

BS sleeping strategy for energy-delay tradeoff in wireless-backhauling UDN

Pei LI, Yi SHEN, Faisal SAHITO, Zhiwen PAN* & Xiaohu YOU

National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China

Received 25 January 2018/Revised 20 March 2018/Accepted 29 May 2018/Published online 22 February 2019

Abstract Ultra-dense network (UDN) has been recognized as a promising technology for 5G. Although turning off low-load base stations (BSs) can improve energy efficiency, it may cause degradation of delay performance. This makes energy-delay tradeoff (EDT) an important topic. In this paper, a theoretical framework for EDT, in wireless-backhauling UDN, is developed. First, we investigate association probabilities of UEs and transmission probabilities of BSs. Expressions for energy consumption and network packet delay are obtained and the impact that BS sleeping ratio has on energy consumption and packet delay are analyzed. Then, we formulate the EDT problem as a cost minimization problem to select the optimal set of sleeping small cells. To solve the EDT optimization problem, a locally optimal sleeping ratio for EDT is obtained using the dynamic gradient iteration algorithm and we prove that it can converge to the global optimal sleeping ratio. Then, queue-aware and channel-queue-aware sleeping strategies are proposed to find the optimal set of sleeping small cells according to the optimal sleeping ratio. We then see that the simulation and numerical results confirm the effectiveness of the proposed sleeping schemes.

Keywords ultra-dense networks, sleeping ratio, sleeping strategy, wireless backhaul, energy-delay tradeoff

Citation Li P, Shen Y, Sahito F, et al. BS sleeping strategy for energy-delay tradeoff in wireless-backhauling UDN. *Sci China Inf Sci*, 2019, 62(4): 042303, <https://doi.org/10.1007/s11432-018-9494-9>

1 Introduction

With the popularization of smart devices and Internet of things, fifth generation (5G) cellular networks will have to deal with extremely dense data traffic and connections and ever-increasing capacity demands. Ultra-dense network (UDN) has been recognized as a promising solution for 5G [1], where overlaying macro cells with densely distributed small cells in hot spots constitutes an UDN. Due to the presence of a large number of small cells, we can obtain a higher spectral efficiency and throughput in UDN. However, this has placed a heavy burden on energy saving. Meanwhile, an increasing number of small cells and fiber backhaul links create not only significant cost problems, but also backhaul installation obstacle of small cells for network operators.

Since base station (BS) accounts for 60%–80% of the total system energy consumption [2], turning off (i.e., switching over to sleep or low-power mode) low-load BSs can save energy [3]. However, the decreased capacity might result in a negative impact on packet delay of user equipment (UE). Thus, energy-delay tradeoff (EDT) is an important topic in UDN.

Wireless-backhauling network is introduced to connect small cells to the gateway through wireless channels and subsequently manage the expensive backhaul deploying obstacles [4,5]. The deployment of dense small cells under the cover of macro cells in UDN will give rise to enormous data transmissions over

* Corresponding author (email: pzw@seu.edu.cn)

wireless backhaul links. The energy consumption and packet delay of the backhaul links might degrade the system performance, such as energy efficiency, packet delay, and so on [6]. Therefore, it is essential to consider the impact of wireless backhaul link while optimizing UDN system performance.

1.1 Related works

Many sleeping schemes have been proposed to improve energy efficiency [7–12]. A random sleeping strategy has been proposed to switch off each BS independently with a fixed probability ratio under the constraint of network coverage probability [7]. A distance-aware sleeping strategy identifies small cells present at undesirable interference spots and selects them for deactivation [8]. In [9], a traffic-aware cell management scheme is proposed to switch low-load BSs into sleeping mode and transfer their traffic load to the neighboring cells. In [10], a distributed learning mechanism for opportunistic sleeping mode is established based on game theory. A greedy on/off sleeping strategy for UDN is proposed to maximize energy efficiency based on traffic load, network topology, and user requirements [11]. In [12], iterative algorithms are proposed to facilitate a downlink cooperative transmission system by adaptively switching off BSs and antenna along with power allocation.

Recently, exploiting BS sleeping strategy for EDT as a method to reduce energy consumption while ensuring tolerable delay has garnered a lot of interest [13–16]. In [13], a random sleeping strategy and a traffic-aware sleeping strategy are analyzed to maximize energy efficiency under delay constraint. In [14], a random sleeping strategy, based on N -policy M/G/1 queueing model, is introduced to save energy while satisfying the delay requirement, where each BS enters the sleep mode independently without considering of the state of other BSs. Authors in [15] take into account BS sleeping and cell association to analyze EDT problem. Greedy-on and greedy-off sleeping strategies are investigated to achieve flexible EDT without any consideration of dynamic interference. In our previous work [16], a random sleeping strategy was used to analyze the EDT problem. The impact of sleeping time on energy consumption and delay are discussed and the optimal sleeping time for EDT problem is also investigated. However, in these cases [13–16], wired backhaul is adopted, hence energy consumption and delay of backhaul link are not considered for EDT in UDN.

Energy efficiency of wireless-backhauling networks is investigated in [17–19]. Ref. [17] proposes a cell selection scheme for small cell networks with constrained backhaul link. Power and backhaul bandwidth allocation are investigated to maximize energy efficiency for OFDMA heterogeneous small cell networks in [18]. In [19], a forward and backhaul link energy efficiency optimization scheme to improve system energy efficiency of small cell networks is proposed.

The impact of backhaul delay on system performance has been investigated in some recent studies. In [20], packet transmission delay of backhauling and radio access networks are analyzed, and a delay-minimum association scheme is proposed based on cell range expansion. A delay aware user association scheme for the backhaul constrained small cell networks is investigated in [21], an M/M/1 queueing model is employed to analyze the packet delay. In [22], delay limited and delay tolerant link allocation policy for multiuser systems with hybrid radio frequency and free space optical backhaul are proposed to assign the transmission time, nevertheless, delay is not analyzed and the objective function is maximizing system throughput under the constraint of delay.

However, there have been no studies on EDT related issues for wireless-backhauling UDNs.

1.2 Motivation and contribution

Sleeping strategy is adopted in UDN to reduce energy consumption. However, this may cause a negative impact on packet delay of UEs. Thus, EDT is an important problem. Ref. [23] analyzes the EDT for wired-backhauling networks, where authors decompose the EDT problem into delay minimization and energy cost minimization sub-problem. However, the distributed energy efficiency-oriented partial spectrum reuse scheme cannot achieve the social optimum and the BSs are always in active mode. Moreover, compared to wired backhaul, EDT problem is more challenging in wireless-backhauling UDNs and there have been no studies on EDT related problems for wireless-backhauling UDNs.

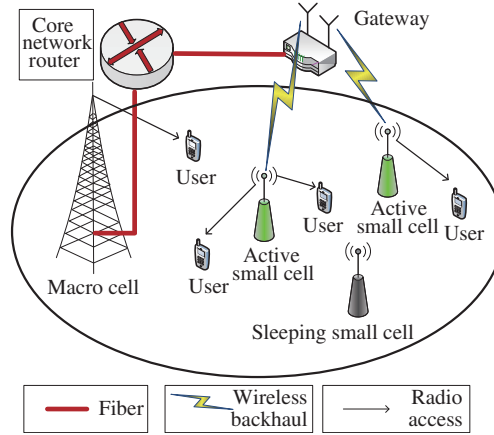


Figure 1 (Color online) Network model for two-tier wireless-backhauling UDN.

In this paper, we analyze the packet delay and energy consumption for both radio access and backhaul link of wireless-backhauling UDNs, with BS sleeping strategy, along with the impact of energy consumption and packet delay, in a backhaul link, on system energy consumption and network packet delay. Furthermore, we also analyze the EDT for wireless-backhauling networks with dynamic traffic load to identify the set of sleeping small cells.

The main contributions can be summarized as follows.

- Association probabilities of UEs and transmission probabilities for BSs are investigated using statistical analysis.
- A mathematical model is developed to obtain system energy consumption and mean packet delay of UEs. Then, the impact of BS sleeping ratio on system energy consumption and mean packet delay are discussed along with an investigation on the impact of energy consumption and mean packet delay for backhaul link.
- An EDT problem is formulated as a cost minimization problem to find the optimal set of sleeping small cells. To solve the EDT problem, we use a dynamic gradient iteration algorithm to achieve the locally optimal solution of sleeping ratio for the EDT problem in the first step. In the first step we prove that it can converge to the global optimal sleeping ratio and in the second step the sleeping strategy is operated to determine the set of sleeping small cells.
- Low complexity queue-aware and channel-queue-aware sleeping strategies are proposed, in the second step, to determine the set of sleeping small cells according to the optimal sleeping ratio. Compared to queue-aware sleeping strategy, it is safe to say that channel-queue-aware sleeping strategy can ensure better system performance.

