Energy-Saving Oriented On/Off Strategies in Heterogeneous Networks : an Asynchronous Approach with Dynamic Traffic Variations

Lun Tang, Weili Wang and Qianbin Chen*
Key Laboratory of Mobile Communication,
Chongqing University of Posts and Telecommunications
Chongqing 400065, China
[e-mail: tangl@cqupt.edu.cn, 1961797154@qq.com, cqb@cqupt.edu.cn]
*Corresponding author: Qianbin Chen

Abstract

Recent works have validated the possibility of reducing the energy consumption in wireless heterogeneous networks, achieved by switching on/off some base stations (BSs) dynamically. In this paper, to realize energy conservation, the discrete time Markov Decision Process (DTMDP) is developed to match up the BS switching operations with the traffic load variations. Then, an asynchronous decision-making algorithm, which is based on the Bellman equation and the on/off priorities of the BSs, is firstly put forward and proved to be optimal in this paper. Through reducing the state and action space during one decision, the proposed asynchronous algorithm can avoid the "curse of dimensionality" occurred in DTMDP frequently. Finally, numerical simulations are conducted to validate the effectiveness and advantages of the proposed asynchronous on/off strategies.

Keywords: Wireless heterogeneous networks, energy conservation, traffic load variations, DTMDP, asynchronous on/off strategies

---

1. Introduction

1.1 Motivation and Related Work

For the environmental and economic benefits, saving energy has been the shared goal among every walk of life. With the explosive popularity of smartphones and tablets, the information and communication technology (ICT) has accounted for 2%~10% of the world's total energy consumption and the proportion is growing year by year [1-3]. Therefore, some energy efficiency-based algorithms have been designed recently for different network environments to ease the current situation.

In [4], to improve the throughput and energy efficiency of the system simultaneously, a Cat swarm optimization (CSO) was used to realize the energy harvesting and spectrum sharing in 5G networks. In [5], a quantum fireworks algorithm (QFA) was designed to acquire high energy efficiency in energy harvesting cognitive radio. An efficient iterative algorithm based on stochastic geometry was presented in [6] for optimal energy efficiency in wireless powered communication networks.

Notably, the wireless heterogeneous networks have been introduced as the promising solutions to satisfy the requirements of user equipments (UEs) while reducing the energy consumption of networks through deploying the high-capacity and low-cost small BSs (SBSs) over macro BSs (MBSs) [7-9].

The SBSs in the heterogeneous networks are usually deployed to satisfy the traffic load in peak period irrespective of its heavily dynamic variations [10,11]. Therefore, when the networks are lightly loaded, turning off the underutilized SBSs has huge potential to decrease the energy consumption. Through defining the indicators of on-state SBSs as 1,and off states as 0, different 0-1 programming models have been formulated to achieve energy conservation when the SBSs are underutilized [12-14].

In [12], a distributed switching-on/off based energy saving algorithm was proposed based on the additional load increments brought to the neighboring BSs by the switching off ones. In [13], two greedy algorithms were proposed to minimize the number of active BSs under the limitation of the minimal rate requirements of UEs. In [14], the lower limit of load was deduced for the switching on/off mechanism to reduce the energy consumption while guaranteeing the service performance of the heterogeneous networks. However, for the stochastic nature of traffic arriving process in practical networks, the traffic distribution after switching off some BSs is uncertain, and the traffic variation near the off-state BSs would increase the risk of unstable networks. Based on these, the DTMDP is an effective means to decide the on/off states of BSs based on the dynamics of traffic load.

Different DTMDP-based algorithms have been designed to achieve energy conservation in the heterogeneous networks [15-17]. In [15], the optimal sleep/wakeup schemes were derived based on the information of user localization and traffic load. In [16], a transfer actor-critic learning algorithm was adopted to obtain the BS switching operations. However, the two works always face the challenging issue of "curse of dimensionality" when the size of state space expands. In [17], a compact state representation was developed to reduce the state space for avoiding the high computational complexity.

1.2 Contributions and Organizations

Inspired by the work [17], a novel asynchronous on/off strategy is proposed to save energy while overcoming the "curse of dimensionality ". The mechanism proposed in the paper can be applied in the dense heterogeneous networks with dynamic traffic variations.

The key contributions of the paper can be summarized as follows:

• Firstly, the UEs in the networks are classified according to their associated BSs, and when an underutilized SBS is switched off, its UEs will be transferred to MBS for it is always in the active state to ensure coverage. To reduce the energy consumption during the whole period, the minimization model of the long-term average energy consumption under the constraint of the blocking probability is formulated in the paper.

• Secondly, through defining the UE numbers of all BSs as the states and the switching operations of all SBSs as the actions, the load-based on/off strategies have been mapped into DTMDP. The on/off priorities of the SBSs are arranged according to their UE numbers. Based on the resulting priorities and Bellman equation, an asynchronous solving algorithm is finally designed in the paper to obtain the optimal BS switching operations.

• Finally, the paper evaluates the performance of the proposed asynchronous on/off strategies through extensive simulations, and compares it with four other approaches, which are the conventional network operation, the heuristic sleep approach proposed in [9], a similar asynchronous on/off strategy to the proposed one but without priorities and the general synchronous on/off decisions based on DTMDP. The results demonstrate that the proposed algorithm can significantly reduce the energy consumption of the networks under the limitation of blocking probability.

The remainder of this paper is organized as follows. The system model is introduced in Section 2 and the problem formulation is presented in Section 3. In Section 4, the constrained energy-saving goal is mapped into DTMDP and the optimality of the proposed asynchronous on/off strategies is proved. Section 5 provides the numerical analysis and the conclusions are drawn in Section 6.

2. System Model

2.1 Two-tier Heterogeneous Networks

As shown in **Fig. 1**, a typical scenario of two-tier heterogeneous networks, whose topology consists of a single MBS underlaid with a set of SBSs $\mathcal{K}$, is considered in the paper, and $K = |\mathcal{K}|$ denotes the number of SBSs. All BSs in the scenario are numbered by $0, 1, ..., K$, where 0 represents the MBS. The set of all UEs in the scenario is composed of a finite set of classes, where the class 0 is a cluster of UEs who can only associate with the BS 0 (MBS), and a class k ($k \in \mathcal{K}$) is a population of UEs who can not only associate with the MBS but the SBS k. Assume that the UEs of class k ($k \in \mathcal{K}$) only associate with the SBS k when it is in active state. Therefore, the UEs of class k ($k \in 0, 1, ..., K$) and the UEs of BS k are interchangeably used hereinafter. Furthermore, when an underutilized SBS k is switched off to save energy, the UEs of class k will be adjusted to that of class 0 for the MBS will be always in the active state to ensure coverage.

