$$

where $\text{UA}(\mathcal{B}_{\text{on}}) \doteq \min_{\rho \in \mathcal{F}(\mathcal{B}_{\text{on}})} \sum_{i \in \mathcal{B}_{\text{on}}} (1-q_i)\rho_i P_i^{\max}$, which is the optimal objective value of user association problem.

There is a technical challenge in solving this problem because it can be reduced from a vertex cover problem which is theoretically known as NP-complete [16]. In order to overcome such a high computational complexity, we consider the design of an efficient heuristic algorithm in this section. To that end, we move the static power consumption term in the objective to the constraint with a nonnegative budget.

$$\begin{aligned} \text{[P2]:} \quad & \min_{\mathcal{B}_{\text{on}} \subseteq \mathcal{B}} && \text{UA}(\mathcal{B}_{\text{on}}) \\ & \text{subject to} && \sum_{i \in \mathcal{B}_{\text{on}}} q_i P_i^{\max} \le Z/\lambda. \end{aligned}$$

As can be easily noticed, there is a close relationship between [P1] and [P2] as primal/dual problems with a Lagrangian multiplier λ. In order to further convert [P2], let us introduce an intuitive diminishing returns property that is formalized by the concept of submodularity [17] as follows.

Definition 3.1: For a real-valued set function H, we define the discrete derivative at $\mathcal{A} \subseteq \mathcal{S}$ in direction $s \in \mathcal{S}$ as $d_s(\mathcal{A}) = H(\mathcal{A} \cup s) - H(\mathcal{A})$. The H is said to be *submodular* if

$$\mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \mathcal{S} \Rightarrow d_s(\mathcal{A}_1) \ge d_s(\mathcal{A}_2) \text{ for all } s \in \mathcal{S} \setminus \mathcal{A}_2.$$

Now we rewrite [P2] in the standard form of submodular maximization problem as follows.

$$\begin{aligned} \text{[P3]:} \quad & \max_{\mathcal{A} \subseteq \mathcal{B} \setminus \mathcal{B}_{\text{init}}} && H(\mathcal{A}) \\ & \text{subject to} && c(\mathcal{A}) = \sum_{i \in \mathcal{A}} c(i) \le C. \end{aligned}$$

where $\mathcal{A} = \mathcal{B}_{\text{on}} \setminus \mathcal{B}_{\text{init}}$, $H(\mathcal{A}) = \text{UA}(\mathcal{B}_{\text{init}}) - \text{UA}(\mathcal{B}_{\text{init}} \cup \mathcal{A})$, $c(i) = q_i P_i^{\max}$ and $C = Z/\lambda - \sum_{i \in \mathcal{B}_{\text{init}}} c(i)$.

It is worthwhile mentioning that there exist a very intuitive yet efficient greedy algorithm for [P3] only if H is a nondecreasing submodular. It works as follows: starting from the empty set $\mathcal{A} = \emptyset$, it iteratively adds the element with the highest value of metric $\frac{H(\mathcal{A} \cup i) - H(\mathcal{A})}{c(i)}$ while the total cost is within the budget C. Mathematically, it has been shown in [17], [18] that this greedy heuristic can give a suboptimal solution with an approximation factor of $(1 - 1/e)$.

We can fortunately show the set function H is nondecreasing submodular under a reasonable assumption that adding or removing one additional BS has marginal impact on the total amount of interference. This implies the greedy heuristic would work well in our active BS selection problem as well. After some tweaking to suit [P1] better, we propose a greedy-style active BS selection algorithm that borrows the metric (i.e., the decrement per unit cost when removing BS i) as follows.

Active BS Selection Algorithm

1: *Initialize*: $\mathcal{B}_{\text{on}} = \mathcal{B}$
2: *Repeat*:
3: Find $i^* = \operatorname{argmin}_{i \in \mathcal{B}_{\text{on}}} \frac{\text{UA}(\mathcal{B}_{\text{on}} \setminus \{i\}) - \text{UA}(\mathcal{B}_{\text{on}})}{q_i P_i^{\max}}$
4: If $\text{UA}(\mathcal{B}_{\text{on}} \setminus \{i\}) - \text{UA}(\mathcal{B}_{\text{on}}) < q_i P_i^{\max}$,
5: then $\mathcal{B}_{\text{on}} \leftarrow \mathcal{B}_{\text{on}} - \{i^*\}$
6: Else,
7: go the Finish.
8: *Finish*: $\mathcal{B}_{\text{on}}$ is the set of active BSs.