1.3 Paper organization

The remainder of this paper is organized as follows. System model and operation model are introduced in Section 2. In Section 3, energy consumption and delay performance is analyzed. Section 4 investigates the EDT optimization problem. Numerical and simulation results are provided in Section 5, and Section 6 relays the conclusion.

2 System model and operation modes

2.1 Network model

As shown in Figure 1, consider a two-tier UDN consisting of radio access and backhauling networks with macro cell, small cell, gateway, and the core network router as its components [20]. In radio access network, UEs can associate with macro or small cell through a wireless link. Small cell can connect to gateway through a wireless backhaul link. Macro cell is directly connected to the core network router

via dedicated fiber, while small cell is linked to the gateway, which is then connected to the core network router via dedicated fiber [20].

Assume that the locations of gateways, macro cells, small cells, and UEs are distributed according to independent Poisson point process (PPP) Φ_g , Φ_m , Φ_s and Φ_u with intensity λ_g , λ_m , λ_s and λ_u ($\lambda_u > \lambda_s > \lambda_g$, $\lambda_u > \lambda_s > \lambda_m$), respectively [7, 20, 23]. Assume that the radio access network and backhauling network operate on different frequencies. All the backhaul links share a dedicated spectrum with bandwidth W_b in the backhauling network. The spectrum of radio access network is orthogonally partitioned and allocated to small cells and macro cells with bandwidth W_s and W_m , respectively.

Take into consideration the downlink of the system, the traffic model of each UE is Poisson traffic [16]. Assume that the data packet of each UE arrives according to an independent PPP with an arrival rate of λ and the average size of packet is l bits [15, 16].

2.2 Transmission model

Macro cells and gateways operate in active mode, while small cells operate in either active or sleeping mode. In the sleeping mode, transceivers and other hardware components of BS are turned off and BS cannot provide service for UEs and operates in low power level [24]. Whereas in active mode, packets arrive according to PPP. Thus, BSs and gateways are in busy state when they have packets to transmit, otherwise, they are in idle state when there is no packet in the buffer. Furthermore, to prevent UE with low signal to interference ratio (SIR) from occupying too much resource, packet transmission is successful only if the SIR of UE is above a specific threshold β [20, 21].

If $|\mathcal{A}|$ denotes the area of the entire network, then the number of gateways, macro cells, and small cells are $\lambda_g|\mathcal{A}|$, $\lambda_m|\mathcal{A}|$, and $\lambda_s|\mathcal{A}|$, respectively. The BS sleeping ratio θ , is defined as the ratio of the number of sleeping small cells to the total number of small cells in the system, that is $\theta = N_{\text{off}}/N_s$, where $N_s = \lambda_s|\mathcal{A}|$ represents the number of small cells and N_{off} is the number of sleeping small cells in the system. $\mathcal{S} = (s_1, s_2, \dots, s_k, \dots, s_{N_s})$ denotes the state set of small cells, where $s_k \in \{0, 1\}$. $s_k = 0$ means that the small cell k is in sleeping mode and $s_k = 1$ represents that the small cell k is in active mode.

Gateways are always in a busy state for active mode with the average transmission probability of 1 due to the enormous data transmissions over wireless backhaul links. If $\bar{\xi}_m(\theta)$ and $\bar{\xi}_s(\theta)$ denote the average transmission probability of macro cell and small cell in active mode (i.e., the probability that the macro or small cell is in a busy state for the active mode). The explicit expressions for $\bar{\xi}_m(\theta)$ and $\bar{\xi}_s(\theta)$ are not available, but equivalent conditions of $\bar{\xi}_m(\theta)$ and $\bar{\xi}_s(\theta)$ will be discussed in the following part.

A biased association policy, based on maximum biased-received-power, is used [25]. The probability that a typical UE associates with small cell is given by [20]

$$\text{Pr}_{\text{SUE}}(\theta) = \frac{(1 - \theta)\lambda_s(A_b P_{\text{st}})^{\frac{2}{\alpha}}}{(1 - \theta)\lambda_s(A_b P_{\text{st}})^{\frac{2}{\alpha}} + \lambda_m P_{\text{mt}}^{\frac{2}{\alpha}}}, \quad (1)$$

where, A_b is the bias factor of small cell, the path loss exponent is α for all links [20, 21]. P_{mt} and P_{st} are the transmit power of the macro and small cell, respectively.

The background noise is neglected [7, 20, 21] and SIR of a typical UE can be expressed as

$$\text{SIR} = \frac{P_t h_{x_0} \|x_0\|^{-\alpha}}{\sum_{x \in \phi \setminus \{x_0\}} P_t h_x \|x\|^{-\alpha}}, \quad (2)$$

where, $\|x_0\|$ and $\|x\|$ indicate the distances from UE to the serving cell and the interfering cell, respectively, and $\phi \setminus \{x_0\}$ refers to the set of interfering cells. BS transmit power is P_t . h_{x_0} and h_x are the channel gains of transmission and interfering link. Rayleigh fading is considered to make model tractable.

2.3 Queue model

As an effective approach of evaluating delay performance, queueing theory is frequently used to tackle the sleeping strategy. M/M/1 queueing model is the simplest model [21], where UEs (customers in

queueing model) arrive according to independent PPP with arrival rate, while service rate (defined as the number of customers served in unit time) follows the Poisson distribution, which is not the case in real system. Similar to the existing studies [14], an M/G/1 queueing model is used to analyze packet delay for wireless-backhauling UDNs in this paper, where the service rate follows general distribution.

Assume that the traffic arrival of packets for each UE follows PPP. The scheduling algorithm of BS is first come first out, which means that all the UEs associated with the same BS are waiting in a queue for service and will be served in order [14, 21]. The packet transmission rate can be obtained based on the Shannon capacity formula and it can be assumed that the service rate of BS follows the general distribution. Moreover, to avoid UE with low SIR occupying too much resource, the packet transmission is successful only if the SIR of UE is above a specific threshold β [20].

Since traffic arrival of packets for each UE is PPP, the total traffic at each BS and gateway is still governed by PPP [26]. As a result, each BS and gateway can be modeled as an M/G/1 queueing system, in other words the transmission process of packets can be viewed as a tandem queueing system and the queue length can be defined as the number of packets waiting for transmission in the buffer of BS.

2.4 Power consumption model

Denote the average power consumption of gateway as \bar{P}_G . Power consumption of macro cell in active mode is $\bar{P}_M = (1 - \bar{\xi}_m(\theta))\bar{P}_{m0} + \bar{\xi}_m(\theta)(\bar{P}_{m0} + \Delta p_m P_{mt})$ [24]. Similarly, the power consumption of a sleeping small cell is \bar{P}_S , the small cell power consumption in active mode is $\bar{P}_A = (1 - \bar{\xi}_s(\theta))\bar{P}_{s0} + \bar{\xi}_s(\theta)(\bar{P}_{s0} + \Delta p_s P_{st})$ [24]. \bar{P}_{s0} ($\bar{P}_{s0} \geq \bar{P}_S$) and \bar{P}_{m0} denote the constant power consumptions of a small cell and macro cell in active mode. Δp_s and Δp_m denote the slope of traffic load dependent power consumption for a small cell and macro cell. P_{st} and P_{mt} are the transmit power of small cell and macro cell, respectively.

3 System energy consumption and packet delay

In this section, we quantitatively analyze system energy consumption and mean packet delay. Since it is difficult to analyze the impact of the set of sleeping small cells on energy consumption and packet delay theoretically, the relationship between BS sleeping ratio, energy consumption, and packet delay are investigated, where the value of sleeping ratio depends on the set of sleeping small cells.

3.1 System energy consumption

System energy consumption consists of radio access network energy consumption and backhauling network energy consumption. Macro cells and small cells account for the main energy consumption in radio access network. Macro cell and gateway are connected to the core network router via dedicated fiber, while small cells are linked to gateway through the wireless backhaul link. Since the backhaul links from gateway and macro cells to the core network router are dedicated fiber links with low delay, their delay and energy consumption are neglected [20, 27]. Therefore, the energy consumption of backhauling network refers to the energy consumption of gateways.