A discrete time model is adopted in the paper, where the time domain is segmented into the same-length time slots of duration T. The term "time slot n" is defined as the time interval $nT \leq t < (n+1)T$, which is equivalent to the term "time t" in the paper.

The evolution of network states across time epochs $t = 1, 2, ...$ is described by a stochastic process $\mathbf{x}(t) \subset X$, where X denotes the state space, then $\mathbf{x}(t)$ is defined as:

$$\mathbf{x}(t) = \left(x_0(t), x_1(t), ..., x_K(t) \right) \tag{1}$$

where each element means the number of UEs of class k during time t. By knowing $\mathbf{x}(t)$, a SBS on/off decision is selected for time $t+1$. A decision set performed in each time t is given by:

$$\mathbf{y}(t) = (y_1(t), y_2(t)..., y_K(t)) \tag{2}$$

where $y_k(t)$ represents the working mode of SBS k, with $y_k(t) = 0$ if SBS k is switched off and $y_k(t) = 1$ otherwise. Let Y denote the action space, then $\mathbf{y}(t) \subset Y$.
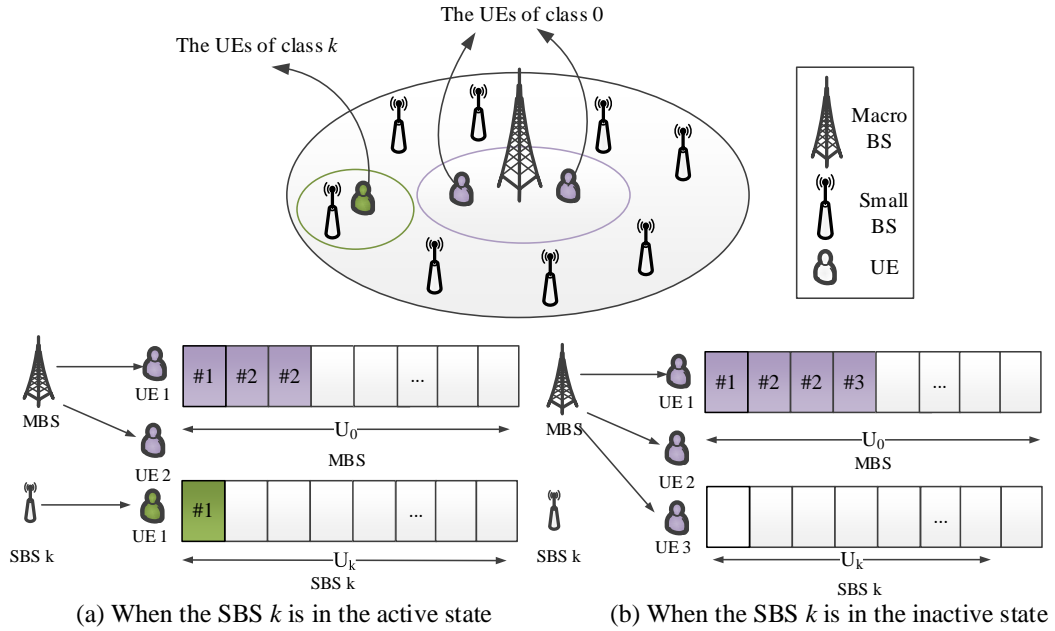

(a) When the SBS k is in the active state   (b) When the SBS k is in the inactive state
Fig. 1. The schematic diagram for BS switching operations.

2.2 Traffic Model

Let $\boldsymbol{\lambda}(t)=(\lambda_0(t),\lambda_1(t),...,\lambda_K(t))$ and $\boldsymbol{\mu}(t)=(\mu_0(t),\mu_1(t),...,\mu_K(t))$ denote the sets of UE arrival rate (UAR) and UE departure rate (UDR) in time t, in which $\lambda_k(t)$ and $\mu_k(t)$ respectively represent the UAR and UDR of class k in time t.

When the SBS k is switched off, its UEs will be delivered into MBS and its UAR and UDR will become 0. Therefore, the actual UAR and UDR of class k can be calculated by:

$$\lambda_0^a(t) = \lambda_0(t) + \sum_{k \in \mathcal{K}} \lambda_k(t) \cdot 1_{\{y_k=0\}}, \mu_0^a(t) \approx \mu_0(t)$$

$$\lambda_k^a(t) = \lambda_k(t) \cdot 1_{\{y_k=0\}}, \mu_k^a(t) = \mu_k(t) \cdot 1_{\{y_k=0\}}, k \in \mathcal{K} \tag{3}$$

where $1_{\{\Theta\}}$ is an indicator function that equals to 1 if Θ is satisfied and 0, otherwise. The smaller the number of the off-state SBSs in the network is , the lower interference the MBS will experience and the higher UDR the MBS will possess. However, the link quality between the MBS and the UEs transferred from the SBSs is worse than that of class 0, which will lower the UDR of MBS. Due to the specific impact of the two cases is difficult to predict, that the negative impact on the UDR can be offset by the positive one is assumed. Therefore, whatever the action $\mathbf{y}(t)$ is taken, the actual UDR of MBS is assumed to approximate the original one.

Referring to the literature [18], the UE arrival and UE departure of any class are two independent stochastic processes in the paper. Specifically, UE arrival of class k in time t is a poisson process with parameter $\lambda_k^{\mathrm{a}}(t)T$, such that the probability of having p UEs in class k within time t is expressed as follows:

$$P\{\,p \text{ users arrive}\} = \frac{(\lambda_k^{\mathrm{a}}(t)T)^p \, e^{-\lambda_k^{\mathrm{a}}(t)T}}{p!} \tag{4}$$

Similar to user arrival, the UE departure of class k during time t follows a Poisson process with parameter $\mu_k^{\mathrm{a}}(t)T$, so the probability of serving q UEs within time t is expressed as follows:

$$P\{q \text{ users depart}\} = \frac{(\mu_k^{\mathrm{a}}(t)T)^q \, e^{-\mu_k^{\mathrm{a}}(t)T}}{q!} \tag{5}$$

3. Problem Formulation

Assume that $r_{k,n}(t)$ denotes the required resource units of UE n served by the BS k and U_k is the total number of resource units in BS k, then the system load of BS k can be obtained by:

$$\rho_k = \left(\sum\nolimits_{n \in x_k(t)} r_{k,n}(t) \right) / U_k, k = 0,1,...,K \tag{6}$$

