Green Small Cell Operation Using Belief Propagation in Wireless Networks

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

Abstract—In this paper we study the energy efficiency maximization problem focused on finding a suitable set of turned-on small cell access points (APs). Since APs in small cells are randomly deployed and thus redundant in many cases, a mechanism of turning on/off APs is needed. We propose a device-assisted framework that exploits feedback messages from the user equipment (UE). To solve the problem, we apply the optimization method using belief propagation (BP) on a factor graph. Furthermore we propose an online algorithm inspired by BP, called DANCE that requires low computational complexity. Extensive simulations confirm that BP enhances energy efficiency significantly. Furthermore, simple but practical DANCE exhibits close performance to BP, and also better performance than other existing methods and the baseline.

Index Terms—Cellular networks, small cell, energy efficiency, belief propagation, optimization.

I. INTRODUCTION

To cope with the mobile traffic explosion expected to reach 11.2 exabytes by 2017 [1], upcoming 5G focuses on the gigabit-scale peak data rate with the higher capacity. Deploying the hyper-dense small cell network is a strong candidate to enhance the capacity up to 100 times [2]. While denser small cells can meet the ends on the capacity requirement, the total power consumption on the network increases proportional to the number of small cell base stations (BS) [3]. In addition we should consider that network infrastructures takes 3% of the world’s electrical energy consumption and BSs consume about 80% of the energy consumption of a mobile network operators (MNOs) [4], [5]. In light of these illustrations, it is indispensable to increase energy efficiency. Thus many efforts are continued in standards of 3GPP and research projects of EARTH, Green Touch, 5GrEEn, etc. [6]. Also in perspective toward 5G, the authors of [7] recently suggest a design framework considering both energy efficiency (EE) and spectral efficiency (SE), called EE-SE co-design.

In literature, a lot of on-going research investigates green BS operations to enhance energy efficiency. In [8], the authors proposed a BS turning-off algorithm using the optimal cell coverage that minimizes BS area power consumption. The authors in [9] proposed cell zooming technique such that the BS turning-off operation reduces the power consumption and other turned-on BSs increase their coverage. In [10]–[12], the on/off methods based on flow-level dynamics are presented. The authors in [10] solved the problem encompassing dynamic BS operation and the related problem of user association. Related to [10], the authors in [11] presented the distributed threshold-based BS off algorithm based on an overlay network using Delaunay triangulation. Also the authors in [12] proposed the energy-efficient user association method by using a population game approach. In [13]–[15], stochastic geometry is used. The authors in [13] proposed macro BS turning-off strategies in homogeneous and heterogeneous networks and showed that the deployment of small cells can increase energy efficiency but the energy efficiency gain is saturated when the small cell density is increased. In two-tier heterogeneous network, the authors in [14] proposed the repulsive cell activation scheme that activates small cells according to user density and offloading macro BS traffic. A deployment strategy considering the density and transmission power of BSs are proposed in [15]. The authors in [16] considered a separation architecture substituting the conventional macro BS into a coverage providing BS and multiple small cells in order to increase energy efficiency. The authors derived the optimal intensity of small cells for the optimal deployment.

In this paper, we study the small cell BS on/off operation to enhance energy efficiency. Our approach is not limited to the type of small cells, e.g., pico BS, femto BS, or even Wi-Fi BS. Hereafter we use the terminology access point (AP) instead of BS to emphasize the small cell BS. Since we assume that users can directly deploy small cell APs, intractable randomness is engendered. Thus devices’ channel measurement is vital to the AP on/off operation because the type of small cells can vary as well. These assumptions motivate us to take a new approach for the AP on/off operation. Our approach exploits belief propagation (BP) to solve an optimization problem on a factor graph.

The main contributions of this paper are summarized as follows. We apply the BP optimization framework into the AP on/off operation. To tackle the randomness of AP deployment, a model exploiting user equipment (UE) feedbacks is proposed. Then by constructing a factor graph representing APs and UEs, the optimization problem is solved with the BP optimization framework. Since BP is an offline algorithm, we propose an online algorithm called DANCE (Device Assisted Networking for Cellular grEening) that is a collection of low complexity...
algorithms inspired by BP. Our extensive simulation results show that BP and DANCE increase energy efficiency significantly, e.g. up to 129%, compared to the baseline where all APs are always on.

