2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing

# Delaunay Triangulation Based Green Base Station Operation for Self Organizing Network

Gilsoo Lee and Hongseok Kim(✉)
Department of Electronic Engineering
Sogang University
Seoul, Korea
Email: {gilsoolee, hongseok}@sogang.ac.kr

Young-Tae Kim and Byoung-Hoon Kim
Advanced Communication Technology Lab.
LG Electronics
Seoul, Korea
Email: {youngtae99.kim, bh.kim}@lge.com

***Abstract*—In order to reduce the energy consumption in the cellular network, green self-organizing network (SON) techniques gain the attention from economical and environmental perspective. We study the base station (BS) turning-off operation that can run without knowing the global knowledge of dynamic traffic distribution in time and space. We propose a green SON framework based on an overlay network using Delaunay triangulation (DT). Specifically, we propose a threshold-based BS off algorithm for reliable and operator-friendly operation. The proposed algorithm can be readily implemented and standard compliant; the existing architecture such as X2 interface can be used with minimal modification. We perform an extensive simulations using a realistic urban BS topology, and the results confirm that the proposed algorithm can significantly reduce the energy consumption, e.g., more than 60%, and the energy-delay trade-off curve is almost close to the optimal one, which can be found by the exhaustive search only.**

***Keywords*—Threshold-based, overlay network, Delaunay triangulation, energy-efficient, base station operation, distributed algorithm, flow-level dynamics.**

## I. Introduction

With the deployment of broadband data services, the demands of mobile traffic is rapidly growing [1], but more energy consumption is unavoidable. We observe that like a wide range of network equipments such as the servers and routers in the Internet, base stations (BSs) also consume *constant* energy irrespective of the traffic level, which leads to energy waste when the traffic level is low, e.g., during the night time. Specifically, considering the spatial and/or temporal mobile traffic variation [2], we can turn off the under-utilized BSs while trying to avoid performance degradation. In doing so, mobile operators can save energy and also reduce $CO_2$ emission. In addition, low operational expenditure (OPEX) is achievable due to the reduced electricity cost. In this regard, many efforts are actively going on in the 3GPP LTE standards (e.g., TS36.300 [3]) and also expected to continue in Rel. 12. In addition to the standardization activity, Green Touch tries to enhance the network energy efficiency by 1,000 times [4], and Toward Green 5G Mobile Network (5GrEEn) group has started research to enhance energy efficiency as well [5].

In the literature, there is a lot of ongoing research about BS on/off operation algorithm, e.g., [6]–[12]. A practical method of constructing location-based virtual grid for BS turning-off operation is proposed, but it was mainly based on simulation [12]. Cell zooming technique was proposed in [9] to avoid coverage hole when BSs are turned off. The authors in [9] suggested centralized and distributed algorithms for the trade-off between reducing energy consumption and outage probability. The authors in [10] proposed a framework for dynamic cell planning and BS operation that increases energy efficiency. As an extension of [10], the smaller and denser cell deployment was proposed in network planning stage when cell zooming is not sufficient [11]. In [6], the authors investigated the energy saving with temporal traffic fluctuation and BS density by modeling the traffic as sinusoidal pattern and using the first-order analysis. Recently, [7] proposed a distributed algorithm for BS switching off operation. They suggested a concept of network-impact, which is an amount of effect induced by turning off a BS; the proposed distributed algorithm is based on the prediction of BS loads, but it may not be valid under inhomogeneous traffic environment because of the nonlinearity effect of traffic load on the network performance, e.g., delay.

In this paper, we study the BS turning-off operation considering the spatial and temporal variability of network traffic under flow-level dynamics. Many research in BS greening assumes that a system with a fixed number of users with infinitely backlogged data which is so called a *saturated system*. However, the saturated system model may not capture the dynamic nature of traffic variation in time and space. As opposed to the saturated systems considered in [8], we use *flow-level dynamics* where new flows (or file transfer requests) arrive into the system at random and leave the system after being served [13]–[17]. The research of dynamic network environment is required to analyze the network with approximation to reality. In doing this, we develop a self organizing network (SON) algorithm for BS off operation that works without global knowledge of traffic distribution and thus that can run without a centralized controller. For SON operation, we consider overlay networks, which are defined over the subsets of BSs, either turned on, turned off, or all.