In each time t, if an on/off decision set $\mathbf{y} = \mathbf{y}(t)$ is executed under a network state $\mathbf{x} = \mathbf{x}(t)$, the power consumption of the MBS can be expressed as:

$$P_0^{(\mathrm{m})}(\mathbf{x},\mathbf{y}) = P_{\mathrm{cst}}^{(\mathrm{m})} + \alpha^{(\mathrm{m})} \rho_0 P_{\mathrm{tx}}^{(\mathrm{m})} \tag{7}$$

where $P_{\mathrm{cst}}^{(\mathrm{m})}$ denotes the fundamental power consumption of MBS caused by the signal processing unit, and $\alpha^{(\mathrm{m})}$ is the linear dependence factor of transmission power $P_{\mathrm{tx}}^{(\mathrm{m})}$. For the SBSs consume no power in the inactive states, the power consumption of the SBS k is given by:

$$P_k^{(s)}(\mathbf{x},\mathbf{y}) = (P_{cst}^{(s)} + \alpha^{(s)}\rho_k P_{tx}^{(s)}).1_{\{y_k=1\}}, k \in 1,2,...,K \tag{8}$$

where $P_{cst}^{(s)}$ denotes the fundamental power consumption of SBSs and $\alpha^{(s)}$ is the dependence factor of transmission power $P_{tx}^{(s)}$.

The total network power consumption in time t can be calculated by:

$$P(\mathbf{x},\mathbf{y}) = P_0^{(m)}(\mathbf{x},\mathbf{y}) + \sum_{k\in\mathcal{K}} P_k^{(s)}(\mathbf{x},\mathbf{y}) \tag{9}$$

One crucial metric for measuring QoS is the blocking probability of the UEs. In time t, the $U_k - u_k(t)$ resource units of BS k are free ,where $u_k(t)$ denotes the number of occupied resource units of BS k in time t. A UE blocking occurs when a new arrival UE n finds the associated BS k less than r_n free resource units. Therefore, the blocking probability of BS k at time t is defined as follows:

$$P_{blocking}^k(t) = \sum_{r_{k,n}}(P\{\text{new user } (k,n) \text{ requires } r_{k,n} \text{ resource}$$
$$\text{units}\} \times P\{U - u_k(t) < r_{k,n}\}), k = 0,1,...,K \tag{10}$$

The primary goal of the paper is to find an on/off strategy $\omega: X \rightarrow Y$ that selects the optimal working modes $\mathbf{y}$ for all SBSs in every network state $\mathbf{x}$, under which the long-term average energy consumption over the whole period can be minimized subject to the blocking probability constraints. Therefore, the optimization objective can be formulated as follows:

$$\min_{\omega\in\Omega}\overline{P} = \limsup_{t\rightarrow\infty}\frac{1}{t}\sum_{\tau=0}^{t-1}E\{|P(\mathbf{x}(\tau),\mathbf{y}(\tau))|\}$$
$$s.t. \ C1: P_{blocking}^k(t) < P^{th}, \forall t, k \in 0,1,...,K \tag{11}$$

where Ω denotes the set of all available on/off strategies and P^{th} is the blocking probability threshold to guarantee the QoS of UEs.

In each time epoch, the working states of SBSs in the networks need to be decided to realize the minimization of average energy consumption over the whole period. Accordingly, the DTMDP is an effective decision-making algorithm that can choose the relevant actions in each time epoch to obtain the optimal long-term expected cost or reward.

4. The Solution to BS Switching Operations

4.1 Map the Problem to DTMDP

Referring to **Fig. 2**, to solve the long-term average optimization goal formulated in (11), a DTMDP $< X, Y, A, C >$ can be defined as follows:

• **State space :** $\{\mathbf{x}: \forall \mathbf{x} \in X\}$, the network load vectors;

- **Action space :** $\{\mathbf{y} : \forall \mathbf{y} \in Y\}$, the on/off operations of all SBS;
- **Transition probability matix :** $\{P(\mathbf{x}^{\cdot} \mid \mathbf{x}, \mathbf{y}) \in A\}$;
- **Cost Function :** $\{C(\mathbf{x}, \mathbf{y}) = P(\mathbf{x}, \mathbf{y})\}$ as defined in (7).

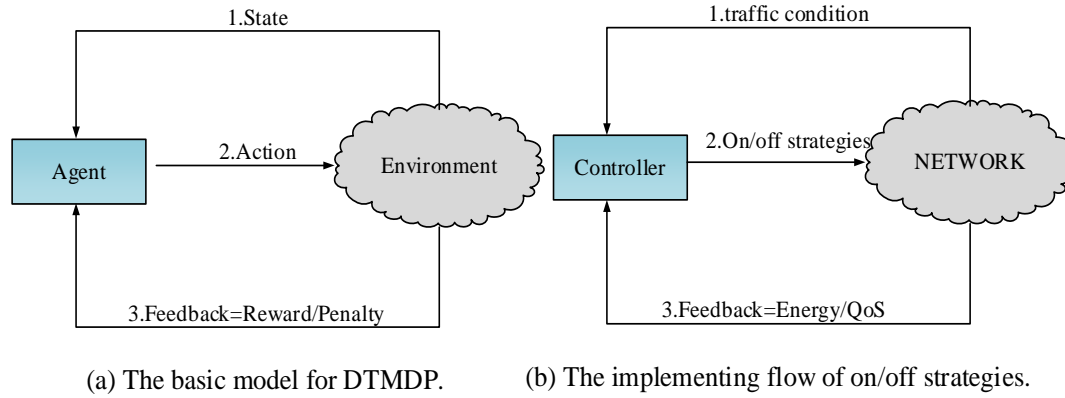(a) The basic model for DTMDP.   (b) The implementing flow of on/off strategies.

Fig. 2. The rationale for using DTMDP in BS switching operations.

In DTMDP, every network state $\mathbf{x} \in X$ is associated with an action $\mathbf{y} \in Y$, a corresponding state transition $P(\mathbf{x}^{\cdot} \mid \mathbf{x}, \mathbf{y}) \in A$, and a cost function $C(\mathbf{x}, \mathbf{y}) \in C$. The network controller observes the current network state $\mathbf{x}(t)$, to which an action $\mathbf{y}(t)$ is associated, then the both elements of the current time t transforms the network state from $\mathbf{x}$ to $\mathbf{x}^{\cdot}$ with a corresponding transition probability $P(\mathbf{x}^{\cdot} \mid \mathbf{x}, \mathbf{y})$, and a value of cost $C(\mathbf{x}, \mathbf{y})$ is generated as a feedback for the network controller.