The remainder of this paper is organized as follows. In Section II, assumptions under our system model and problem formulation are presented. Section III is devoted to the graphical modeling of our framework and the BP algorithm. In Section IV, we propose online algorithms. In Section V, the performance by using our algorithms compared to the baseline. Mitigation of turning off APs can also increase the data rate when APs of service that is mainly determined by data rate. Interestingly, solution when APs are densely deployed. However, while saving energy, it is important to maintain the acceptable level of quality of service that is mainly determined by data rate. Interestingly, turning off APs can also increase the data rate when APs are densely deployed because interference among APs can be mitigated depending on the set of turned-off APs. In Section VI, we show that higher or similar average data rate is achieved by using our algorithms compared to the baseline.

We assume that the set of APs $\mathcal{J}$ are located in the macro base station coverage, and the set of UE $\mathcal{I}$ is in the coverage provided by APs where $j \in \mathcal{J}$ is the index of AP and $i \in \mathcal{I}$ is the index of UE.

**Definition 1 (AP On/Off State Indication Vector):** When $x_j$ stands for the turned-on or off state of AP $j \in \mathcal{J}$, the AP on/off state indication vector $x = \{x_1, \cdots, x_j, \cdots, x_J\}$ is defined by

$$x_j = \begin{cases} 1, & \text{if AP } j \text{ is turned on,} \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where $J$ is the number of APs in the set $\mathcal{J}$. Note that the AP state indication vector shows a possible state of each AP. Once the SCC determines the set of turned-on APs, the vector indicates whether each AP is turned on or not. Suppose that if an AP is turned on, it consumes the operational power $P_{j,Op}$ that includes the transmission power $P_{j,Tx}$ dissipated into the air. Thus the operational power vector of AP $j \in \mathcal{J}$ is $P_{Op}(x) = \{P_{1,Op}(x_1), \cdots, P_{j,Op}(x_j), \cdots, P_{J,Op}(x_J)\}$ where

$$P_{j,Op}(x_j) = \begin{cases} P_{j,Op}, & \text{if } x_j = 1, \\ 0, & \text{otherwise,} \end{cases}$$

and the transmission power vector of AP $j$ becomes $P_{Tx}(x) = \{P_{1,Tx}(x_1), \cdots, P_{j,Tx}(x_j), \cdots, P_{J,Tx}(x_J)\}$ where

$$P_{j,Tx}(x_j) = \begin{cases} P_{j,Tx}, & \text{if } x_j = 1, \\ 0, & \text{otherwise.} \end{cases}$$

We define the association matrix $S = [s_{i,j}] \in \mathbb{R}^{I \times J}$ that shows user association between APs and UEs given by

$$s_{i,j} = \begin{cases} 1, & \text{if UE } i \text{ is connected to AP } j, \\ 0, & \text{otherwise.} \end{cases}$$

**B. Problem Formulation**

When UE $i$ is associated with AP $j$, we have the signal-to-interference ratio (SINR) between UE $i$ and AP $j$ given by

$$SINR_{i,j}(x) = \frac{g_{i,j} P_{j,Tx}(x_j)}{N_0 + \sum_{j' \in \mathcal{J}\setminus\{j\}} |g_{i,j'} P_{j',Tx}(x_{j'})|}$$

where $g_{i,j}$ is the channel gain between UE $i$ and AP $j$, and $N_0$ is the noise power. Then the maximum spectral efficiency of UE $i$ is $log_2(1 + SINR_{i,j}(x))$. We assume that when an AP $j$ has multiple associated UEs, the associated UEs are scheduled as temporally fairly or equally when sharing the frequency channel; thus, the data rate of UE $i$ is given by

$$R_i(x) = \log_2 \left(1 + SINR_{i,j}(x) \right) / S_j \quad (2)$$

where $S_j = \sum_{i=1}^{I} s_{i,j}$. Energy consumption of APs can be calculated by summing the operational power of turned-on APs. Thus the total power consumption of the network becomes $\sum_{j=1}^{J} (P_{j,Op} \cdot 1(S_j))$ where an indicator function is

$$1(S_j) := \begin{cases} 1 & \text{if } S_j > 0, \\ 0 & \text{if } S_j = 0. \end{cases}$$