If $|\mathcal{A}|$ denotes the area of the entire network, then the number of gateways, macro cells, and small cells is given by $\lambda_g|\mathcal{A}|$, $\lambda_m|\mathcal{A}|$, and $\lambda_s|\mathcal{A}|$, respectively. Now, the average energy consumption of the system can be represented by

$$\bar{P}(\theta) = |\mathcal{A}|(\lambda_g \bar{P}_G + \lambda_m(\bar{P}_{m0} + \bar{\xi}_m(\theta)\Delta p_m P_{mt}) + (1 - \theta)\lambda_s(\bar{P}_{s0} + \bar{\xi}_s(\theta)\Delta p_s P_{st}) + \theta\lambda_s \bar{P}_S), \quad (3)$$

$\bar{\xi}_m(\theta)$ increases with the increase of θ , $\bar{\xi}_s(\theta)$ is an increasing function of θ as explained in Appendix A, assume $P_{mt} \geq P_{st}$, $\Delta p_m \geq \Delta p_s$ [24]. We can prove that system energy consumption is a decreasing function of θ [17]. Besides, increasing gateway density of UDN will lead to an increase of system energy consumption.

3.2 Mean packet delay

In this paper, packet delay is defined as the time difference between packet transfer request arriving and the transmission of the packet being served. In radio access network, UEs can associate with either small cell or macro cell, different association cases may lead to different packet delays. In a backhauling network, the links from macro cells and gateways to the core network router are dedicated backhaul links with low delay and it is ok to assume that the delay of the dedicated backhaul links can be neglected [20].

Therefore, when UE associates with a macro cell, the packet delay \bar{D}_m refers to the packet delay of radio access link from macro cell to UE. When UE associates with small cell, the packet delay \bar{D}_s refers to the packet delay of radio access link from small cell to UE \bar{D}_{sr} and the packet delay of backhaul link from gateway to small cell \bar{D}_{sb} , i.e., $\bar{D}_s = \bar{D}_{sr} + \bar{D}_{sb}$. If $\bar{D}(\theta)$ denotes the mean network packet delay from core network router to UE. Obviously, the packet delay of backhaul link \bar{D}_{sb} impacts the packet delay of small cell \bar{D}_s and network packet delay $\bar{D}(\theta)$ significantly.

Since the transmission process of packet is modeled as a tandem queueing system, the delay in a wireless network consists of transmission delay and queueing delay, that is $\bar{D}_m = \bar{D}_{Tm} + \bar{D}_{Qm}$ and $\bar{D}_s = \bar{D}_{Tsr} + \bar{D}_{Qsr} + \bar{D}_{Tsb} + \bar{D}_{Qsb}$.

The packet transmission delay can be expressed as $D_T = l/R$, the average packet size is l bits, R is the instantaneous transmission rate of a packet. Using the results in M/G/1 queueing model, the average queue length (i.e., average number of waiting packets in queue) is calculated to be $(\lambda^2\sigma^2 + \rho^2)/2(1-\rho)$ [28]. ρ represents the system utilization (or traffic intensity [16]) defined as the probability of BS being occupied in active state and $\rho = \lambda/\mu$. λ is the arrival rate of packet, and $\mu = R/l$ is the service rate of BS. σ^2 denotes the variance of service rate. In this paper, we work with the average packet transmission rate, hence the service rate of BS is a fixed value with a variance of 0, that is $\sigma^2 = 0$ [28]. Using Little's Law, the waiting queueing delay can be given by, $D_Q = \rho D_T / 2(1 - \rho)$ [28].

Based on the preceding analysis, we can obtain the packet delay for different cases.

3.2.1 Mean packet delay for UEs in a macro cell \bar{D}_m

(i) Mean transmission delay for UEs in a macro cell \bar{D}_{Tm} . Since we work with Poisson traffic model, discrete packets arrive in BSs according to PPP. $\bar{\xi}_m(\theta)$ is the average transmission probability of a macro cell in active mode. Besides, the packet transmission is successful in preventing a UE with low SIR from occupying too much resource, only if the SIR of UE is above a specific threshold β [20].

The SIR of a typical UE is a random variable, the complementary cumulative distribution function (CCDF) of SIR for UE is given by

$$\Pr_{th}(SIR > \eta) = (1 + \bar{\xi}_m(\theta)Z(\eta))^{-1}, \quad (4)$$

and

$$Z(\eta) = \eta^{\alpha_0} \int_{\eta^{-\alpha_0}}^{\infty} (1 + u^{\alpha_0})^{-1} du, \quad \alpha_0 = 2\alpha^{-1}, \quad (5)$$

where η is a random variable. α is the path loss exponent for all links [20, 21]. Therefore, if SIR of a typical UE is above the threshold β , the CCDF of SIR for UE is formulated as

$$\Pr_{Ast}(SIR > \eta) = (1 + \bar{\xi}_m(\theta)Z(\beta))(1 + \bar{\xi}_m(\theta)Z(\eta))^{-1}, \quad \eta \geq \beta. \quad (6)$$

Furthermore, we can get \bar{D}_{Tm}

$$\begin{aligned} \bar{D}_{Tm} &= \int_0^{\infty} \Pr(T > t) dt \\ &= \int_0^{\frac{l}{W_m \log(1+\beta)}} \Pr_{Ast} \left(\frac{l}{W_m \log(1+\beta)} > t \right) dt \\ &= \int_0^{\frac{l}{W_m \log(1+\beta)}} \Pr_{Ast} \left(SIR < \exp \left(\frac{l}{W_m t} \right) - 1 \right) dt \end{aligned}$$

$$= \int_0^{\frac{l}{W_m \log(1+\beta)}} \left(1 - \frac{1 + \bar{\xi}_m(\theta)Z(\beta)}{1 + \bar{\xi}_m(\theta)Z(\exp(\frac{l}{W_m t}) - 1)} \right) dt. \quad (7)$$

Since each BS is modeled as the M/G/1 queueing system, $\bar{\xi}_m(\theta)$ is equal to the utilization of M/G/1 queueing system ρ according to the definition [27], that is

$$\bar{\xi}_m(\theta) = E(\lambda_{mu}E(D_{Tm})) = E(\lambda_{mu})\bar{D}_{Tm}, \quad (8)$$

where $E(\lambda_{mu})$ denotes the average arrival rate of packet at each macro cell and can be expressed as

$$E(\lambda_{mu}) = \lambda\lambda_u\lambda_m^{-1}(1 - \text{Pr}_{\text{SUE}}(\theta))\text{Pr}_{\text{th}}(\text{SIR} > \beta), \quad (9)$$

where λ_u and λ_m are the densities of UEs and macro cell, λ denotes the arrival rate of packets for each UE, $\text{Pr}_{\text{SUE}}(\theta)$ is given by (1).

Substituting \bar{D}_{Tm} given by (7) and $E(\lambda_{mu})$ given by (9) into (8), we get

$$\frac{\lambda_m}{\lambda\lambda_u}(1 - \text{Pr}_{\text{SUE}}(\theta))^{-1} = (1 + \bar{\xi}_m(\theta)Z(\beta))^{-1} \int_0^{\frac{l}{W_m \log(1+\beta)}} \frac{y_m - Z(\beta)}{1 + \bar{\xi}_m(\theta)y_m} dt, \quad (10)$$

and

$$y_m = Z(\exp(l(W_m t)^{-1}) - 1), \quad (11)$$

where λ_u and λ_m are the densities of UEs and macro cell, respectively. λ is the arrival rate of packets for each UE and W_m is the bandwidth of the macro cell and l is the packet size.

Although closed-form expression of $\bar{\xi}_m(\theta)$ is difficult to derive due to complicated integral operation in (10), we can get $\bar{\xi}_m(\theta)$ using the bisection method in region $(0, 1)$ [29].

(ii) Mean queueing delay for UEs in macro cell \bar{D}_{Qm} is given by

$$\bar{D}_{Qm} = \frac{1}{2}\bar{\xi}_m(\theta)\bar{D}_{Tm}(1 - \bar{\xi}_m(\theta))^{-1}. \quad (12)$$

(iii) Mean packet delay for UEs in macro cell \bar{D}_m is given by

$$\bar{D}_m = \bar{D}_{Tm} + \bar{D}_{Qm}. \quad (13)$$

3.2.2 Mean packet delay for UEs in small cell \bar{D}_s

(i) Mean packet delay of radio access network for UEs in small cell \bar{D}_{sr} . \bar{D}_{sr} is the same as that of macro cell. Mean transmission delay of radio access network for UEs in a small cell \bar{D}_{Tsr} is

$$\bar{D}_{\text{Tsr}} = \int_0^{\frac{l}{W_s \log(1+\beta)}} \left(1 - \frac{1 + \bar{\xi}_s(\theta)Z(\beta)}{1 + \bar{\xi}_s(\theta)Z(\exp(\frac{l}{W_s t}) - 1)} \right) dt, \quad (14)$$

$\bar{\xi}_s(\theta)$ satisfies the following condition:

$$\frac{(1 - \theta)\lambda_s}{\lambda\lambda_u}\text{Pr}_{\text{SUE}}(\theta)^{-1} = (1 + \bar{\xi}_s(\theta)Z(\beta))^{-1} \int_0^{\frac{l}{W_s \log(1+\beta)}} \frac{y_s - Z(\beta)}{1 + \bar{\xi}_s(\theta)y_s} dt, \quad (15)$$

where

$$y_s = Z(\exp(l(W_s t)^{-1}) - 1), \quad (16)$$

λ_s is the density of small cells. W_s denotes the bandwidth of the small cell.