For the UE states $i = x_k(t)$ and $j = x_k(t+1)$ of class k, let $z = j - i$ denote the UE net arrivals within time t. Note that the value of z can be positive or negative. It is positive when the arrival UEs are more than the depart ones, and negative otherwise. Hence, the user transition probability $p_{i,j,k}(t) = P(x_k(t+1) = j \mid x_k(t) = i, \mathbf{y}(t))$ of class k can be determined as follows:

$$p_{i,j,k}(t) = \begin{cases} P\{z \text{ net arrivals}\}, \text{if } j - i = z \\ 0, \qquad \text{otherwiswe} \end{cases} \tag{12}$$

where $P\{z \text{ net arrivals}\}$ is calculated by:

$$P\{z \text{ net arrivals}\} = P\{p \text{ arrivals}\} - P\{q \text{ departures}\}$$
$$= \frac{(\lambda_k^{\mathrm{a}}(t)T)^p e^{-\lambda_k^{\mathrm{a}}(t)T}}{p!} - \frac{(\mu_k^{\mathrm{a}}(t)T)^q e^{-\mu_k^{\mathrm{a}}(t)T}}{q!} \tag{13}$$

From (13), it can be noticed that z is the difference between two independent Poisson random variables, such that the probability distribution of z can be expressed as the Skellam distribution [18] :

$$P\{z \text{ net arrivals}\} = e^{-(\lambda_k^a(t)+\mu_k^a(t))T}(\frac{\lambda_k^a(t)}{\mu_k^a(t)})^{z/2}I_{|z|}(2T\sqrt{\lambda_k^a(t)\mu_k^a(t)}) \tag{14}$$

where $I_z(\cdot)$ denotes the modified Bessel function of the first kind and is defined as follows:

$$I_z(x) = \sum_{m=0}^{\infty}\frac{1}{m!\Gamma(m+z+1)}(\frac{x}{2})^{(2m+z)} \tag{15}$$

Combining (12) and (14), the complete UE transition probability expression of class k can be obtained as given in (16):

$$p_{i,j,k}(t) = \begin{cases} e^{-(\lambda_k^a(t)+\mu_k^a(t))T} \cdot (\frac{\lambda_k^a(t)}{\mu_k^a(t)})^{z/2} \cdot I_{|z|}(2T\sqrt{\lambda_k^a(t)\mu_k^a(t)}), \text{if } z = j-i \\ \\ \quad\quad\quad 0, \quad\quad\quad\quad \text{otherwise} \end{cases} \tag{16}$$

The transition probability matrix A can be determined by combining the user transition probability of every class.

The goal is to obtain a strategy ω, which maps a state $\mathbf{x}$ to an action $\omega(\mathbf{x})$, i.e., $\mathbf{y}$, to minimize the total expected discounted cost starting from the state $\mathbf{x}$. Generally, the cost of this kind is called a state value function, which is defined as follows [19]:

$$V^{\omega}(\mathbf{x}) = E_{\omega}\left[C(\mathbf{x},\mathbf{y}) + \gamma \sum_{\mathbf{x}\in X} P(\mathbf{x}'\mid \mathbf{x},\mathbf{y})V^{\omega}(\mathbf{x}') \right] \tag{17}$$

where the parameter γ ($\gamma \in (0,1)$) represents the discount factor that maps the impact of future cost to the current state. Importantly, the state value function defined in (17) is a recursive form of the long-term expected energy cost expressed in (11). Given the blocking probability constraints, i.e. C1, the optimal strategy ω^* for the long-term average goal (11) can be obtained by the following constrained Bellman equation [19]:

$$\omega^* = \omega^*(\mathbf{x}) = \arg\min_{\mathbf{y}\in Y}\left\{ E_{\omega^*}\left[C(\mathbf{x},\mathbf{y}) + \gamma \sum_{\mathbf{x}\in X} P(\mathbf{x}'\mid \mathbf{x},\mathbf{y})V^{\omega^*}(\mathbf{x}') \right]\Big| C1 \right\} \tag{18}$$

4.2 A DTMDP-based Asynchronous On/off Strategy

Assume that the paper adopts the synchronous on/off strategies for all the SBSs, and the minimum number of states for any class is M in the next time slot, then the size of state space is more than M^{K+1}. Furthermore, to avoid turning the SBSs on and off repeatedly in a short time, the value of duration T can not be very small, which will cause the relatively large value of M. In this case, the modest-scaled number of SBSs would cause the very large state space of the networks, which leads to the "curse of dimentionality" occurred in DTMDP frequently. Given these, the paper first proposes a DTMDP-based asynchronous on/off strategy, which

can lower the solution complexity of MDP greatly. The specific steps of the asynchronous on/off strategy are shown in **Fig. 3**, where $\mathbf{H}(t)$ denotes the on/off priority set of all SBSs and $x_0^k(t)$ represents the temporary state of MBS when the on/off decision of SBS k is performed.

For the capacity of MBS is limited and the blocking probability threshold is set to limit the number of switched-off SBSs, an on/off priority for SBSs is expected to be developed during the execution of the asynchronous on/off strategy. The priorities $\mathbf{H}(t)$ are arranged according to the UE numbers of all SBSs. The fewer the number of UEs of the SBS is, the earlier the on/off strategy of the SBS implements, such that more SBSs can be switched off and more network energy can be saved for the existence of the fixed power consumption.
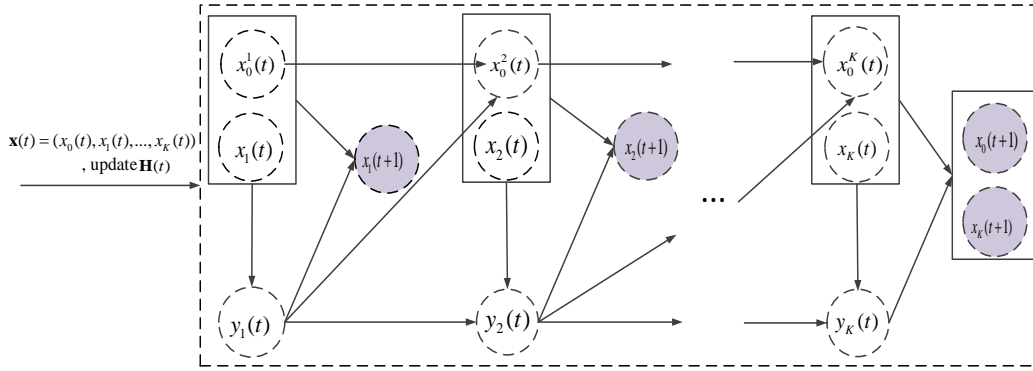

Fig. 3. The specific steps to the DTMDP-based asynchronous on/off strategies.