Our objective function is the energy efficiency $F(x)$ also
called as the global function given by
\[ F(x) = \sum_{i=1}^{I} f_i(x). \]

Our problem is to find the on/off state vector \( x \) that maximizes
the global function \( F(x) \). Thus, our optimization problem is
\[
\max_{x,S} \quad F(x) \tag{3}
\]
\[ \text{s.t} \quad x_j \in \{0,1\}, \; \forall j \in J \tag{4} \]
\[ \sum_{i=1}^{I} \sum_{j \in \{i|x_i=0\}} s_{i,j} = 0, \tag{5} \]
\[ \sum_{j=1}^{J} s_{i,j} = 1, \; \forall i \in I. \tag{6} \]

The first constraint (4) is from determining a set of turned-on APs. The second constraint (5) means that turned-off APs should not be associated with any UEs. The third constraint (6) shows that all UEs should be associated with only one of the turned-on APs. To tackle the the problem (3), we determine the set of turned-on APs. Thus, it is known that the problem is NP-hard implying that it is hard to find the optimal solution of (3) in a polynomial time; the exhaustive search should be done. In the next section we apply BP to our framework and find a solution that is close to the optimal performance.

III. BELIEF PROPAGATION

We adapt BP to solve our optimization problem expecting that a simple but practical algorithm can provide an approximated solution of our problem. BP originally proposed by Pearl [18] is a message-passing algorithm widely used in the inference problems. BP has successfully been demonstrated in many applications such as error-correcting codes or artificial intelligence. By using BP, the marginals of a joint probability distribution can be calculated with iterations of message passing on graphical models including Bayesian networks and Markov random fields. We specifically focus on running BP on a factor graph, which also can be readily converted into other forms of graphical models like Bayesian networks. For details, please refer to [19] that shows the general overview of BP, and [20–22] that provide explanations about BP on a factor graph. The optimization method using BP was used by the authors of [23]. The authors applied BP to intercell interference coordination in femtocell network to solve the problems approximately.

By using the elements of the AP on/off state indication vector, we define a set of discrete random variables \( X = \{X_1, \cdots, X_j, \cdots, X_J\} \). Then, \( X_j, \forall j \in J \), can be seen as a random variable denoting the turned-on or turned-off state. The realizations of the random variable \( X_j \) is denoted by \( x_j \in \{0,1\}, \forall j \in J \), which is consistent with the definition (1). Let us use the abbreviated notation \( x = \{x_1, \cdots, x_j, \cdots, x_J\} \) denoting \( \{X_1 = x_1, \cdots, X_j = x_j, \cdots, X_J = x_J\} \). By using these definitions above, the joint probability distribution \( p(X_1 = x_1, \cdots, X_j = x_j, \cdots, X_J = x_J) \) is expressed in the shorter form \( p(x) \).

To define the joint probability distribution \( p(x) \), the concept of Boltzmann’s law in thermal equilibrium from statistical mechanics is used. Our framework can be seen as a system with the assumption that the system has \( J \) particles, and a state of each particle is denoted by \( x_j \). By substituting the negative of free energy part of Boltzmann’s law into our global function, \( F(x) \) is defined as
\[ p(x) = \frac{1}{z} \exp \left( \frac{1}{\tau} F(x) \right) \tag{7} \]
\[ = \frac{1}{z} \exp \left( \frac{1}{\tau} \sum_{i=1}^{I} f_i(x) \right) \tag{8} \]
where \( z \) is a normalization constant and \( \tau \) denotes the temperature. When BP is applied to the non-physical system, the temperature can be seen as a constant parameter appropriately chosen.