Mean queueing delay of radio access network for UEs in small cell \bar{D}_{Qsr} is given by

$$\bar{D}_{\text{Qsr}} = \frac{1}{2}\bar{\xi}_s(\theta)\bar{D}_{\text{Tsr}}(1 - \bar{\xi}_s(\theta))^{-1}. \quad (17)$$

Thus, \bar{D}_{sr} is given as

$$\bar{D}_{\text{sr}} = \bar{D}_{\text{Tsr}} + \bar{D}_{\text{Qsr}}. \quad (18)$$

(ii) Mean packet delay of backhauling network for UEs in small cell \bar{D}_{sb} . \bar{D}_{sb} is similar to \bar{D}_{sr} .

Since we assume that the gateways are always in active mode with an average transmission probability of 1, in wireless backhaul links from gateway to small cells, the CCDF of SIR is given by

$$\Pr_g(\text{SIR} > \beta) = (1 + Z(\beta))^{-1}. \quad (19)$$

The average transmission delay of wireless-backhauling network for UEs in small cell \bar{D}_{Tsb} is

$$\bar{D}_{Tsb} = \int_0^{\frac{l}{W_b \log(1+\beta)}} \left(1 - \frac{1 + Z(\beta)}{1 + Z(\exp(\frac{l}{W_b t}) - 1)} \right) dt, \quad (20)$$

where, W_b is the bandwidth of gateway. The utilization of each gateway in backhaul link can be expressed as follows:

$$\mu_g(\theta) = \lambda \lambda_u \lambda_g^{-1} \Pr_{\text{SUE}}(\theta) \Pr_g(\text{SIR} > \beta) \bar{D}_{Tsb}, \quad (21)$$

where, λ_g is density of gateway. $\Pr_{\text{SUE}}(\theta)$ is given by (1).

Moreover, the mean queueing delay of wireless-backhauling network for UEs in a small cell is

$$\bar{D}_{Qsb} = \frac{1}{2} \mu_g(\theta) \bar{D}_{Tsb} (1 - \mu_g(\theta))^{-1}. \quad (22)$$

Thus, \bar{D}_{sb} is given as

$$\bar{D}_{sb} = \lambda_s \lambda_g^{-1} (1 - \theta) (\bar{D}_{Tsb} + \bar{D}_{Qsb}). \quad (23)$$

(iii) Mean packet delay for UEs in small cell \bar{D}_s is

$$\bar{D}_s = \bar{D}_{sr} + \bar{D}_{sb}. \quad (24)$$

3.2.3 Mean network packet delay for UEs in the system $\bar{D}(\theta)$

Considering the probability that a typical UE associates with either a small cell or a macro cell, $\bar{D}(\theta)$ is given as

$$\bar{D}(\theta) = \Pr_{\text{SUE}}(\theta) \bar{D}_s + (1 - \Pr_{\text{SUE}}(\theta)) \bar{D}_m. \quad (25)$$

Although it is difficult to analyze the impact of sleeping ratio θ on mean network packet delay $\bar{D}(\theta)$, the simulation results show that $\bar{D}(\theta)$ is not a monotonic function of θ .

4 EDT optimization problem

— *에너지 절약과 지연 성능 향상*

The objective of EDT optimization problem is to find the optimal set of sleeping small cells that improve energy saving while ensuring a good delay performance. To solve the EDT optimization problem, we first analyze the optimal sleeping ratio of small cells. Then, the optimal sleeping ratio is used to find the set of sleeping small cells.

4.1 EDT problem formulation

Since the BS sleeping strategy has a significant impact on energy consumption and packet delay, both energy consumption and packet delay for radio access and backhaul link of wireless-backhauling UDN should be considered while finding the optimal set of sleeping small cells for EDT. To control the EDT in UDN more flexibly, as seen in the existing studies [14, 16], we define the cost function as

$$F(\theta) = \bar{P}(\theta) + \omega \bar{D}(\theta), \quad (26)$$

where $\bar{P}(\theta)$ is the average system energy consumption, $\bar{D}(\theta)$ is the mean network packet delay of UEs, and ω denotes the positive weighting factor. When ω is zero, we focus only on the energy saving, however, as ω grows, more emphasis is placed on the delay performance.

We aim to find the optimal set of sleeping small cells to minimize the cost function with ω as the weighting factor. Therefore, the EDT problem can be formulated as

$$\arg \min_{\mathcal{S}^*} F(\theta) = \bar{P}(\theta) + \omega \bar{D}(\theta), \quad (27)$$

$$\text{s.t.} \quad \begin{cases} 0 < \theta < 1, \\ 0 < \bar{\xi}_s(\theta) < 1, \\ 0 < \bar{\xi}_m(\theta) < 1, \end{cases} \quad (28)$$

where \mathcal{S} denotes the state set of small cells, the optimal state set of small cells \mathcal{S}^* corresponds to the optimal set of sleeping small cells $\mathcal{S}_{\text{off}}^*$ ($\mathcal{S}_{\text{off}}^* \subseteq \mathcal{S}^*$) constituted of the sleeping small cell k and $s_k = 0$. $\theta = N_{\text{off}}/N_s = \sum_{k=1}^N s_k/N_s$. $N_s = \lambda_s |\mathcal{A}|$ represents the number of small cells in the system and N_{off} is the number of sleeping small cells in the system. $\bar{\xi}_s(\theta)$ and $\bar{\xi}_m(\theta)$ are the average transmission probabilities of the small cell and macro cell respectively.

4.2 Solution for EDT optimization problem

The objective of the EDT problem is to find the optimal set of sleeping small cells to minimize the cost function. This problem is a challenging combinatorial optimization problem and the optimal solution can be found by conducting exhaustive study of $O(2^{N_s})$ possible cases.

To reduce the complexity in the process of finding the optimal set of sleeping small cells, EDT optimization problem is solved in two steps. First, a locally optimal sleeping ratio is derived to minimize the cost function using a dynamic gradient algorithm and then we prove that it can converge to global optimal sleeping ratio. Now, two sleeping strategies with low complexities are proposed to determine the suboptimal set of sleeping small cells according to the optimal sleeping ratio.

4.2.1 First step: find optimal sleeping ratio θ^*

Based on (28), we can get the feasible region $(\theta_{\min}, \theta_{\max})$ for θ :

$$\theta_{\min} = \min \left\{ 1 - \frac{\lambda \lambda_u}{\lambda_s} X_1 + \frac{\lambda_m}{\lambda_s} \left(\frac{P_{\text{mt}}}{A_b P_{\text{st}}} \right)^{\frac{2}{\alpha}}, 1 - \lambda_s^{-1} (A_b P_{\text{st}})^{-\frac{2}{\alpha}} \left(X_2 - \lambda_m P_{\text{mt}}^{\frac{2}{\alpha}} \right) \right\}, \quad (29)$$

$$\theta_{\max} = \max \left\{ 1 - \frac{\lambda \lambda_u}{\lambda_s} Y_1 + \frac{\lambda_m}{\lambda_s} \left(\frac{P_{\text{mt}}}{A_b P_{\text{st}}} \right)^{\frac{2}{\alpha}}, 1 - \lambda_s^{-1} (A_b P_{\text{st}})^{-\frac{2}{\alpha}} \left(Y_2 - \lambda_m P_{\text{mt}}^{\frac{2}{\alpha}} \right) \right\}, \quad (30)$$

where X_1 , X_2 , Y_1 , Y_2 , and detailed derivations are given in Appendix A. The cost function is approximately convex function of sleeping ratio. For proof, please refer to Appendix B.

A dynamic gradient descent iterative algorithm is used to obtain the local optimal sleeping ratio θ^* [29, 30]. Since the cost function is approximately convex function, the iteration method can converge it to the global optimal sleeping ratio in the feasible region [30].