Theorem 1: The asynchronous on/off strategies are equivalent to the optimal ones.

Proof : Let $\omega_k^*(t), k = 1, 2, ..., K$ represent the optimal strategy of SBS k. For convenience, some temporary variables are defined as follows:

$$
\begin{aligned}
\mathbf{x}_k &= (x_0^k(t), x_k(t)) = (x_0^k, x_k), \forall k \in \mathcal{K}; \\
\mathbf{x}_k^{'} &= (x_0^{k+1}(t), x_k(t+1)) = (x_0^{k+1}, x_k^{'}), k = 1, ..., K-1; \\
\mathbf{x}_K^{'} &= (x_0^1(t+1), x_K(t+1)) = (x_0^{'}, x_K^{'}).
\end{aligned}
\tag{19}
$$

According to the Bellman equation, the optimal on/off strategy of SBS k is determined as follows:

$$
\omega_1^* = \underset{y_1 \in Y}{\arg\min} \left[c^1 + \gamma \sum_{\mathbf{x}_1^{'} \in \mathcal{X}} P(\mathbf{x}_1^{'} \mid \mathbf{x}_1, y_1) V^{\omega_1^*}(\mathbf{x}_1^{'}) \mid C1 \right]
$$

$$
...
$$

$$
\omega_k^* = \underset{y_k \in Y}{\arg\min} \left[c^k + \gamma \sum_{\mathbf{x}_k^{'} \in \mathcal{X}} P(\mathbf{x}_k^{'} \mid \mathbf{x}_k, y_k) V^{\omega_k^*}(\mathbf{x}_k^{'}) \mid C1 \right]
$$

$$
...
$$

$$
\tag{20}
$$

where $c^k = P_0^{(m)}(\mathbf{x}, \mathbf{y}) + P_k^{(s)}(\mathbf{x}, \mathbf{y})$ denotes the temporary cost function after the decision of SBS k.

The entire policy $\omega^* = (\omega_1^*, \omega_2^*, ..., \omega_K^*)$ can be obtained by combined the K expressions of (20) in (21):

$$
\begin{aligned}
\omega^* &= \arg\min_{\mathbf{y} \in Y} \left\{ \begin{bmatrix} KP_0^{(m)}(\mathbf{x}, \mathbf{y}) + \sum_{k \in \mathcal{K}} P_k^{(s)}(\mathbf{x}, \mathbf{y}) + \gamma \sum_{\mathbf{x}' \in X} P(\mathbf{x}_1') \mid \mathbf{x}_1, y_1) \\ \cdot P(\mathbf{x}_2' \mid \mathbf{x}_2, y_2) \cdot ... \cdot P(\mathbf{x}_K' \mid \mathbf{x}_K, y_K) \cdot V^{\omega^*}(\mathbf{x}') \end{bmatrix} \middle| C1 \right\} \\
&= \arg\min_{\mathbf{y} \in Y} \left\{ \begin{bmatrix} KP_0^{(m)}(\mathbf{x}, \mathbf{y}) + \sum_{k \in \mathcal{K}} P_k^{(s)}(\mathbf{x}, \mathbf{y}) + \gamma \sum_{\mathbf{x}' \in X} P(x_0^2, x_1') \mid x_0^1, x_1, y_1) \\ \cdot P(x_0^3, x_2' \mid x_0^2, x_2, y_2) \cdot ... \cdot P(x_0', x_K' \mid x_0^K, x_K, y_K) \cdot V^{\omega^*}(\mathbf{x}') \end{bmatrix} \middle| C1 \right\}
\end{aligned}
\tag{21}
$$

The value function is a nested model, i.e., $V^{\omega_k^*}(\mathbf{x}_k')$ has the same structure as its front part in formula (20). Therefore, if the front part of formula (21) and that of (18) are equivalent, so are the asynchronous on/off strategies and the optimal strategies.

For the coefficient before $P_0^{(m)}(\mathbf{x}, \mathbf{y})$ has no impact on the optimal strategies and $x_1', x_2', ..., x_K' \in \mathbf{x}'$ are independent from each other, furthermore, the current action y_k of SBS k only influences the next state x_0' of MBS and x_k' of its own, so the equation (22) can be derived as follows:

$$
\begin{aligned}
&\sum_X P(x_0^2, x_1' \mid x_0^1, x_1, y_1) \cdot P(x_0^3, x_2' \mid x_0^2, x_2, y_2) \cdot ... \cdot P(x_0', x_K' \mid x_0^K, x_K, y_K) \\
&= \sum_X P(x_0', x_1', x_2', ..., x_K' \mid x_0, x_1, x_2, ..., x_K, y_1, y_2, ..., y_K) \\
&= \sum_{\mathbf{x}' \in X} P(\mathbf{x}' \mid \mathbf{x}, \mathbf{y})
\end{aligned}
\tag{22}
$$

Therefore, the optimality of the proposed asynchronous on/off strategies is proved.

The optimal asynchronous strategies can be obtained based on the on/off priorities and Bellman equation, whose detailed steps are shown in **Algorithm 1**.

Notably, the proposed asynchronous solving algorithm has a complexity of $O(KM^2)$, which is much lower than $O(M^{K+1})$. Therefore, the proposed algorithm not only can obtain the optimal on/off strategies but has a stable performance when the network size expands.

5. Simulation Results

A relative simple but representative two-tier heterogeneous network, which consists of one MBS and 14 SBSs is considered in the numerical simulations of this paper. According to the practical network condition, the variables about the power consumption are set as follows: $P_{cst}^{(s)} = 4.8W$; $P_{tx}^{(s)} = 0.05W$; $P_{cst}^{(m)} = 130W$; $P_{tx}^{(m)} = 20W$; $\alpha^{(m)} = 4.7$; $\alpha^{(s)} = 8$ [17]. The

---

Algorithm 1: The DTMDP-based asynchronous on/off strategies

Input: The network state: $\mathbf{x} = (x_0(t), x_1(t), ..., x_K(t))$,

the transition probability matrix: $\{p_{i,j,k}(t), k \in 0,1,...,K\}$

Output: the optimal on/off strategy of the SBSs: $\omega^*(t)$

1: Obtain the on/off priority set $\mathbf{H}(t)$ based on the UE numbers of all SBSs

2: Number the SBSs by $k \in 1,2,...,K$ according to the on/off priority set $\mathbf{H}(t)$