The main contributions of this paper are summarized as follows:

1) First, we develop a framework for BS off operation originated from *overlay networks*. We construct overlay networks using Delaunay triangulation (DT) for dynamic BS off operation. When BSs are turned off, the DT graph of turned-off BSs successfully reflects the topology change of BSs because DT graph can be reconstructed in a distributed manner. Hence, the proposed method is implementable in a


978-0-7695-5046-6/13 $26.00 © 2013 IEEE
DOI 10.1109/GreenCom-iThings-CPSCom.2013.27

distributed way and thus highly scalable. Under the dynamically changing BS topology, DT unambiguously defines the neighboring BSs to cooperate for BS off operation. Furthermore, the proposed overlay network is based on X2 interface of 3GPP LTE, so the architecture of LTE needs not be changed.

2) Second, we propose a *threshold-based* operator-friendly BS off algorithm in solving the NP-hard optimization problem. The utilization of a BS, defined as the busy fraction of time, is the *key* parameter to decide turning-off BS. We will see that utilization is related with user association policy and load balancing. Relying only on the BS utilization to decide the turning-off BS perhaps looks simple, but it works *surprisingly* well because the utilization effectively captures the traffic variation information in space. The threshold-based turning-off operation is also operator-friendly; the choice of lower and upper thresholds determines the operating point on the trade-off curve between energy and delay.

3) Third, extensive simulations confirm that the proposed BS off algorithm can significantly reduce the energy consumption, e.g., more than 60% compared to the baseline. Our simulation is based on a real 3G deployment environment of a specific urban area. The proposed distributed algorithm surprisingly outperforms the previous one that works with the global traffic information [15]. In addition, the simulation results confirm that the proposed algorithm almost achieves the optimal energy/delay trade-off curve.

The remainder of this paper is organized as follows. In Section II, the system model and problem formulation are presented. In Section III, we propose the threshold-based BS off algorithm. In Section IV, the performance of the proposed algorithm is demonstrated with extensive simulation. Finally, we conclude the paper in Section V.

## II. System Model

### A. Assumptions

We consider an infrastructure-based wireless network such as, but are not limited to, 3GPP LTE-Advanced. For the network modeling, a region $\mathcal{L} \subset \mathbb{R}^2$ is served by a set of BSs. At a location $x \in \mathcal{L}$, best effort flows arrive following the Poisson point process. When the average arrival rate of flows at $x$ is $\lambda(x)$ and the average file size is $1/\mu(x)$, the traffic load density is given by

$$\gamma(x) := \frac{\lambda(x)}{\mu(x)}, \tag{1}$$

which captures the spatial traffic variation. When the flow at location $x$ is serviced by BS $i \in \mathcal{B}$, the achievable rate of BS $i$ is

$$c_i(x) = \log_2(1 + SINR_i(x)) \tag{2}$$

where $SINR_i$ is the received signal to interference plus noise ratio of BS $i$. From the definition, $c_i(x)$ is a function of $x$ and determined by $SINR$. We simply consider that the sum of interference experienced by BS $i$ is location-dependent thanks to randomized interference and/or the frequency reuse technique. It should be noted that $c_i(x)$ is not merely determined by the distance from BS $i$, which means it can capture effects of the shadowing by estimating channel gain from BS $i$ to the mobile terminal at location $x$ [13]. The system load density experienced by BS $i$ is defined as

$$\varrho_i(x) := \frac{\gamma(x)}{c_i(x)}. \tag{3}$$

The system load density implies the fraction of time to serve the traffic per unit area at location $x$ with the data rate of $c_i(x)$. Now we define the BS utilization $\rho_i$ as the busy fraction time of BS $i$ to serve the traffic within its coverage,

$$\rho_i = \int_{\mathcal{L}} \varrho_i(x) p_i(x) dx \tag{4}$$

where $p_i(x)$ is user association probability, i.e., the flow at a location $x$ is associated with BS $i$ with probability $p_i(x)$ [13].

### B. Problem formulation

*Definition 2.1 (Feasibility):* Let $\rho = (\rho_1, \cdots, \rho_b)$ be a load vector of all BSs where $b$ is the number of BSs. A set of all BSs in the area $\mathcal{L}$ is denoted by $\mathcal{B}_{all}$. A set of powered-on BSs is denoted by $\mathcal{B}_{on} \subseteq \mathcal{B}_{all}$, and a set of powered-off BSs is denoted by $\mathcal{B}_{off} \subset \mathcal{B}_{all}$. The set $\mathcal{F}(\mathcal{B}_{on})$ is a *feasible set* of utilizations of BSs, which is given by

$$\mathcal{F}(\mathcal{B}_{on}) = \Big\{ \rho \mid \rho_i = \int_{\mathcal{L}} \varrho_i(x) p_i(x) dx, \forall i \in \mathcal{B}_{all}, \tag{5}$$

$$0 \le \rho_i \le 1, \forall i \in \mathcal{B}_{all}, \tag{6}$$

$$\sum_{i \in \mathcal{B}_{all}} p_i(x) = 1, \tag{7}$$

$$0 \le p_i(x) \le 1, \ \forall i \in \mathcal{B}_{all}, \ \forall x \in \mathcal{L}, \tag{8}$$

$$p_i(x) = 0, \ \forall i \in \mathcal{B}_{off} = \mathcal{B}_{all} \backslash \mathcal{B}_{on} \Big\}. \tag{9}$$