The gradient function of the cost function is

$$\begin{aligned} \frac{\partial F(\theta)}{\partial \theta} &= |\mathcal{A}| \lambda_m \Delta p_m P_{\text{mt}} \frac{\partial \xi_m(\theta)}{\partial \theta} + (1 - \theta) |\mathcal{A}| \lambda_s \Delta p_s P_{\text{st}} \frac{\partial \xi_s(\theta)}{\partial \theta} \\ &\quad - |\mathcal{A}| \lambda_s (\bar{P}_{s0} - \bar{P}_S + \xi_s(\theta) \Delta p_s P_{\text{st}}) + \omega \frac{\partial \bar{D}(\theta)}{\partial \theta}, \end{aligned} \quad (31)$$

$\xi_s(\theta)$ and $\xi_m(\theta)$ are the mappings from θ to $\bar{\xi}_s(\theta)$ and $\bar{\xi}_m(\theta)$, respectively. $\partial \bar{D}(\theta)/\partial \theta$, $\partial \xi_m(\theta)/\partial \theta$ and $\partial \xi_s(\theta)/\partial \theta$ are given in Appendix B.

Combined with the constraint conditions, θ^* can be given by the following dynamic gradient descent

iterative algorithm:

$$\theta_{n+1} = \begin{cases} \theta_{\min}, & \theta_n - \delta \frac{\partial F(\theta_n)}{\partial \theta_n} < \theta_{\min}, \\ \theta_n - \delta \frac{\partial F(\theta_n)}{\partial \theta_n}, & \theta_{\min} < \theta_n - \delta \frac{\partial F(\theta_n)}{\partial \theta_n} < \theta_{\max}, \\ \theta_{\max}, & \theta_n - \delta \frac{\partial F(\theta_n)}{\partial \theta_n} > \theta_{\max}, \end{cases} \quad (32)$$

where δ is a sufficiently small step size.

Proof. Please see Ref. [29].

4.2.2 Second step: BS sleeping strategy

In the first step, the optimal sleeping ratio θ^* is derived to minimize the cost function from the point of view of EDT. In the second step, θ^* is converted into the number of sleeping small cells N_{off} ,

$$N_{\text{off}} = [\theta^* N_s] = [\theta^* \lambda_s |\mathcal{A}|], \quad (33)$$

where λ_s denotes the small cell density and $|\mathcal{A}|$ is the area of the entire network. Two sleeping strategies are proposed to find N_{off} small cells and switch them to the sleeping mode based on θ^* , thus, we can obtain the suboptimal set of sleeping small cells $\mathcal{S}_{\text{off}}^*$ for the EDT optimization problem.

Algorithm 1: queue-aware sleeping strategy. BS sleeping strategy is designed to switch off some BSs for energy saving. Traffic-aware sleeping strategies are widely researched in existing studies. The simplest traffic-aware sleeping strategy is a random sleeping policy that switches off each BS independently when there is no active UE in it [7, 13, 14, 16].

As one of the simplest traffic-aware sleeping strategies, queue-aware sleeping strategy is proposed based on the average queue length of small cells. That is, the small cell with the shortest average queue length is turned off to save energy. Queue length, in this paper, is defined as the average number of packets waiting for transmission in the buffer of BS. Long queue length means heavy traffic load.

The queue length for small cell BS k at the t -th time slot is given by

$$Q_k(t) = \max\{0, Q_k(t-1) - R_k(t)\Delta t + A_k(t)\}, \quad (34)$$

where $R_k(t)$ denotes the average transmission rate of a small cell k at the t -th time slot, $A_k(t)$ denotes the number of packets arriving in the small cell k at the t -th time slot, and Δt is the time interval.

For queue-aware sleeping strategy, the optimal sleeping ratio is converted into the number of sleeping small cells N_{off} according to (33). Then, we switch off the first N_{off} small cells with the smallest mean queue length. A more detailed queue-aware sleeping strategy is presented in Algorithm 1.

Algorithm 2: channel-queue-aware sleeping strategy. Since the average queue length depends on not only the packet arrival rate, but also the BS service rate, a channel-queue-aware sleeping strategy with consideration for both channel state and traffic load is introduced, as shown in Algorithm 2. For a channel-queue-aware sleeping policy, we choose the small cell with the smallest product of mean queue length and average transmission rate for a small cell as the sleeping small cell. As in queue-aware sleeping strategy, the optimal sleeping ratio is converted into the number of sleeping small cells N_{off} , according to (33). Now, the first N_{off} small cells with the smallest product of mean queue length and average transmission rates are turned off. A more detailed channel-queue-aware sleeping strategy is given in Algorithm 2.

Complexity analysis. In Algorithm 1, to calculate the average transmission rate and queue length, TN_s calculations are required in steps 9–13. Now, the suboptimal set of sleeping small cells can be directly obtained through $N_{\text{off}}N_s$ comparisons. The complexity of Algorithm 2 is the same as that of Algorithm 1.

이거 insight 얻은 것 같아 좋게 해

Algorithm 1 Queue-aware sleeping strategy

Input: SBS set \mathcal{B}_S , MBS set \mathcal{B}_M , UE set \mathcal{U} , θ^* , T , Δt , $\forall i \in \mathcal{U}$, $\forall j \in \{\mathcal{B}_S, \mathcal{B}_M\}$, $\forall k \in \{\mathcal{B}_S\}$.

Output: Optimal state set of SBS \mathcal{S}^* .

```

1: Initialize: all MBSs and SBSs are active,  $\mathcal{S} = (1, 1, 1, \dots, 1)$  and  $n = 1$ ;
2: Calculate  $N_{\text{off}}$  according to (33);
3: Select the serving BS  $j^* = \arg \max_{j \in \{\mathcal{B}_S, \mathcal{B}_M\}} \{\text{RSRP}_j\}$  for each UE  $i$ ;
4: Find the set of UEs  $\mathcal{U}_j$  that can be served by each BS  $j$ ;
5: for each  $t \in [1, T]$  do
6:   Calculate transmission rate  $R_k(t)$ , update queue length according to (34), for small cell BS  $k$ ;
7: end for
8: Calculate  $\bar{Q}_k = \frac{\sum_{t=1}^T Q_k(t)}{T}$  for each small cell BS  $k$ ;
9: while  $n \leq N_{\text{off}}$  then
10:   for each small cell BS  $k$  do
11:     if  $\mathcal{S}(1, k) = 1$  and  $k = \min_{k \in \mathcal{B}_S} \{\bar{Q}_k\}$ ;
12:        $\mathcal{S}(1, k) = 0$ , assign UEs in  $\mathcal{U}_k$  to neighboring BSs;
13:     end if
14:   end for
15:    $n = n + 1$ ;
16: end while

```

Algorithm 2 Channel-queue-aware sleeping strategy

Input: SBS set \mathcal{B}_S , MBS set \mathcal{B}_M , UE set \mathcal{U} , θ^* , T , Δt , $\forall i \in \mathcal{U}$, $\forall j \in \{\mathcal{B}_S, \mathcal{B}_M\}$, $\forall k \in \{\mathcal{B}_S\}$.

Output: Optimal state set of SBS \mathcal{S}^* .

```

1: Initialize: all MBSs and SBSs are active,  $\mathcal{S} = (1, 1, 1, \dots, 1)$  and  $n = 1$ ;
2: Calculate  $N_{\text{off}}$  according to (33);
3: Select the serving BS  $j^* = \arg \max_{j \in \{\mathcal{B}_S, \mathcal{B}_M\}} \{\text{RSRP}_j\}$  for each UE  $i$ ;
4: Find the set of UEs  $\mathcal{U}_j$  that can be served by each BS  $j$ ;
5: for each  $t \in [1, T]$  do
6:   Calculate transmission rate  $R_k(t)$ , update queue length according to (34), for small cell BS  $k$ ;
7: end for
8: Calculate  $\bar{Q}_k = \frac{\sum_{t=1}^T Q_k(t)}{T}$  and  $\bar{R}_k = \frac{\sum_{t=1}^T R_k(t)}{T}$  for each small cell BS  $k$ ;
9: while  $n \leq N_{\text{off}}$  then
10:   for each small cell BS  $k$  do
11:     if  $\mathcal{S}(1, k) = 1$  and  $k = \min_{k \in \mathcal{B}_S} \{\bar{Q}_k \bar{R}_k\}$ ;
12:        $\mathcal{S}(1, k) = 0$ , assign UEs in  $\mathcal{U}_k$  to neighboring BSs;
13:     end if
14:   end for
15:    $n = n + 1$ ;
16: end while

```

Table 1 System parameters

Parameter	Value	Parameter	Value	Parameter	Value	Parameter	Value
λ_g	5×10^{-6}	P_{s0}	4.8 W	W_b	20 MHz	λ	0.5 s^{-1}
λ_m	1×10^{-5}	P_{m0}	10 W	W_m	10 MHz	Δp_m	10
λ_s	5×10^{-5}	P_S	2.4 W	W_s	10 MHz	Δp_s	8
λ_u	2×10^{-4}	P_G	100 W	l	0.1 MB	β	5

5 Numerical and simulation results

In this section, numerical and simulation results are provided to validate the proposed schemes. System parameters are consistent with those in [31]. The area of the system is $250000\pi \text{ m}^2$. Other system parameters are listed in Table 1. For each configuration, the system is simulated for 10000 time slots.