3: Initialize $k = 0$

4: **while** ($P^i_{\text{blocking}}(t) < P^{\text{th}}, i \in 0,1,...K$) **do**

6: Determine the on/off state of all SBSs one by one: $k = k+1$

7:     Obtain the optimal on/off strategy of the SBS k according to the Bellman equation:

$$\omega_k^*(t) = \arg\min_{y_k \in Y} \left\{ c^k(t) + \gamma \sum_{\mathbf{x}_k(t+1) \in X} P(\mathbf{x}_k(t+1) \mid \mathbf{x}_k(t), y_k(t)) V^{\omega_k^*}(\mathbf{x}_k(t+1)) \right\}$$

8: **end**

9: **for** $l = k : K$ **do**

10:    The current on/off state of SBS l is same as the previous one: $\omega_l^*(t) = \omega_l^*(t-1)$

11: **end**

12: Obtain the optimal on/off strategy set: $\omega^*(t) = (\omega_1^*(t), \omega_2^*(t), ...\omega_K^*(t))$

---

duration T is set to be 1 min to achieve a trade-off between the dimension of state space and the state change times of SBSs. The proposed asynchronous on/off algorithm is compared with the conventional network operation, in which all BSs are always in active state and there is no switching operation being conducted. For further comparisons, the paper considers the other three strategies, the heuristic sleep approach proposed in [9], a similar asynchronous on/off strategy to the proposed one but without priorities and the general synchronous on/off decisions based on DTMDP. Matlab is used as the simulation tool to verify the performance of the proposed algorithm. The presented results are averaged a great number of independent runs with different practical configurations.

Fig. 4 shows the changing trend about the number of active SBSs when the arrival rate of the MBS varies. It can be observed that as the arrival rate of the MBS increases, the number of active SBSs in the network increases accordingly. Furthermore, the change trends of the two variables are not in synchrony for the proposed DTMDP-based asynchronous on/off strategies, which can bring the effect of future cost into the current decisions, will adjust the number of active SBSs prior to the exhaustion of the resource units of MBS.

Fig. 5 presents the converging process of the energy consumption for different arrival rates of the MBS as the iteration times in one epoch increase. Through analyzing the converging trend, the total energy consumption of the network will decrease gradually for the increased number of inactive SBSs as the asynchronous on/off strategy performs iteratively. Besides, it can be found that the simulation configurations with large arrival rates need few iteration times. It is because when the new arrival UEs increase, the number of the SBSs that can be turned off will decrease, and the iteration times for the convergence of the proposed strategy will decrease accordingly.
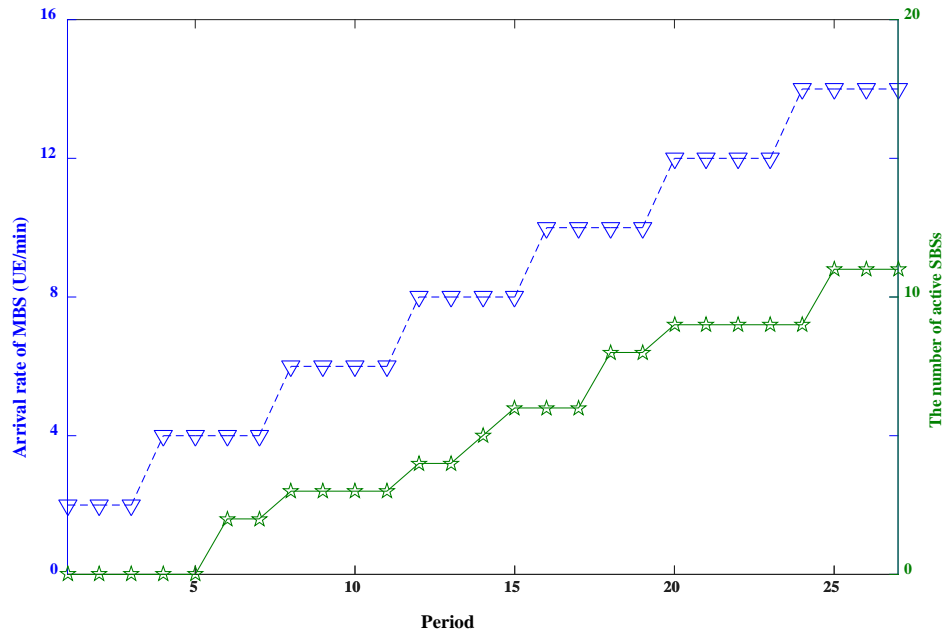
Fig. 4. The changing trend about the number of active SBSs when the arrival rate of the MBS varies.
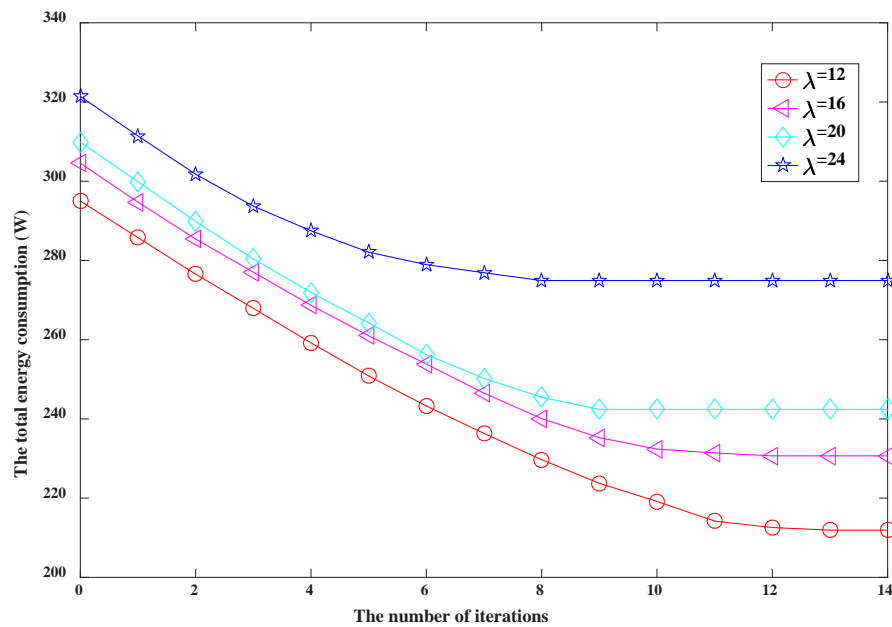

Fig. 5. The converging process of the energy consumption for different arrival rates of the MBS.