A. Factor Graph

Observing (7) and (8), we see that (7) including the global function \( F(x) \) is factorized into the product of the local functions \( f_i(x) \). Thus (8) can be shown in a factor graph. A factor graph \( G = (V,E) \) is a bipartite graph with vertices \( V \) and edges \( E \). Vertices, i.e., nodes of the graph, correspond to APs and UEs. There are two kinds of nodes: the variable nodes and the factor nodes. The variable nodes representing APs, shown as circle in the graph, are referred to \( X_j, j \in J \). The factor nodes denoting UEs, shown as square in a graph, are referred to \( \exp \left( \frac{1}{\tau} f_i(x) \right) \) where the local function is included. Edges mean the channel between APs and UEs. To solve our optimization problem, we exploit the constructed factor graph in BP message passing.

B. Belief Propagation Optimization Algorithm

Following the results of [24], when \( \frac{1}{\tau} \rightarrow \infty, p(x) \) concentrates around the maxima of \( F(x) \), and then we have
\[ \lim_{\frac{1}{\tau} \rightarrow \infty} \hat{x} = \arg\max_{x} F(x) \tag{9} \]
where \( \hat{x} = \{\hat{x}_1, \cdots, \hat{x}_j, \cdots, \hat{x}_J\} \), and \( \hat{x}_j = E[X_j] \) means the marginal expectation of the random variable \( X_j \). In principal, standard BP can be seen as a process to calculate the estimated marginal probability distribution with respect to the random variable \( X_j \). Thus, when the marginal probability distribution of \( p(x) \) is estimated, an approximated solution of the optimization problem can be found.

Definition 2 (Belief): The belief \( b_j(x_j) \) is the estimated marginal probability distribution of \( p(x) \) with respect to \( X_j \).

Thus we leverage the BP algorithm to find \( b_j(x_j) \). On the constructed factor graph, belief messages are passed through the edges between APs and UEs. After the belief message passing is repeated between APs and UEs, the set of turned-on APs is determined according to the estimated marginal probability distribution. It is well known that when a factor graph has no cycle, \( b_j(x_j) \) converges to the true marginal probability distribution of \( p(x) \) with respect to \( X_j \) [23]. However, as shown in the Fig. 2, the constructed factor graph usually has cycles, e.g., AP 2 - UE 3 - AP j - UE i - AP 2. In this case, BP yields an approximated result; thus, we compare our simulation results with the optimal value obtained from the exhaustive search in Section V.

Now we present the algorithm below.

1) Initialization. At time \( t = 0 \), messages \( b_{j \to i}(t, x_j) \) for \( \forall j \in J \) and \( \forall i \in I \) are set to have an arbitrary distribution. For instance, we use the uniform distribution such as

\[
b_{j \to i}(t, x_j) \sim U(0, 1).
\] (10)

2) UE update. After the initialization, UE \( i \in I \) generates and sends the belief messages to its neighboring APs \( j \in N(i) \) where \( N(i) \) denotes the neighbor APs of UE \( i \).

At \( t > 0 \), the message update is defined by

\[
b_{i \to j}(t, x_j) = E\left[\frac{1}{x_j} f_i(X_j)\right]|x_j]
\] (11)

\[
= \sum_{x \in \{X_j\}} \left(\exp\left(\frac{1}{\eta} f_i(x)\right)\right) \times \prod_{j' \in N(i) \setminus \{j\}} b_{j' \to i}(t, x_{j'})
\] (12)

where for given \( X_j = x_j \), the expectation (11) is calculated over independent \( x_{j'} \), \( \forall j' \in N(i) \). Also in (12) \( \exp\left(\frac{1}{\eta} f_i(x)\right) \) is calculated for given \( X_j = x_j \). Since (12) has both sum and product, BP is called the sum-product algorithm. Equation (12) is the product of the factor with belief messages from AP \( j' \in N(i) \), \( j' \neq j \), to UE \( i \), marginalized over all random variables except \( x_j \). To calculate \( f_i(x) \), at \( t = 1 \) we assume that \( S_j = 1, \forall j \in J \).

At \( t > 2 \), \( S_j, \forall j \in J \), is initialized to 1 and incremented by the number of associated UEs where each UE is associated with AP \( j \) is \( \text{argmax}_{x_j} b_{j \to i}(t, x_{j'}) = 1 \).