5.1 Mean packet delay

The simulation and numerical results for mean packet delay vs. sleeping ratio θ are presented in Figure 2, the numerical results match the simulation results very well. From Figure 2 we can see that the mean network packet delay $\bar{D}(\theta)$ first decreases and then increases with the increase of θ . This is due to the fact that the presence of a large numbers of active small cells can lead to serious interference, thus turning

parameter table의 의미가 생각됨
보여줌

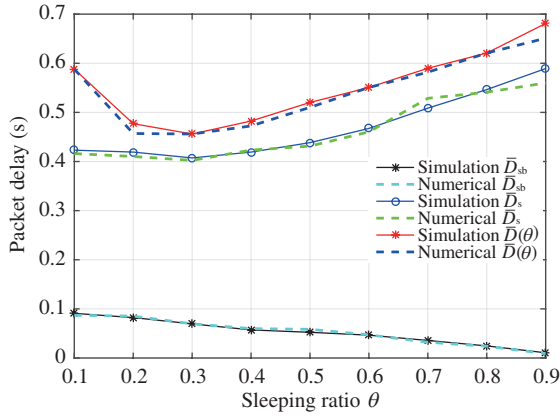


Figure 2 (Color online) Simulation and numerical results for mean packet delay vs. sleeping ratio θ .

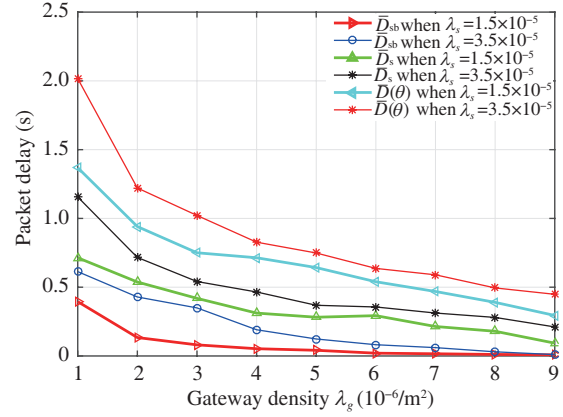


Figure 3 (Color online) Numerical results for mean packet delay vs. gateway density λ_g .

off some BSs may lead to decrease of delay [29]. Meanwhile, when a large number of BSs are turned off, the increase in delay is directly proportional to increase of θ due to the decreased probability for UE to associate with small cells. Since a small cell can provide high quality service as compared to the macro cell, as seen in Figure 2, \bar{D}_s is smaller than $\bar{D}(\theta)$. In addition to this, \bar{D}_{sb} decreases with increase of θ , this is because the increasing value of θ signifies the decreasing number of UEs in each small cell BS, which decreases the traffic load in each gateway of the backhauling network.

Figure 3 represents the numerical results for mean packet delay vs gateway density λ_g . We can see that λ_g has as significant impact on the packet delay of the backhauling network \bar{D}_{sb} , the larger the value of λ_g , the smaller the value of \bar{D}_{sb} . Since \bar{D}_{sb} decreases with an increase in λ_g , \bar{D}_s and $\bar{D}(\theta)$ decrease as λ_g increases. On the other hand, a larger small cell density λ_s leads to an increase in the value of \bar{D}_{sb} due to the increased packet traffic in the backhaul link, and also \bar{D}_s and $\bar{D}(\theta)$ increase with an increase in λ_s .

5.2 System energy consumption

Figure 4 shows the impact of θ on system energy consumption. With the increase of θ , the energy consumption decreases due to the decreased number of active BSs. Compared to the scheme without considering the sleeping strategy when $\theta = 0$, less energy is consumed in UDN with the sleeping strategy. Besides, larger λ_g leads to a larger number of gateways in system, thus energy consumption increases with an increase in the value of λ_g .

5.3 Energy delay tradeoff

System energy consumption vs. mean network packet delay for different θ is given in Figure 5. It is obvious that an increase in θ leads to the decrease in energy consumption. On the contrary, the relationship between mean network packet delay and θ is not monotonic, as shown in Figures 2 and 3. Thus, we can say that the relationship between energy consumption and mean network packet delay deviates from monotonic curve, which means that the sacrificing delay cannot always be energy saving in return.

Figure 6 represents the numerical results for cost function $F(\theta)$ of EDT vs. θ for different ω . We can see that $F(\theta)$ is approximately convex for different θ , which means that the dynamic gradient descent iterative algorithm can converge to the global optimal sleeping ratio θ^* . As shown in Figure 6, the values of $F(\theta)$ for different ω are different, therefore value of θ^* that can minimize the cost function is different.

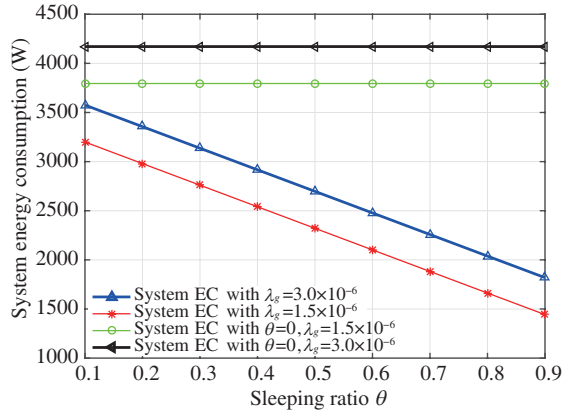


Figure 4 (Color online) Numerical results for system energy consumption vs. sleeping ratio θ .

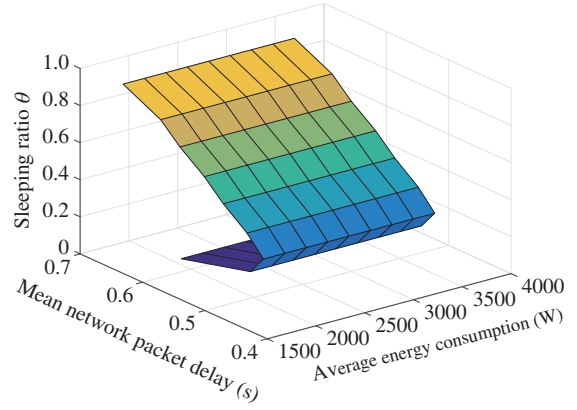


Figure 5 (Color online) Energy consumption vs. mean network packet delay for different θ .

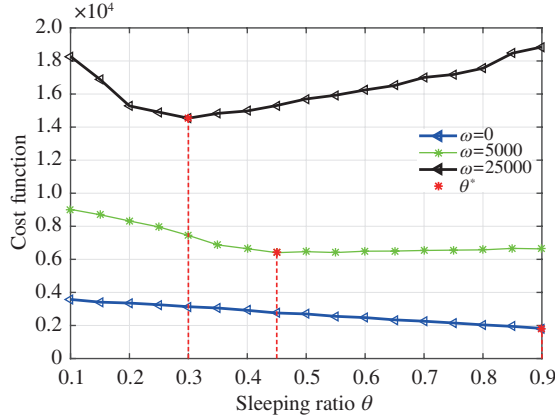


Figure 6 (Color online) Numerical results for cost function of EDT problem vs. BS sleeping ratio θ .

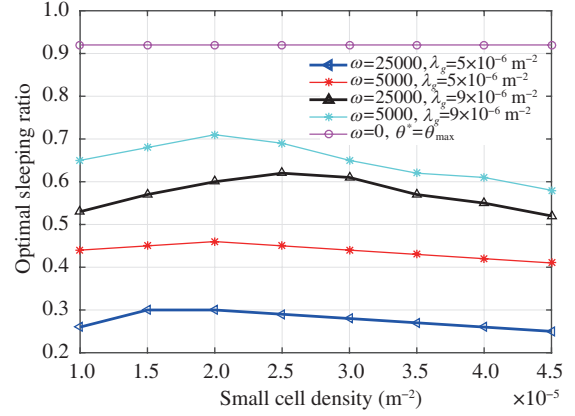


Figure 7 (Color online) Optimal sleeping ratio vs. small cell density for different weighting factor.