Fig. 6 shows the influence of blocking probability thresholds on the total energy consumption and QoS of the networks. By comparing the subfigures (a) and (b), it can be found that when the blocking probability threshold is set larger, the total energy consumption will decrease and the average number of blocking UEs will increase. Through sacrificing the

service quality of the networks, the higher blocking probability threshold can bring more energy conservation for more SBSs can be switched off to save energy. Therefore, the value of blocking probability threshold can be dynamically adjusted based on the different sensitivity of different networks to the energy consumption and QoS.
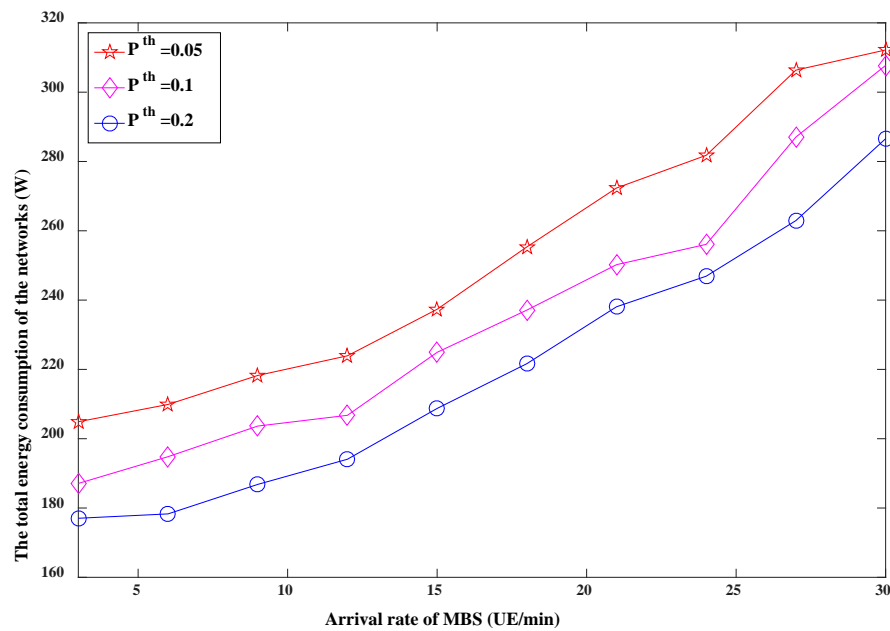
(a) The relationship between the blocking probability thresholds and the total energy consumption
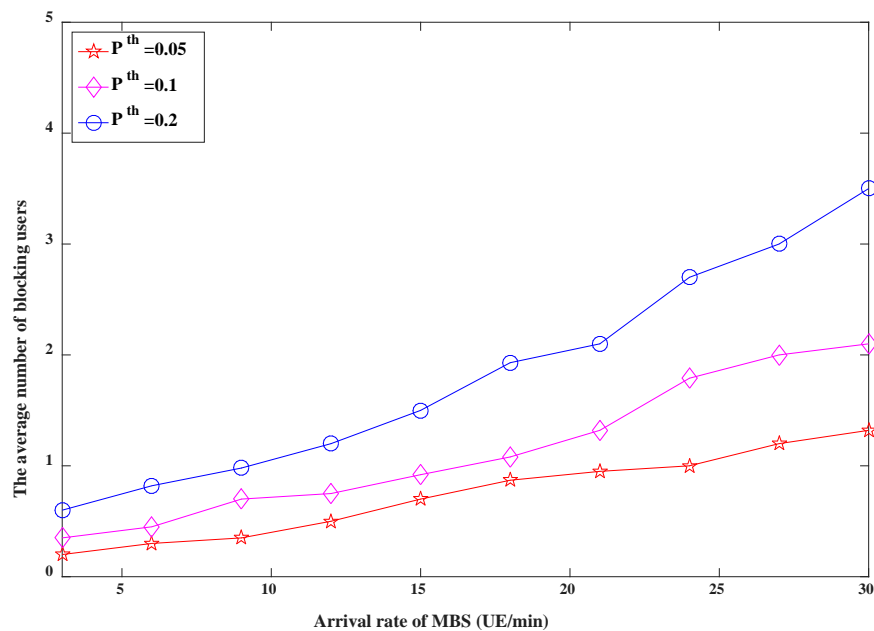
(b) The relation between the blocking probability thresholds and the blocking UE numbers.
Fig. 6. The influence of Pth on the total energy consumption/ the average number of blocking UEs.

Fig. 7 compares the different on/off strategies about the total energy consumption when the arrival rate of MBS varies. As expected, the proposed asynchronous on/off strategy performs much better than the conventional network operation and the similar asynchronous algorithm to the proposed approach but without priorities. It can be observed that the proposed asynchronous on/off strategy can achieve a better performance compared with the heuristic sleep approach, which obtains the sub-optimal on/off strategy according to the load of BSs and the interference relation matrix. Furthermore, when the arrival rate of MBS is low, the energy-saving performance of the heuristic sleep approach is a little worse than the asynchronous on/off algorithm without priorities, but when the arrival rate of MBS increases, the situation is reversed. It is because the priority, which decides the switching order of the SBSs, has a greater influence on the energy-saving performance as the SBSs that can be switched off decrease. It is noteworthy that such a simple strategy with low computational complexity can obtain the same performance with the general synchronous on/off decisions, whose complexity is exponential level and should be avoided in practical networks. Therefore, the availability and effectiveness of the proposed strategies are further validated.
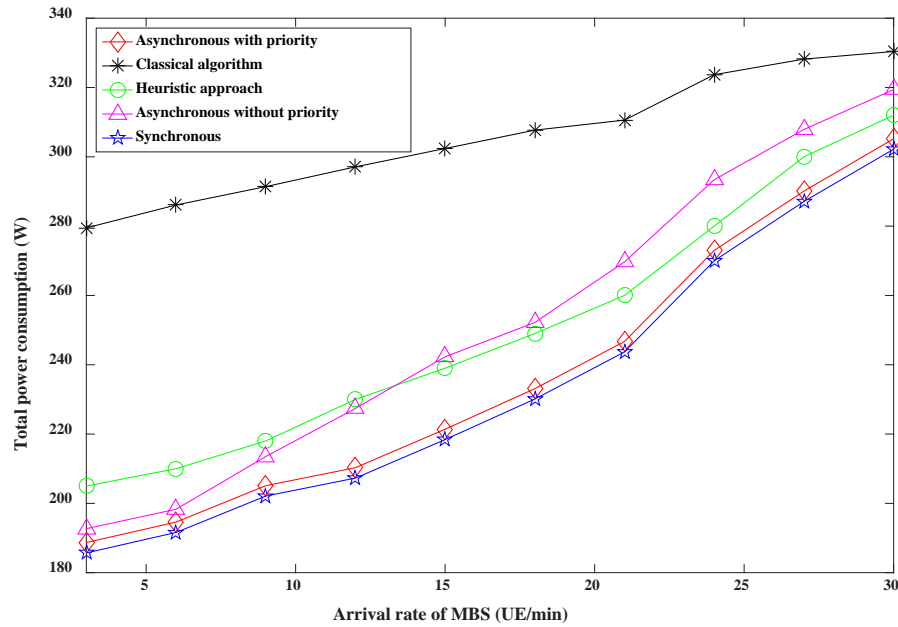

Fig. 7. The comparison of the total energy consumption among different strategies.