Our problem is to find an optimal set of turned-on BSs, $\mathcal{B}_{on}$ considering the trade-off between flow-level performance and the energy consumption. We formulate our problem as a *vector optimization* problem [18], which is given by

$$\min_{\rho \in \mathcal{F}(\mathcal{B}_{on}),\ \mathcal{B}_{on} \subseteq \mathcal{B}_{all}} \{(\Phi_\alpha(\rho, \mathcal{B}_{on}), \Psi(\rho, \mathcal{B}_{on}))\} \tag{10}$$

where $\Phi_\alpha$ is a flow-level performance such as delay, and $\Psi$ is energy consumption.

It can be shown that the feasible set $\mathcal{F}(\mathcal{B}_{on})$ is a convex set, and the objective function (10) is a convex function of $\rho$ given $\mathcal{B}_{on}$. Thus it becomes a convex optimization problem given given $\mathcal{B}_{on}$. However, the choice of $\mathcal{B}_{on}$ essentially makes our problem NP hard.

Our work is based on the framework of [15], but [15] uses the scalarization to find an optimal point considering two competing objectives in (10), which is shown as

$$\min_{\rho \in \mathcal{F}(\mathcal{B}_{on}),\ \mathcal{B}_{on} \subseteq \mathcal{B}_{all}} \Phi_\alpha(\rho, \mathcal{B}_{on}) + \eta \Psi(\rho, \mathcal{B}_{on}) \tag{11}$$

where the parameter $\eta \ge 0$ determines the trade-off between the flow-level performance and the energy consumption. To solve the problem (11), the authors of [15] proposed an algorithm that requires $\gamma(x)$ and $c_i(x)$ over all $x \in \mathcal{L}$, which may be a significant obstacles in implementing the algorithm in the real system; it is hard to know the traffic load density $\gamma(x)$

in practice, and the algorithm in [15] are hardly implementable, which motivates our work.

The cost function of energy consumption $\Phi_\alpha$ is given as

$$\Phi_\alpha(\rho, \mathcal{B}_{on}) = \sum_{i \in \mathcal{B}_{on}} \frac{(1-\rho_i)^{1-\alpha} - 1}{\alpha - 1} \quad (12)$$

where $\alpha \geq 0$ specifies the desired degree of load balancing; when $\alpha = 0$ is rate optimal, $\alpha = 1$ is throughput-optimal, and $\alpha = 2$ is delay-optimal, etc [13]. When $\alpha = 2$, (12) becomes

$$\Phi_2(\rho, \mathcal{B}_{on}) = \sum_{i \in \mathcal{B}_{on}} \frac{\rho_i}{1-\rho_i} \quad (13)$$

which corresponds to minimizing the average delay when the active number of flows changes dynamically, and the flows served by the BS are scheduled in a temporally fair way assuming M/GI/1 multi-class processor sharing model [13].

The energy consumption of BSs $\Psi$ can be modeled in two parts: the fixed energy consumption and the utilization-proportional energy consumption, which is given as

$$\Psi(\rho, \mathcal{B}_{on}) = \sum_{i \in \mathcal{B}_{\text{on}}} \Big[(1-q_i)\rho_i P_i + q_i P_i\Big] \quad (14)$$

where $q_i \in [0, 1]$ is the portion of the fixed power consumption for BS $i$, and $P_i$ is the maximum energy consumption of BS $i$ when it is fully utilized [15]. In the current cellular systems, $q_i$ is very close to 1, which implies the current BS is hardly energy-proportional, and this is because the power consumption of cooling equipments and power amplifier during turned-on period is substantial. For example, in the case of 3G network, $q_i$ is surprisingly around 0.95 [12]. The higher $q_i$ the more beneficial to turn off the underutilized BSs.

We assume that the time scale of turning on/off BSs is much larger than the user association, so first we consider the user association problem, and then we determine the BS turning-off operation. To consider the user association problem, we adopt the distributed $\alpha$-optimal user association [13],

$$\min_{\rho \in \mathcal{F}(\mathcal{B}_{\text{on}})} \Phi_\alpha(\rho, \mathcal{B}_{\text{on}}), \quad (15)$$

which considers the cell load balancing and decides the user association deterministically using the distributed iterations that converge to an optimal point. In other words, each user equipment (UE) makes the user association decision considering the load balancing.