3) AP update. AP \( j \in J \) generates and sends the belief messages to all neighboring UE \( i \in N(j) \) where \( N(j) \) denotes the neighbor UEs of AP \( j \). The message is defined by

\[
b_{j \to i}(t + 1, x_j) = \frac{1}{z'} \prod_{i' \in N(j) \setminus \{i\}} b_{i' \to j}(t, x_j)
\] (13)

where \( z' \) is the normalizing constant making \( \sum_{x_j \in \{0, 1\}} b_{j \to i}(t + 1, x_j) = 1 \). The belief message (13) is the product of the belief messages from UE \( i' \in N(j) \setminus \{i\} \), all neighboring node except \( i \). Then, the steps ‘UE update’ and ‘AP update’ are repeated until the number of iteration is satisfied.

4) Decision-making. After finishing the iteration procedure, The estimated marginal probability distribution with respect to \( X_j \) is finalized as

\[
b_{j}(x_j) = \frac{1}{z''} \prod_{i \in N(j)} b_{i \to j}(t, x_j)
\] (14)

where \( z'' \) is the normalizing constant to make \( b_j(x_j) \) as the probabilistic distribution. To describe our association rule, we define an integer-valued function \( \Phi_j(\cdot) \) that determines the position of \( j \)-th element when the elements of the input vector are sorted in an increasing order. Then the association rule of UE \( i \) is given by

\[
j^*(i) = \text{argmax}_j [\Phi_j(\{\text{SINR}_{i,j}(x)\}) + \eta \Phi(\{p_j(x_j)\})]
\]

where \( x_j = 1, \forall j \in J \), and \( \eta \) is a weighting parameter considering the channel condition and the estimated marginal probability from BP. For example if \( \eta = 0 \), each UE is associated with the AP that provides the highest SINR among APs in \( N(i) \). If \( \eta \to \infty \), UE is associated with the AP \( j \) whose \( p_j(x_j) \) is the greatest. It could be possible that SCC determines \( \eta \) to maximize the energy efficiency by using adaptive learning methods. After user association is done for all UEs, the APs with no UEs are turned off.

While BP framework is readily well-defined on many graphical models, the drawback of BP framework is complexity that is \( O(2^{N(i)}) \) in view point of the UE. Hence, the complexity of BP could be the barrier to scale up to implement for a huge number of APs and/or UEs. Thus we propose online algorithms for dense cellular network in the next section.

IV. ONLINE ALGORITHM: DANCE

We propose online algorithms that are motivated by the BP framework. In BP the node having high certainty of belief can influence the final estimated marginal probability distribution. To mimic this behavior, DANCE considers priority to the AP or UE whose impact can be great. Since the ultra-dense network (UDN) environment is considered where APs are densely deployed in UDN, DANCE is designed to require lower computational complexity.

As mentioned in Section II, UEs measure the APs’ periodic beacon signal and send feedbacks to SCC. We assume that UEs can send feedback messages as shown in Fig. 1. The feedback message can contain only the connectivity information if using only 1 bit, or can have more detailed information such as the maximum achievable data rate or appropriate adaptive
modulation and coding (AMC) using more bits. It is assumed that connectivity exists if $SINR_{i,j}(x)$ is greater than some threshold. In the SCC, the feedback messages form the feedback matrix $A = [a_{i,j}]_{I \times J}$.

DANCE solves the same problem of (3). By using matrix-based algorithms, DANCE determines the turned-off APs and the user association together. Once the target AP is chosen, and then the association is automatically decided on the matrix $A$.

DANCE includes into 3 algorithms: AP First, UE First, and Proximity-ON (Prox-ON) algorithm.

Two algorithms, AP First and UE First, are designed to choose the most influential AP or UE. Depending on the type of feedback, both AP First and UE First have several variations: AP-N/AP-1 and UE-N/UE-1 where the number implies the length of the feedback bits. Note that 1 bit feedback only implies the connectivity possibility. Following the categorization, AP First and UE First are described in Algorithm 1 and 2, respectively.