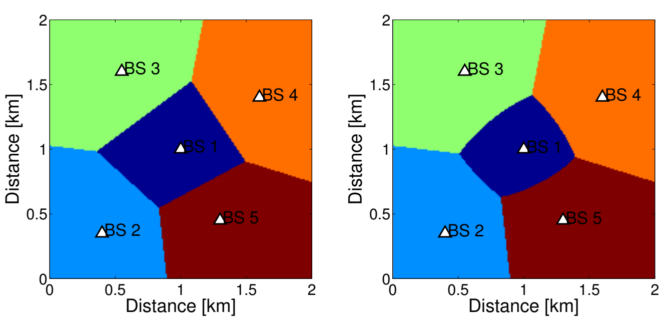
Our proposed algorithm starts from the point where all BSs are turning on and finds the best BS as a turning-off candidate for energy conservation in line 3. Note that the denominator is *the amount of static power consumption saving* from turning off when BS i. On the other hand, the numerator is *the increment of dynamic power consumption*, which comes from the fact that MTs originally associated with the switched-off BS would see possibly lower transmission rate $c_i(x)$ due to father distance to the new serving BS. In line 4, the algorithm checks whether there is a net energy saving (in other words, the decrement in static power consumption is larger than the increment in dynamic power consumption). If so, we turn off BS i. Otherwise, we stop the algorithm.

IV. Simulation Results

We evaluate the performance of the proposed DCR framework though simulations. Typical transmit power for macro BSs and their maximum operational power are considered to be 20W and 865W according to [14], respectively. The static power portion q_i is assumed to be 0.5 unless explicitly mentioned, but we will examine the effect of varying this parameter in Section IV-C.

A. Load balancing via Penalty-based User Association

Consider a simple network composed of five active BSs and spatially heterogeneous traffic loads, i.e., the required rate $\propto (\max(r) - r)^5$ where r is the distance from the center. So the area in the middle, mostly covered by BS 1, can be interpreted as hotspot. In order to see how the proposed user association algorithm balances traffic loads, we plot Fig. 1 illustrating snapshots of BSs' coverage areas for the cases (a) without and (b) with penalty function. We can easily notice the effect of introducing the penalty function L_i into the reconfiguration algorithm by comparing two figures. With penalty, some users leave the congested BS 1, as indicated by the shrinking of cell 1 in Fig. 1, and associate with neighboring BSs 2-5, which are actually under-utilized.



(a) Without penalty ($\rho_{th} = 1$) (b) With penalty ($\rho_{th} = 0.5$)

Fig. 1. Snapshots of cell coverage: $L_{\max} = \sum_{i \in \mathcal{B}} P_i^{\max}$ and $\beta = 2$.

Such a load balancing comes at the cost of slight increase in dynamic power consumption. In Fig. 2, we calculate the average delay to transfer a filesize of 100 Kbyte and the worst delay in the network as a yardstick of load balancing. The less delay means the less congestion (i.e., the more effective load balancing). As shown in Fig. 2, the power cost is marginal compared to the delay benefit we can expect. For example, in the case of $\rho_{th} = 0.7$, there are 39% and 47% reductions in the average and worst delay, with 0.56% (2838W to 2854W) increase in power consumption. Note that this tradeoff graph may also be used to choose ρ_{th} in practice based on the maximum tolerable delay. In the following simulation of Section IV-B, we set $\rho_{th} = 0.7$.