## III. Utilization threshold-based BS off Algorithm

### A. Constructing Overlay Networks Using DT Graphs

For SON operation, BSs need to recognize the network topology change by exchanging messages among BSs if necessary. When BSs are turned off, the set of turned-on BSs, i.e., $\mathcal{B}_{on}$ dynamically changes; thus, the message exchange is needed regularly, so that BSs are aware of the topology change. A general message exchanging procedure is defined in 3GPP LTE standards via X2 interface. Since the architecture of LTE does not assume the centralized controller like the radio network controller (RNC) of the universal mobile telecommunication system (UMTS), the BSs should cooperate themselves and exchange messages via X2 interface for balancing the cell loads or mitigating the inter-cell interference. We exploit the architecture of X2 interface for BS off operation and propose a novel technique using overlay networks based on DT[1]. The proposed technique is well aligned with SON operation and effectively works with the spatial mobile traffic variation.

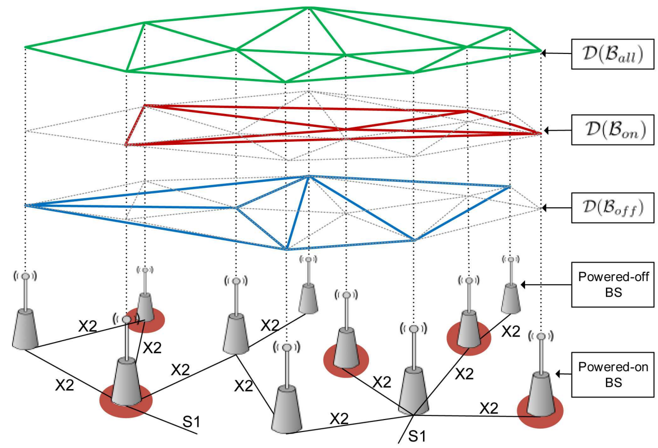



Fig. 1. Overlay networks using DT : $\mathcal{D}(\mathcal{B}_{all})$ with all BSs, $\mathcal{D}(\mathcal{B}_{on})$ with powered-on BSs, and $\mathcal{D}(\mathcal{B}_{off})$ with powered-off BSs.

We construct three different overlay networks, i.e., three DT graphs as shown in Fig. 1. $\mathcal{D}(\mathcal{B}_{all})$ is the DT with all BSs, $\mathcal{D}(\mathcal{B}_{on})$ is for the powered-on BSs, and $\mathcal{D}(\mathcal{B}_{off})$ is for the powered-off BSs, respectively. Due to the space limitation, we focus on the BS off algorithm using $\mathcal{D}(\mathcal{B}_{on})$ in this paper.[2] To construct the overlay networks using DT work based on only local interaction, a set of neighbors should be defined unambiguously. Once a set of neighbors is defined, each BS should be able to dynamically maintain its neighbor cell list (NCL).

*Definition 3.1 (NCL):* NCL of BS $i$ is a list of BSs that are 1-hop away from BS $i$ on $\mathcal{D}(\mathcal{B}_{on})$. Each BS $i$ maintains its NCL.

NCL information is necessary to make the turning-off decision in a cooperate manner with neighboring BSs. For the stable BS operation, when BS $i$ makes the turning-off decision, BS $i$ should consider how to distribute its traffic load into neighboring BSs and how to control a large number of handover messages induced by the UEs in the turning-off BSs. Thanks to the property that the average number of neighbors in DT is six, the size of memory table per BS to keep connectivity remains quite small (six on the average). In addition the limited size of NCL makes the message signaling overhead negligible.

*Definition 3.2 (Cluster):* On $\mathcal{D}(\mathcal{B}_{on})$, a *cluster* of BS $i$ denoted by $\mathcal{C}_i$ is defined as a set of BSs that are within $n$-hop away from BS $i$ including BS $i$.

[1]DT is dual graph of Voronoi diagram and constructed by connecting centers of cells adjacent to each cell of Voronoi diagram.

[2]Our future work will exploit both $\mathcal{D}(\mathcal{B}_{off})$ and $\mathcal{D}(\mathcal{B}_{all})$ for BS on/off algorithms together.

After BSs construct overlay networks, BS $i$ broadcasts its utilization $\rho_i$. To keep the message broadcasting range within its cluster, the message includes a hop counter which is similar to the time to live (TTL) in the IP packet. The hop counter in the message decrements by one when it passes one hop. A concept of the cluster is desired when BS turning-off operation need to occur locally in a self organizing manner. From the definition, BS $i$ is always the center of cluster $\mathcal{C}_i$. Note that each BS's cluster is different. In other words, $\mathcal{C}_i \neq \mathcal{C}_{j \neq i}$ when $j$ is in the NCL of BS $i$.