**Algorithm 1: AP First**

1: Initialize $s_{i,j} = 0, \forall i \in I, \forall j \in J$.
2: while $A \neq \emptyset$
3: 
4: 
5: 
6: else $s_{i,j} = 0$.
7: end for
8: Eliminate the column $j^*$ of $A$.
9: Eliminate $i'$ rows of $A$ where $i' \in \{i|a_{i,j^*} > 0\}$.
10: end while

**Algorithm 2: UE First**

1: Initialize $s_{i,j} = 0, \forall i \in I, \forall j \in J$.
2: while $A \neq \emptyset$
3: 
4: 
5: 
6: else $s_{i,j} = 0$.
7: end for
8: Eliminate the column $j^*$ of $A$.
9: Eliminate $i'$ rows of $A$ where $i' \in \{i|a_{i,j^*} \neq 0\}$.
10: end while

As a counterpart, **UE First** algorithms, UE-1 and UE-N, consider UEs foremost. For example, in UE-1 algorithm we firstly find the UE having the smallest number of association-possible APs. It is because that UE could be possibly the worst case UE. Then among the APs that can connect to the chosen UE, we select the AP $j^* = \arg \max_j \sum_i a_{i,j}$ where $\forall j \in \{j'|a_{i,j'} > 0\}$ in order to minimize the number of turned-on APs. Note that UE First is a generalization of [25] where only a single bit feedback was assumed.

The other is Proximity-ON algorithm. In this algorithm, each UE associates with the AP that provides the greatest SINR where non-selected APs is turned off.

**V. Numerical Results**

We verify the proposed optimization algorithms through extensive simulations. We assume that 9 APs and 20 UEs are randomly distributed in a hexagonal cell area with 100 m radius. From Alcatel-Lucent device (model 9361 home cell v2), the transmission power of the AP is 13 dBm, and the operational power of the AP is 9 W. According to IEEE 802.16m evaluation methodology document (EMD), we use Winner model assuming non-line-of-sight (NLOS) case and propagation through light walls with 2.1 GHz carrier frequency. The total bandwidth is 20 MHz, and the power spectral density of thermal noise is $-174$ dBm/Hz.

Fig. 3 shows the average energy efficiency of 4,000 UEs obtained through 200 iterations of 20 UE case, and results are summarized in Table I. In Fig. 4, we compare the average power consumption and the average data rate obtained from all iterations. For every iteration of the proposed BP algorithm, UE update and AP update is repeated 5 times where $\frac{r}{\eta} = 1$ and $\eta = 5.5$. The baseline is the case when all APs are turned on. Using the exhaustive search, we find the case that maximizes energy efficiency where every UE is associated with the AP providing the highest SINR for the every candidate set of turned-on APs.

First, we compare the average energy efficiency among BP, DANCE, and the baseline in Fig. 3. Table I represents the
improvement of the average energy efficiency. Results show that BP and DANCE increase energy efficiency significantly. For example, BP enhances energy efficiency 129% compared to the baseline. To compare BP and DANCE, BP is better than DANCE. However, the gap is not much, and if one considers that DANCE is a low complexity algorithm, the improvement obtained by DANCE is noticeable; for example, in the case of UE-N, it is 102%. In addition, our series of algorithms, BP, AP-1, AP-N, UE-1, and UE-N outperform Proximity-ON and the baseline.

Fig. 4 shows the average of power consumption and data rate. We observe that BP and DANCE consume less than the average power consumption of the baseline. As trade-off of BP’s significant power reduction, it achieves the lower average data rate. DANCE however achieves the similar or even higher average data rate because the interference is mitigated by the AP off operation and the set of turned-on APs are well chosen to provide high data rate.

VI. CONCLUSION

In this paper, we studied the mechanism to enhance energy efficiency in small cell environment. We formulated the energy efficiency optimization problem and solved the problem by BP. To apply BP, we constructed a factor graph and BP is used to derive the estimated marginal probability. Then according to our association rule, the set of turned-on APs are determined. Whereas BP is an offline algorithm, we also proposed its online version called DANCE, which is inspired by BP. Extensive simulations confirmed that the proposed BP algorithm significantly enhances energy efficiency, and DANCE requiring low complexity not only achieves close performance to BP, but also increases the average data rate with the lower average power consumption.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT and Future Planning (NO. NRF-2014R1A1A1006551).

REFERENCES