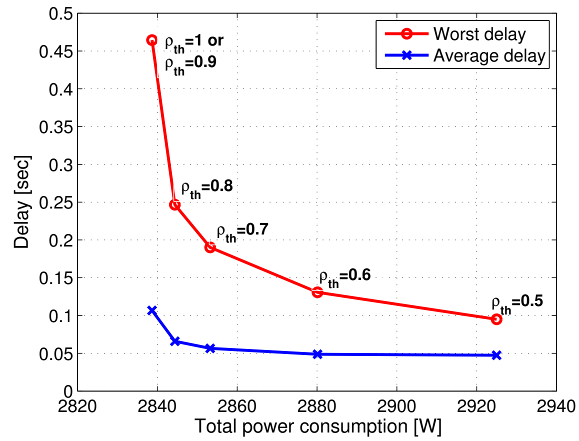


Fig. 2. Tradeoff between delay and total power consumption by varying the BS utilization threshold ρ_{th} from 0.5 to 1.0.

B. Energy Saving by the Proposed DCR Framework

Now let us investigate the performance of the whole DCR framework including both user association and active BS selection. To have more realistic results, a topology with fifteen BSs in 4.5×4.5 km^2, a part of 3G network in metropolitan area [19], is adopted (see Fig. 3). For comparison, we also consider three other schemes:

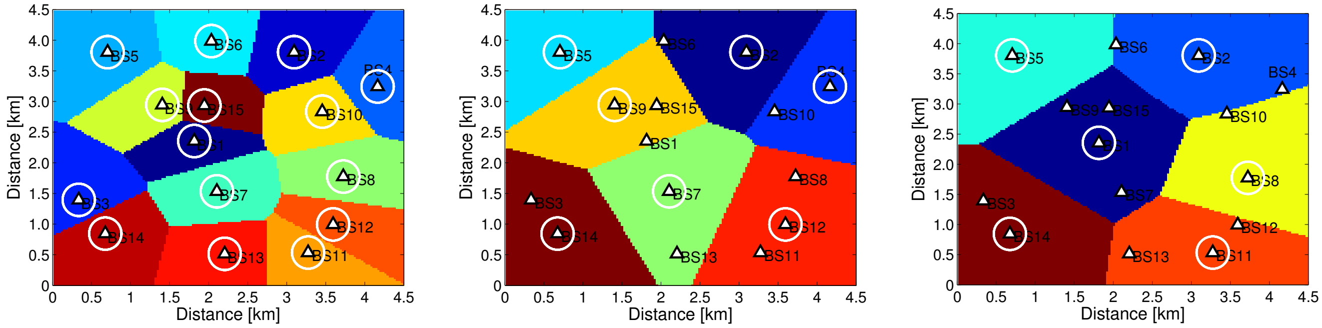
- *All-On (baseline scheme)*: always turning on all BSs
- *Util-based*: turning off the least utilized BS one by one which is shown to be an effective heuristic in [8]
- *Exhaustive*: the optimal solution from an exhaustive search

Fig. 3 shows snapshots of the active BSs and their coverage areas at the normalized traffic load[1] = 0.3. All-On still keeps turning on all BSs at such a low load, which naturally leads to the energy inefficiency. However, DCR and Exhaustive turn off eight and nine BSs for energy conservation, respectively. As a consequence, the remaining BSs dynamically reconfigure their cells (i.e., cell zooming). In our simulations under various configurations, DCR often finds a suboptimal solution that has the same number of active BSs as Exhaustive and just one or two more in the worst case. It is also worthwhile investigating the static and dynamic power consumption breakdown: DCR (4.46kW = 3.03kW + 1.43kW) vs. Exhaustive (4.25kW = 2.60kW + 1.65kW).

Fig. 4 shows the total power consumption of the cellular network as a function of the normalized traffic load in both (a) uniform and (b) non-uniform[2] traffic distribution. Our results show that a brute-force Util-Based works well in the uniform environment, but not in non-uniform environment. However, the proposed DCR always outperforms Util-Based, and moreover its performance is very close to that of the exhaustive search solution. Compared to static All-On, DCR

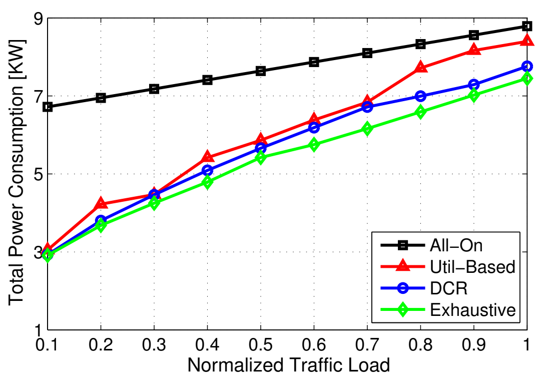
[1] In our simulation, the normalized traffic load [no unit] is the traffic load normalized by the traffic load at peak time.