### B. Algorithm Description

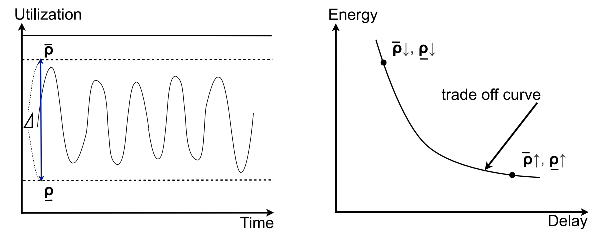
The utilization $\rho_i$ is passed over DT graphs via X2 interface and used for turning-off decision. The utilization $\rho_i$ in (4) is the integration of traffic load density $\varrho_i(x)$ multiplied by the user association probability $p_i(x)$. Thus, the utilization contains the information about spatial traffic distribution. At first glance, it seems hard to implement due to the complexity of the integral formula in (4). However, it is easy to implement in practice; instead of computing the average utilization from (4), BS utilization in OFDM system such as LTE-A can be computed by the average usage rate of resource block the scheduler manages, so it is conveniently measurable [13]. Additionally, we assume that the powered-off BSs can still exchange messages via X2 interface because it requires very little energy consumption.

In the algorithm, we introduce two parameters, *upper threshold* $\bar{\rho} \in [0, 1]$ and *lower threshold* $\underline{\rho} \in [0, 1]$. Intuitively $\bar{\rho} \geq \underline{\rho}$, and these are used for controlling the turning-on and turning-off operations. BSs $i$ could be turned off when $\rho_i \leq \underline{\rho}$, i.e., when the traffic level is low as the traffic level changes like sinusoidal pattern during a day. Similarly, if BS $i$'s utilization becomes $\rho_i \geq \bar{\rho}$, then one of BS $i$'s neighbors need to wake up to unload the BS $i$. This will typically occur when the average traffic level arises as the morning comes while many BSs have been turned off during the night.

Proper setting of two thresholds can prevent so-called *ping-pong effect* that a subset of BSs are repeatedly turned-off and turned-on. For example, when $\Delta = (\bar{\rho} - \underline{\rho})$ becomes large in Fig. 2 (a), ping-pong effect will be mitigated. Mobile operators can properly set up the values, $\bar{\rho}$ and $\underline{\rho}$, to achieve reliable BS off operation considering the energy-delay trade-off, as shown in Fig. 2 (b). When $\underline{\rho}$ and/or $\bar{\rho}$ is large, BS $i$ consumes little energy in average, but it results in higher average flow delay. It is intuitivley correct that more BSs are off when $\underline{\rho}$ is large, and also more powered-off BSs wait to initiate turning-on process until $\rho_i \geq \bar{\rho}$ due to large $\bar{\rho}$.

Now we describe our BS turning off algorithm. Our algorithm can be implemented independent of the size of network because BS off operation is performed locally within the cluster. Note that $\mathcal{B}_{on}$ and $\mathcal{B}_{off}$ evolve dynamically so do the associated DTs as the turning-off operation proceeds.

*1) Threshold-based BS off Algorithm:* We describe the turning-off algorithm that distributedly finds BSs to turn off. Powered-on BSs ($\mathcal{B}_{on}$) construct overlay network, $\mathcal{D}(\mathcal{B}_{on})$. Note that there are many methods of constructing DT in a distributed way, but out of the scope of this paper. If $\rho_i \leq \underline{\rho}$, the BS $i$ broadcasts the `utilization` $\rho_i$, and the neighboring BSs $j$ in the $i$'s cluster $\mathcal{C}_i$ get the `utilization`.




(a) Temporal traffic fluctuation (b) Energy-delay trade-off curve

Fig. 2. The mobile operators can choose the specific operation point on the trade-off curve. $\Delta = \bar{\rho} - \underline{\rho}$ can be used to avoid the ping-pong effect.