5.4 Operation for sleeping strategy

Figure 7 represents the optimal sleeping ratio θ^* of EDT problem for different small cell density obtained by the dynamic gradient descent iterative algorithm. θ^* for $\omega = 0$, $\omega = 5000$ and $\omega = 25000$ when gateway density $\lambda_g = 5 \times 10^{-6}$ is 0.91, 0.45 and 0.3, respectively, which are in accordance with the results in Figure 6. Besides, the larger the value of λ_g , the larger will be the value of θ^* owing to the fact that the increasing number of gateways reduces \bar{D}_{sb} and $\bar{D}(\theta)$, thus more small cells will be switched off to save energy.

The objective of EDT problem is to find the optimal set of sleeping small cells to minimize the cost function. To solve the EDT problem, the dynamic gradient descent iterative algorithm is used in first step to find the optimal sleeping ratio for cost function as shown in Figure 7, and then queue-aware and channel-queue-aware sleeping strategy are proposed to select the small cells to be turned off. The suboptimal state set of small cells corresponding to the suboptimal set of sleeping small cells under the optimal sleeping ratio for Algorithms 1 and 2 are presented in Figure 8. It can be seen that the set of sleeping small cells are different for the two different sleeping strategies.

5.5 System performance of proposed sleeping strategy

System energy consumption vs. small cell density λ_s for different sleeping schemes based on the optimal sleeping ratio θ^* is presented in Figure 9. The system energy consumption increases with the increase

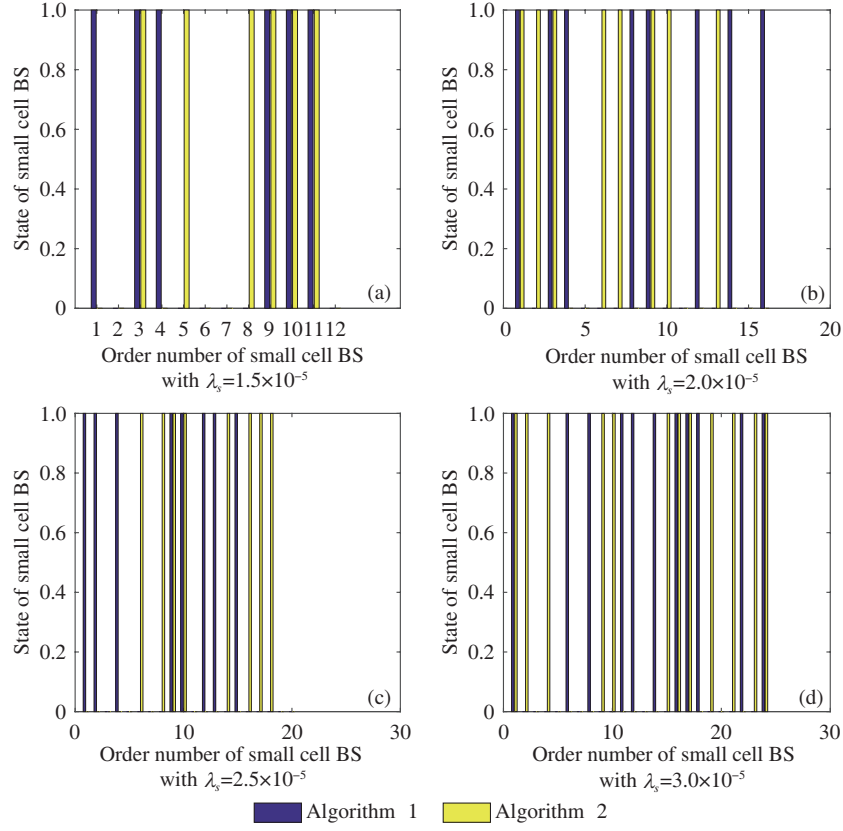


Figure 8 (Color online) Optimal state set of small cells for different small cell density. (a) $\lambda_s = 1.5 \times 10^{-5}$; (b) $\lambda_s = 2.0 \times 10^{-5}$; (c) $\lambda_s = 2.5 \times 10^{-5}$; (d) $\lambda_s = 3.0 \times 10^{-5}$.

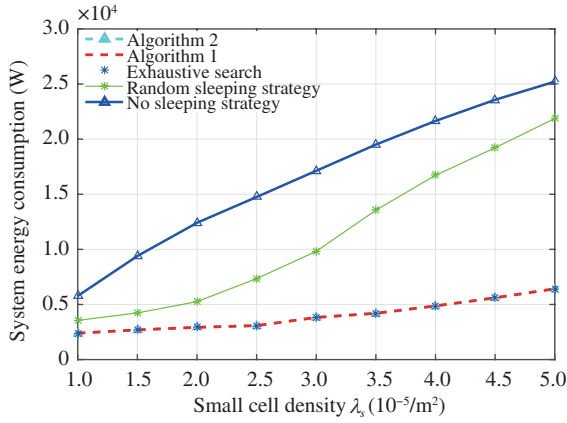


Figure 9 (Color online) System energy consumption vs. λ_s with θ^* for different sleeping schemes.

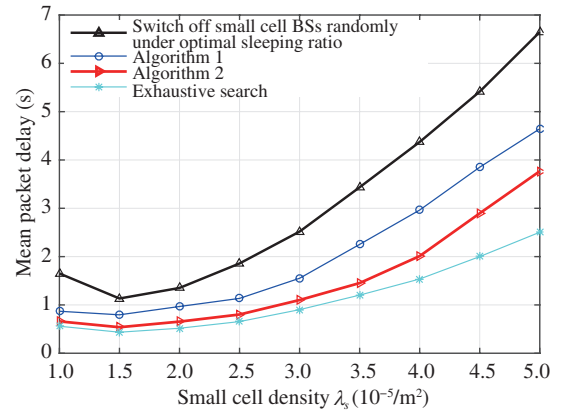


Figure 10 (Color online) Mean delay vs. small cell density with the optimal sleeping ratio for different sleeping schemes.

in the value of λ_s although the sleeping strategy is used. From Figure 9, it can be seen that the energy consumption for Algorithm 1 is the same as that of Algorithm 2 and exhaustive search due to the same sleeping ratio under our model assumption. Compared with the scheme without considering the sleeping strategy, random sleeping strategy brings better performance for energy conservation. Moreover, simulation results indicate that energy saving can be further improved significantly by our proposed queue-aware sleeping and channel-queue-aware sleeping scheme.

Figure 10 gives the comparisons of simulation results of mean network packet delay vs. small cell

density λ_s under the optimal sleeping ratio θ^* for different sleeping schemes. From this figure, it can be observed that the mean network packet delay first slightly decreases and then increase with the increase of λ_s . The decreasing delay due to the increasing BS density can provide more opportunities for UEs to associate with small cells, the increasing delay is caused by an increasing interference of active BSs even though the sleeping strategy is employed [29]. As benchmarks, exhaustive search and a sleeping strategy that selects the sleeping small cells randomly with θ^* are presented. Benefitting from information of queue length, mean network packet delay can be improved by Algorithm 1. Since the channel state information and queue length are both considered, Algorithm 2 has a better delay performance with low complexity.

6 Conclusion

In this paper, EDT problem for two-tier UDN with wireless backhaul is studied. A mathematical model is introduced to analyze system energy consumption and mean network packet delay. Then, the EDT problem is formulated as a cost minimization problem to find the set of sleeping small cells. A dynamic gradient iteration algorithm is used to achieve a locally optimal solution of sleeping ratio and we prove that it can converge to a global optimal sleeping ratio. Furthermore, queue-aware and channel-queue-aware sleeping strategies are proposed to select the set of sleeping small cells according to the optimal sleeping ratio. Simulation and numerical results confirm the effectiveness of the schemes.

Acknowledgements This work was partially supported by National Major Project (Grant No. 2017ZX03001002-004), National Natural Science Foundation Project (Grant No. 61521061), 333 Program of Jiangsu (Grant No. BRA2017366), and Huawei Technologies Co., Ltd.