[2] A linearly decreasing traffic along the diagonal direction from left top to right bottom in Fig. 3 is considered to generate non-uniform environment.

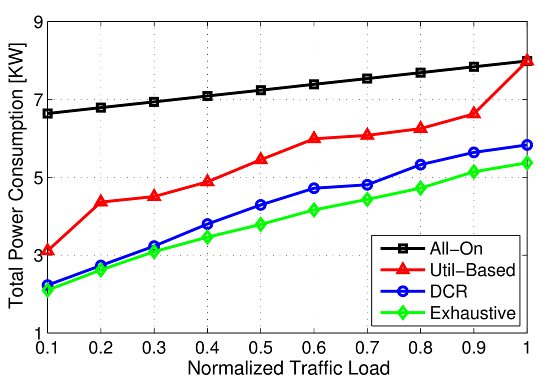


(a) All-On (b) DCR (c) Exhaustive

Fig. 3. Snapshots of the active BSs and their coverage areas at the normalized traffic load = 0.3 for different schemes.



(a) Uniform traffic distribution



(b) Non-uniform traffic distribution

Fig. 4. Normalized traffic load vs. total power consumption.

yields the potential energy savings of 10-60% depending on the amount of traffic and its spatial distribution.

In order to calculate the overall energy saving per day from DCR, in TABLE I, we analyze the fraction of time that the BS utilization is observed to be in different ranges between 0 and 1 based on the sample traffic traces from [3]. Using these numbers as weighting factors along with the results of total power consumption in Fig. 4, we can obtain that about 30-40% savings are realistically achievable during one day.

TABLE I
FRACTION OF TIME IN THE RANGES OF BS UTILIZATION BASED ON THE SAMPLE TRAFFIC TRACES FROM [3]

BS Utilization	Fraction of Time
0.0∼0.1	0.313
0.1∼0.2	0.061
0.2∼0.3	0.077
0.3∼0.4	0.083
0.4∼0.5	0.049
0.5∼0.6	0.038
0.6∼0.7	0.103
0.7∼0.8	0.047
0.8∼0.9	0.184
0.9∼1.0	0.045

C. Effect of the portion of static power consumption q_i

Fig. 5 illustrate the effect of static power fraction q_i on the maximum energy saving (at a very low load ~ 0.1). As expected, there is no gain at $q_i = 0$ because those energy-proportional BSs have no standby power dissipation. However, we can obtain much saving when the static power consumption contributes to a significant portion of the total consumption, e.g., nearly 70% possible at $q = 1$. Given that recent measurement in [14] (implying high q_i) that the BS power consumption varies only about 5% over time regardless of its utilization, the proposed DCR framework would bring huge benefit to the current cellular networks.

V. CONCLUSION AND FUTURE WORK

In this paper, we proposed a novel dynamic cell reconfiguration framework for BS energy saving that includes both user association and active BS selection algorithms in cellular wireless networks. Through analytical and simulation studies, we showed that our DCR framework is not only computationally efficient but also can achieve the performance close to the optimum solution. We also made several interesting observations that high energy savings are expected, especially, when the average traffic load is low and/or the portion of static power consumption is high.

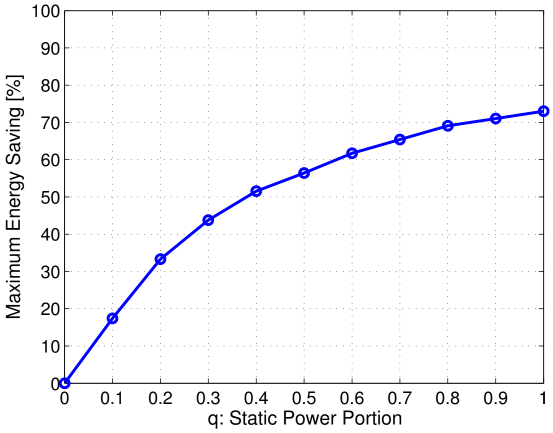


Fig. 5. Effect of static power portion q_i on the maximum energy saving.