TABLE I. MESSAGES FOR BS OFF ALGORITHM

| Message Type | Main Contents |
|---|---|
| `utilization` | $\rho_i$ |
| `off-initiation` | Notification |
| `off-agree` | Confirmation |
| `off-complete` | Notification |

After broadcasting $\rho_i$, BS $i$ can know the lowest utilization in $\mathcal{C}_i$ since other BSs having utilization lower than $\underline{\rho}$ also broadcast the `utilization` periodically. Once BS $i$ knows that it is the lowest utilized BS in $\mathcal{C}_i$, BS $i$ broadcasts the `off-initiation` so that other BSs know who is going to turn off and thus wait until the BS $i$ finishes the turning-off procedure. In our algorithm, we do not allow two BSs in the same cluster are turned off simultaneously for reliable operation. The neighboring BSs $j$ check their utilization condition $\rho_j \leq \bar{\rho}, \forall j \in \mathcal{C}_i, j \neq i$. If BS $j$ still has the spare capacity to accept the handover request from BS $i$, then BS $j$ sends the `off-agree` to the turning off BS $i$. Then BS $i$ broadcasts the `off-complete` and is turned off. BS $i$ is removed from $\mathcal{B}_{on}$, and now neighboring BSs memorize $i \in \mathcal{B}_{off}$. The BS off algorithm is summarized below.

**Threshold-based Distributed BS off Algorithm**

**while** $\rho_i \leq \underline{\rho}$ and $\rho_j \leq \bar{\rho}$ where $\forall j \in \mathcal{C}_i$
    BS $i \in \mathcal{B}_{on}$ broadcasts `utilization` to $\mathcal{C}_i$.
    **if** $\rho_i = \min\{\rho_j\}$, **then**
        BS $i$ broadcasts `off-initiation` to $\mathcal{C}_i$.
        **if** BS $i$ receives `off-agree` from BS $j(\neq i)$, **then**
            {BS $i$ broadcasts `off-complete` to $\mathcal{C}_i$, and is turned off. }
    **else** maintain the current state.
    (Re)construct $\mathcal{D}(\mathcal{B}_{on})$ and $\mathcal{D}(\mathcal{B}_{off})$
**end while**

*Remark 3.1:* It can be shown that the algorithm works without a deadlock when the number of BSs is finite.

## IV. NUMERICAL RESULTS

We first verify the proposed threshold-based BS off algorithm by extensive simulations based on pratical network configuration. We compare the proposed turning-off algorithm with greedy-off algorithm (GOFF) [15] and the exhaustive

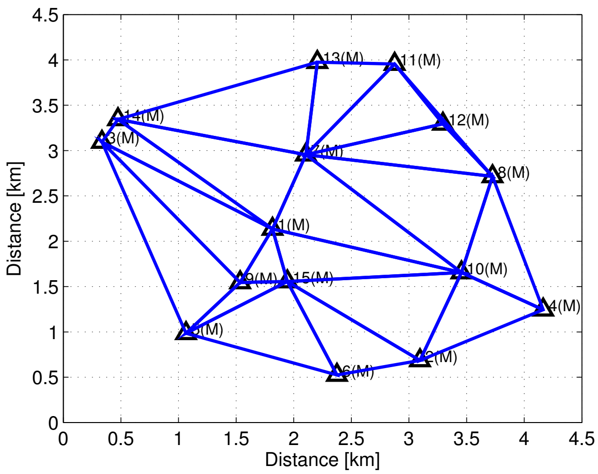



Fig. 3. The BS deployment topology in the urban area (15 BSs)

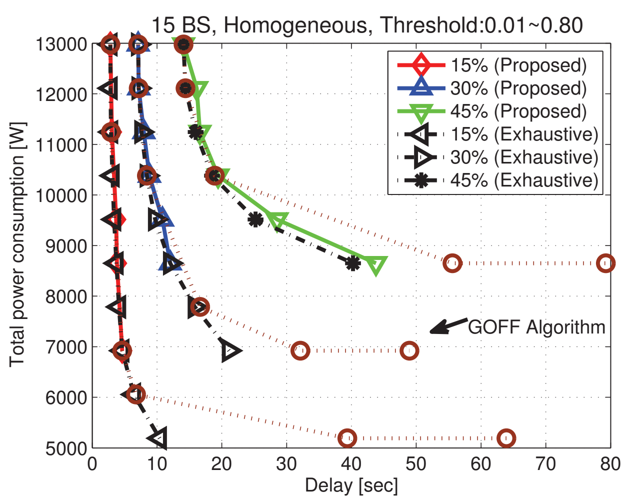



Fig. 4. Energy-delay trade-off for the comparison the threshold-based BS off algorithm, the centralized GOFF algorithm, and the exhaustive search. As $\underline{\rho}$ increases, energy saving can be higher at the cost of higher delay.

search algorithm. Note that GOFF algorithm shows good performance, almost close to the optimal one, which can be found by the exhaustive search only because our problem is originally NP hard. As can be seen in (11), the drawback of GOFF is such that it requires the global knowledge of traffic distribution $\varrho_i(x)$ to calculate the objective functions, which can be hardly known in practice, and GOFF algorithm requires the centralized controller that has global traffic distribution.