6. Conclusion

A novel asynchronous on/off strategy has been first presented in the paper to save energy by switching off some underutilized SBSs. Specifically, the BS switching operations was formulated as a DTMDP under the dynamic traffic load. Besides, the optimal asynchronous algorithm, which was based on the on/off priorities of the SBSs and the Bellman equation, has been designed to obtain the BS switching solution for decreasing the overall energy consumption. The DTMDP-based asynchronous on/off strategies can avoid the "curse of dimensionality "through reducing the state and action space during one decision. Finally, the optimality of the proposed algorithm was proved. The extensive simulation results have validated the effectiveness and robustness of the asynchronous on/off strategies under various

practical configurations.

References

[1] H. Zhang, A. Gladisch, M. Pickavet, Z. Tao and W. Mohr, "Energy efficiency in communications," *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 48–49, 2010. Article (CrossRef Link)

[2] Z. Niu, "TANGO: traffic-aware network planning and green operation," *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 25-29, 2011. Article (CrossRef Link)

[3] AA. Abdulkafi, TS. Kiong and IK. Sileh, "A Survey of Energy Efficiency Optimization in Heterogeneous Cellular Networks," *Ksii Transactions on Internet & Information Systems*, vol. 10, no. 2, pp. 462-483, 2016. Article (CrossRef Link)

[4] H. Gao, W. Ejaz and M.Jo, "Cooperative Wireless Energy Harvesting and Spectrum Sharing in 5G Networks," *IEEE Access*, pp.3647-3658, 2017. Article (CrossRef Link)

[5] H. Gao, Y. Du and C. Li, "Quantum fireworks algorithm for optimal cooperation mechanism of energy harvesting cognitive radio," *Journal of Systems Engineering and Electronics*, vol. 29, no. 1, pp.18-30, 2018. Article (CrossRef Link)

[6] J. Zhang, Q. Zhou, D. W. K. Ng and M. Jo, "Optimal Energy Efficiency Fairness of Nodes in Wireless Powered Communication Networks," *Sensors*, vol.17, no. 9, 2125, 2017. Article (CrossRef Link)

[7] Q. Wang and J. Zheng, "A Distributed base station On/Off Control Mechanism for energy efficiency of small cell networks," in *Proc. of IEEE Int. Conf. Commun. (ICC)*, pp. 3317-3322, 2015. Article (CrossRef Link)

[8] N. Yu, Y. Miao and L. Mu, "Minimizing Energy Cost by Dynamic Switching ON/OFF Base Stations in Cellular Networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7457-7469, 2016. Article (CrossRef Link)

[9] G. Su, B. Chen and X. Lin, "User association and base station sleep management in dense heterogeneous cellular networks.," *Ksii Transactions on Internet & Information Systems*, vol. 11, no.4, pp. 2058-2074, 2017. Article (CrossRef Link)

[10] J. Kim, P.Y. Kong, N.O. Song, J.K.K. Rhee and S. Al-Araji, "MDP Based Dynamic Base Station Management for Power Conservation in Self-Organizing Networks," in *Proc. of IEEE Wireless Comm. and Networking Conf. (WCNC)*, pp. 2561-2567, 2014. Article (CrossRef Link)

[11] E. Oh, K. Son and B. Krishnamachari, "Dynamic Base Station Switching-On/Off Strategies for Green Cellular Networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2126-2136, 2013. Article (CrossRef Link)

[12] S. Samarakoon, M. Bennis and W. Saad, "Dynamic Clustering and on / off Strategies for Wireless Small Cell Networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2164-2178, 2015. Article (CrossRef Link)

[13] Y. Yang, L. Chen and W. Dong, "Active Base Station Set Optimization for Minimal Energy Consumption in Green Cellular Networks," *IEEE Trans. Veh. Technol.*, 2015, vol. 64, no. 11, pp. 5340-5349, 2015. Article (CrossRef Link)

[14] L. Tang, W.Wang, Y.Wang and Q. Chen, "An energy-saving algorithm with joint user association, clustering and on/off strategies in dense heterogeneous networks," *IEEE Access*, pp.12988-13000, 2017. Article (CrossRef Link)

[15] L. Sakar, L. E. Elayoubi, R. Combes and T. Chahed, "Optimal Control of Wake Up Mechanisms of Femtocells in Heterogeneous Networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp.664-672, 2012. Article (CrossRef Link)

[16] R. Li, Z. Zhao, X. Chen, J. Palicot and H. Zhang, "TACT: A transfer actor-critic learning framework for energy saving in cellular radio access networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 2000–2011, 2014. Article (CrossRef Link)

[17] X. Chen, J. Wu, Y. Cai, H. Zhang and T. Chen, "Energy-efficiency oriented traffic offloading in wireless networks: a brief survey and a learning approach for heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 4, 627-640, 2015. Article (CrossRef Link)

[18] P.Y. Kong, "Optimal probabilistic policy for dynamic resource activation using markov decision process in green wireless networks," *IEEE Trans. Mobile Comput.*, vol. 13, no.10, pp. 2357-2368, 2014. Article (CrossRef Link)

[19] R. Sutton and A. Barto, Reinforcement Learning: An Introduction. Cambridge University Press, 1998. Available: Article (CrossRef Link)

LUN TANG received the Ph.D. degree in communication and information systems from Chongqing University Posts and Telecommunications, Chongqing, China. He is currently a Professor with the School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications. His research interests include 5G cellular networks, interference management, and small cell networks.

WEILI WANG received the B.Sc. degree from Chongqing University of Posts and Telecommunications, Chongqing, China, in 2015, where she is currently pursuing the M.S. degree in communication and information systems. Her research interests are in 5G cellular networks, with an emphasis on resource allocation techniques for energy efficiency in dense heterogeneous networks.

QIANBIN CHEN received the Ph.D. degree in communication and information systems from the University of Electronic Science and Technology of China, Chengdu, China, in 2002. He is currently a Professor with the School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, and the Director of the Chongqing Key Laboratory of Mobile Communication Technology. He has authored or co-authored more than 100 papers in journals and peer-reviewed conference proceedings, and has co-authored seven books. He holds 47 granted national patents.