The proposed framework brings about many interesting future research opportunities. We are currently working on integrating transmit power control into a unified framework to further improve the energy efficiency of BSs. Additionally, we would like to investigate the impacts presented by DCR on the cellular uplink transmissions.

REFERENCES

[1] K. Son, H. Kim, Y. Yi, and B. Krishnamachari, "Toward energy-efficient operation of base stations in cellular wireless networks," a book chapter of *Green Communications: Theoretical Fundamentals, Algorithms, and Applications, CRC Press, Taylor & Francis*, 2012.

[2] M. A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Optimal energy savings in cellular access networks," in *Proc. GreenComm*, Jun. 2009.

[3] E. Oh, B. Krishnamachari, X. Liu, and Z. Niu, "Towards dynamic energy-efficient operation of cellular network infrastructure," *IEEE Comm. Mag.*, pp. 56–61, Jun. 2011.

[4] K. Son, E. Oh, and B. Krishnamachari, "Energy-aware hierarchical cell configuration: from deployment to operation," in *Proc. IEEE INFOCOM 2011 - GCN Workshop*, Apr. 2011, pp. 289–294.

[5] P. Rost and G. P. Fettweis, "Green communications in cellular networks with fixed relay nodes," in *Proc. GreenComm*, Dec. 2009.

[6] W. Guo and T. O'Farrell, "Capacity-energy-cost tradeoff in small cell networks," in *Proc. IEEE VTC*, May 2012.

[7] F. Richter, A. J. Fehske, and G. P. Fettweis, "Energy efficiency aspects of base station deployment strategies for cellular networks," in *Proc. IEEE VTC*, Sep. 2009.

[8] 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.*, pp. 1525–1536, Sep. 2011.

[9] S. Zhou, J. Gong, Z. Yang, Z. Niu, and P. Yang, "Green mobile access network with dynamic base station energy saving," in *Proc. ACM MobiCom (Poster)*, Beijing, China, Sep. 2009, pp. 1–3.

[10] L. Chiaraviglio, D. Ciullo, M. Meo, M. A. Marsan, and I. Torino, "Energy-aware UMTS access networks," in *Proc. WPMC Symposium*, Lapland, Finland, Sep. 2008.

[11] J. Kwak, K. Son, Y. Yi, and S. Chong, "Impact of spatio-temporal power sharing policies on cellular network greening," in *Proc. WiOpt*, Princeton, NJ, May 2011, pp. 167–174.

[12] K. Son and B. Krishnamachari, "Speedbalance: Speed-scaling-aware optimal load balancing for green cellular networks," in *Proc. IEEE INFOCOM*, Orlando, FL, Mar 2012.

[13] Z. Niu, Y. Wu, J. Gong, and Z. Yang, "Cell zooming for cost-efficient green cellular networks," *IEEE Comm. Mag.*, vol. 48, pp. 74–79, Nov. 2010.

[14] 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*, Jun. 2010.

[15] H. Kim, G. de Veciana, X. Yang, and M. Venkatachalam, "Distributed α-optimal user association and cell load balancing in wireless networks," *IEEE/ACM Trans. Networking.*, vol. 10, no. 1, pp. 177–190, Feb. 2012.

[16] R. M. Karp, "Reducibility among combinatorial problems," *Complexity of computer computations*, vol. 40, no. 4, pp. 85–103, 1972.

[17] G. Nemhauser, L. Wolsey, and M. Fisher, "An analysis of the approximations for maximizing submodular set functions-I," *Mathematical Programming*, vol. 14, no. 1, pp. 265–294, Dec. 1978.

[18] M. Sviridenko, "A note on maximizing a submodular set function subject to knapsack constraint," *Operations Research Letters*, vol. 32, 2004.

[19] K. Son, S. Lee, Y. Yi, and S. Chong, "REFIM: A practical interference management in heterogeneous wireless access networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 6, pp. 1260–1272, Jun. 2011.