For the simulation, we use the BS deployment topology consisting of 15 macro BSs of some 3G operator as shown in Fig. 3. We use the modified 231 path loss model and other parameters specified in IEEE 802.16m evaluation methodology document [19]. For modeling BS power consumption, we assume the non energy-proportional BS that consumes the constant energy when it is on. The transmit power of BS is 43dBm, and the total power consumption of the BS is 865W [15], [20], [21]. Note that applying the proposed algorithm to the energy-proportional BS is also feasible.

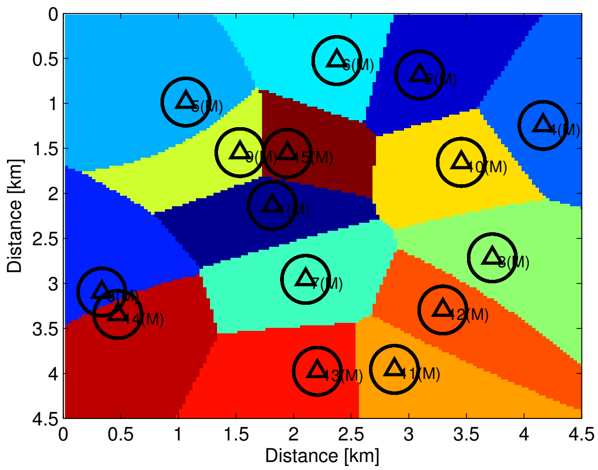
For simplicity, we assume that all 15 BSs are in one cluster, and traffic is distributed homogeneously in space even though our algorithm works for inhomogeneous traffic distribution as well. The results are for three cases of the average utilizations, 15%, 30%, and 45%, respectively. In the real environment, 15% may imply the low utilization during the night while 30% or 45% may imply the increased traffic in the morning and daytime. Our setup is conservative in the sense that the actual average BS utilization can be less than 10% at night [7], so further energy saving is possible depending on the traffic load.

The proposed turning-off algorithm determines the operating point depending on the lower threshold $\underline{\rho}$. In Fig. 4, we sweep the value of $\underline{\rho}$ from 0.01 to 0.80. To focus on the impact of energy saving by turning off the BSs, we set the upper threshold $\bar{\rho}$ as 1. As can be seen, when $\underline{\rho}$ increases, more BS are turned off, and total energy consumption is decreased. However, the delay becomes large. Note that turning off too many BSs induces some BS's utilization to exceed $\bar{\rho}$, i.e., those BSs are overloaded, and we cannot continue turning-off process. Fig. 4 shows that trade-off curves found by the exhaustive search and the proposed algorithm, which are very close. To compare with GOFF algorithm, $\eta$ in (11) varies as 0, $10^{-4}$, $10^{-3}$, $10^{-2}$, $10^{-1}$, and $10^{0}$, respectively. By increasing $\eta$, the energy consumption becomes low, but the delay increases fast, which makes the service unfavorable.

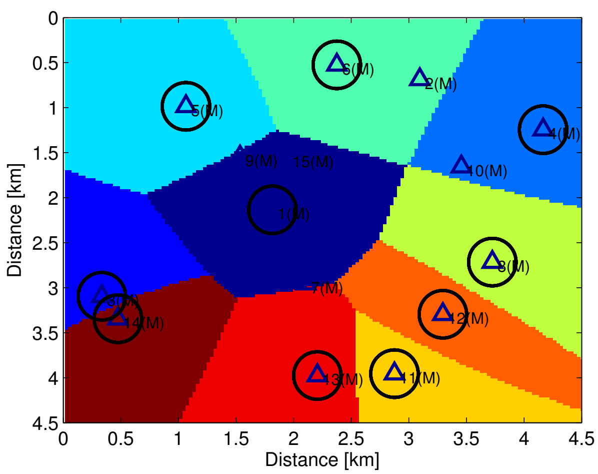
In addition, Fig. 5 shows the coverage of each BS. Fig. 5 (a) is the initial coverage when fifteen BSs are powered on. After running the proposed algorithm, Fig. 5 (b) shows five BSs are turned off. We use $\eta = 0$, $\underline{\rho} = 0.2$, $\alpha = 2$, and the initial average BS utilization is 15%.

## V. Conclusion and Future Work

In this paper we proposed a novel technique for green BS operation. Specifically, our algorithm is intended for the distributed implementation of BS off without knowing the global spatial traffic distribution a priori. Considering the fact that the green BS operation is essentially an NP hard optimization problem and previous approaches are mainly based on the centralized controller, our proposed method is unique. In doing this, we proposed an overlay network based green SON, which takes the advantage of DT. We observe that DT has several nice properties for green SON operation. Specifically, when BSs are dynamically on and off, the active set of BSs accordingly evolves so do the neighboring cell lists. Since DT can be updated in a distributed manner, using DT for green SON operation is very desirable. Furthermore, we only assume the X2 interface between BSs for SON operation and minimal message overhead, the proposed algorithm is compliant with the upcoming standards such as the next version of LTE-A. For the operator friendly operation, we proposed the utilization threshold-based BS off algorithm over DTs. The intuition is that the utilization of BSs implicitly captures the spatial traffic distribution. Extensive simulations show that our proposed method achieves the near-optimal energy-delay performance curve whereas the optimal curve can be found by the exhaustive search only. Furthermore, our




(a) Before running the proposed algorithm (All 15 BSs are turned on.)




(b) After running the proposed algorithm (10 BSs are turned on.)

Fig. 5. The comparison of cell coverage areas. After 5 BSs are turned off, the algorithm is automatically stopped by the constraint ($\rho_i \leq \bar{\rho}$) to sustain good delay performance.

algorithm outperforms the previous one assuming the global traffic distribution and the centralized controller, which shows the practical implication of the proposed method. In the future work, we will propose the BS on algorithm and investigate the impact of massive hand-over initiated possibly by many mobile terminals that belong to the turning-off BSs. Finally, applying our technique for dense small cell BS environments will be of interest too.

## ACKNOWLEDGMENT

This research was supported by LG Electronics Co., Ltd.

## REFERENCES

[1] C. V. N. Index, "Global mobile data traffic forecast update, 2012-2017," *Cisco white paper*, 2013.

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

[3] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access (E-UTRAN); Overall description; Stage 2," 3rd Generation Partnership Project (3GPP), TS 36.300, Mar. 2013. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/36300.htm

[4] "Green touch." [Online]. Available: http://www.greentouch.org/.

[5] "5GrEEN." [Online]. Available: http://www.eitictlabs.eu/news-events/events/article/5green-towards-green-mobile-networks-stockholm/.

[6] E. Oh and B. Krishnamachari, "Energy savings through dynamic base station switching in cellular wireless access networks," in *Proc. of IEEE Globecom*, Miami, FL, Dec. 2010.

[7] E. Oh, K. Son, and B. Krishnamachari, "Dynamic base station switching-on/off strategies for green cellular networks," *IEEE Trans. Wireless Commun.*, 2013.

[8] S. Zhou, J. Gong, Z. Yang, Z. Niu, and P. Yang, "Green mobile access network with dynamic base station energy saving," in *ACM MobiCom*, vol. 9, no. 262, 2009, pp. 10–12.

[9] Z. Niu, Y. Wu, J. Gong, and Z. Yang, "Cell zooming for cost-efficient green cellular networks," *IEEE Comm. Mag.*, vol. 48, no. 11, pp. 74–79, 2010.

[10] Z. Niu, "Tango: traffic-aware network planning and green operation," *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 25–29, 2011.

[11] X. Weng, D. Cao, and Z. Niu, "Energy-efficient cellular network planning under insufficient cell zooming," in *IEEE VTC-Spring '11, Wksp. Greenet*, Budapest, Hungary, May 2011.

[12] C. Peng, S.-B. Lee, S. Lu, H. Luo, and H. Li, "Traffic-driven power saving in operational 3g cellular networks," in *in Proc. of ACM Mobicom, Las Vegas, Nevada, USA, Sep. 2011*, pp. 121–132.

[13] H. Kim, G. de Veciana, X. Yang, and M. Venkatachalam, "Distributed alpha-optimal user association and cell load balancing in wireless networks," *IEEE/ACM Trans. Networking.*, vol. 20, pp. 177–190, Feb. 2012.

[14] H. Kim and G. de Veciana, "Leveraging dynamic spare capacity in wireless systems to conserve mobile terminals energy," *IEEE/ACM Trans. Networking.*, vol. 18, no. 3, pp. 802–815, Jun. 2010.

[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 Jour. Select. Areas in Comm.*, vol. 29, Sep. 2010.

[16] S. Borst, "User-level performance of channel-aware scheduling algorithms in wireless data networks," in *Proc. IEEE INFOCOM*, Apr. 2003, pp. 321–331.

[17] H. Kim, C.-B. Chae, G. de Veciana, and R. W. Heath Jr., "A cross-layer approach to energy efficiency for adaptive mimo systems exploiting spare capacity," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 2745–53, Aug. 2009.

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

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

[20] A. J. Fehske, F. Richter, and G. P. Fettweis, "Energy efficiency improvements through micro sites in cellular mobile radio networks," in *Proc. of the second International Workshop on Green Communications (GreenComm)*, Honolulu, HI, Dec. 2009.

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