References

- 1 You X H, Pan Z W, Gao X Q, et al. The 5G mobile communication: the development trends and its emerging key techniques (in Chinese). *Sci Sin Inform*, 2014, 44: 551–563
- 2 Ismail M, Zhuang W H. Network cooperation for energy saving in green radio communications. *IEEE Wirel Commun*, 2011, 18: 76–81
- 3 Wu J J, Zhang Y J, Zukerman M, et al. Energy-efficient base-stations sleep-mode techniques in green cellular networks: a survey. *IEEE Commun Surv Tut*, 2015, 17: 803–826
- 4 Ge X H, Cheng H, Guizani M, et al. 5G wireless backhaul networks: challenges and research advances. *IEEE Netw*, 2014, 28: 6–11
- 5 Suarez L, Bouraoui M A, Mertah M A, et al. Energy efficiency and cost issues in backhaul architectures for high data-rate green mobile heterogeneous networks. In: *Proceedings of the 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Hong Kong, 2015. 1563–1568
- 6 Tombaz S, Monti P, Farias F, et al. Is backhaul becoming a bottleneck for green wireless access networks? In: *Proceedings of International Conference on Communications (ICC)*, Sydney, 2014. 4029–4035
- 7 Chang P L, Miao G W. Joint optimization of base station deep-sleep and DTX micro-sleep. In: *Proceedings of IEEE Global Communication Conference Workshops (GlobeCom Workshops)*, Washington, 2017
- 8 Ebrahim A, Alsusa E. Interference and resource management through sleep mode selection in heterogeneous networks. *IEEE Trans Commun*, 2017, 65: 257–269
- 9 Li Z H, Grace D, Mitchell P. Traffic-aware cell management for green ultradense small-cell networks. *IEEE Trans Veh Technol*, 2016, 66: 2600–2614
- 10 Samarakoon S, Bennis M, Saad W, et al. Opportunistic sleep mode strategies in wireless small cell networks. In: *Proceedings of IEEE International Conference on Communications (ICC)*, Sydney, 2016. 2707–2712
- 11 Wu J, Liu J, Zhao H. Dynamic small cell on/off control for green ultra-dense networks. In: *Proceedings of IEEE International Conference on Wireless Communications and Signal Processing (WCSP)*, Yangzhou, 2016
- 12 Zhang Q, Yang C Y, Haas H, et al. Energy efficient downlink cooperative transmission with BS and antenna switching off. *IEEE Trans Wirel Commun*, 2014, 13: 5183–5195
- 13 Liu C, Natarajan B, Xia H X. Small cell base station sleep strategies for energy efficiency. *IEEE Trans Veh Technol*, 2016, 65: 1652–1661
- 14 Liu B, Zhao M, Zhou W Y, et al. Flow-level-delay constrained small cell sleeping with macro base station cooperation for energy saving in hetnet. In: *Proceedings of IEEE Vehicular Technology Conference (VTC-Fall)*, Boston, 2015
- 15 Son K, Kim H, Yi Y, et al. Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks. *IEEE J Sel Areas Commun*, 2011, 29: 1525–1536
- 16 Li P, Jiagn H L, Pan Z W, et al. Energy-delay tradeoff in ultra-dense networks considering BS sleeping and cell association. *IEEE Trans Veh Technol*, 2018, 67: 734–751

- 17 Nie G F, Tian H, Ren C S. Energy efficient cell selection in small cell networks with constrained backhaul links. *IEEE Commun Lett*, 2016, 20: 1199–1202
- 18 Liu H, Zhang H J, Cheng J L, et al. Energy efficient power allocation and backhaul design in heterogeneous small cell networks. In: *Proceedings of IEEE International Conference on Communications (ICC)*, Kuala Lumpur, 2016. 22–27
- 19 Nie G F, Tian H, Sengul C, et al. Forward and backhaul link optimization for energy efficient OFDMA small cell networks. *IEEE Trans Wirel Commun*, 2016, 16: 1080–1093
- 20 Zhang G Z, Quek T, Huang A, et al. Backhaul-aware base station association in two-tier heterogeneous cellular networks. In: *Proceedings of the 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Stockholm, 2015. 390–394
- 21 Han T, Ansari N. User association in backhaul constrained small cell networks. In: *Proceedings of IEEE Wireless Communication and Networking Conference (WCNC)*, New Orleans, 2015. 1637–1642
- 22 Jamali V, Michalopoulos D S, Uysal M, et al. Link allocation for multiuser systems with hybrid RF/FSO backhaul: delay-limited and delay-tolerant designs. *IEEE Trans Wirel Commun*, 2016, 15: 3281–3295
- 23 Cui Z Y, Cui Q M, Zheng W, et al. Energy-delay analysis for partial spectrum sharing in heterogeneous cellular networks with wired backhaul. In: *Proceedings of IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Valencia, 2016. 1563–1568
- 24 Auer G, Giannini V, Desset C, et al. How much energy is needed to run a wireless network? *IEEE Trans Wirel Commun*, 2011, 18: 40–49
- 25 Jo H S, Sang Y J, Xia P, et al. Heterogeneous cellular networks with flexible cell association: a comprehensive downlink SINR analysis. *IEEE Trans Wirel Commun*, 2012, 11: 3484–3495
- 26 Haenggi M, Andrews J G, Baccelli F, et al. Stochastic geometry and random graphs for the analysis and design of wireless networks. *IEEE J Sel Areas Commun*, 2009, 27: 1029–1046
- 27 Li X, Ji H, Wang K, et al. Energy-efficient access scheme with joint consideration on backhualing in UDN. In: *Proceedings of IEEE Vehicular Technology Conference (VTC-Fall)*, Montreal, 2016
- 28 Takagi H. *Queueing Analysis: A Foundation of Performance Evaluation, Volume I: Vacation and Priority Systems*. 1st ed. Amsterdam: Elsevier, 1991
- 29 Li L, Peng M G, Yang C Q, et al. Optimization of base station density for high energy efficient cellular networks with sleeping strategies. *IEEE Trans Veh Technol*, 2016, 65: 7501–7514
- 30 Stephen B, Lieven V. *Convex Optimization*. 1st ed. Cambridge: Cambridge University Press, 2009
- 31 Jo H S, Xia P, Andrews J G. Open, closed, and shared access femtocells in the downlink. *J Wirel Commun Netw*, 2012, 2012: 363–378

Appendix A Analysis of the feasible region for θ

According to (10) and (15), we have

$$\begin{aligned}\theta &= 1 - \lambda_s^{-1} (A_b P_{st})^{-\frac{2}{\alpha}} \left(\frac{\lambda \lambda_u P_{mt}^{\frac{2}{\alpha}}}{1 + \bar{\xi}_m(\theta) Z(\beta)} \int_0^{\frac{l}{W_m \log(1+\beta)}} \frac{y_m - Z(\beta)}{1 + \bar{\xi}_m(\theta) y_m} dt - \lambda_m P_{mt}^{\frac{2}{\alpha}} \right) \\ &= 1 + \frac{\lambda_m}{\lambda_s} \left(\frac{P_{mt}}{A_b P_{mt}} \right)^{\frac{2}{\alpha}} - \frac{\lambda \lambda_u \int_0^{\frac{l}{W_s \log(1+\beta)}} \frac{y_s - Z(\beta)}{1 + \bar{\xi}_s(\theta) y_s} dt}{\lambda_s (1 + \bar{\xi}_s(\theta) Z(\beta))}.\end{aligned}\quad (A1)$$

We can observe that θ increases with the increase of $\bar{\xi}_m(\theta)$ and $\bar{\xi}_s(\theta)$, respectively, thus, $\bar{\xi}_m(\theta)$ and $\bar{\xi}_s(\theta)$ both are increasing function of θ . Substituting $\bar{\xi}_m(\theta) = 0, 1$ into (A1), and $\bar{\xi}_s(\theta) = 0, 1$ into (A1), (29) and (30) are derived, and where

$$X_1 = \int_0^{\frac{l}{W_s \log(1+\beta)}} (y_s - Z(\beta)) dt, \quad (A2)$$

$$X_2 = \lambda \lambda_u P_{mt}^{\frac{2}{\alpha}} \left(\int_0^{\frac{l}{W_m \log(1+\beta)}} (y_m - Z(\beta)) dt \right), \quad (A3)$$

$$Y_1 = (1 + Z(\beta))^{-1} \int_0^{\frac{l}{W_s \log(1+\beta)}} \frac{y_s - Z(\beta)}{1 + y_s} dt, \quad (A4)$$

$$Y_2 = \lambda \lambda_u P_{mt}^{\frac{2}{\alpha}} \left(\frac{1}{1 + Z(\beta)} \int_0^{\frac{l}{W_m \log(1+\beta)}} \frac{y_m - Z(\beta)}{1 + y_m} dt \right). \quad (A5)$$

Appendix B Analysis of cost function

$$Z_1 = \frac{1}{2} (1 - \text{Pr}_{\text{SUE}}(\theta)) \bar{D}_{\text{Tm}} \frac{1}{(1 - \bar{\xi}_m(\theta))^2} \frac{\partial \bar{\xi}_m(\theta)}{\partial \theta} - \frac{\partial \text{Pr}_{\text{SUE}}(\theta)}{\partial \theta} \left(1 + \frac{1}{2(1 - \bar{\xi}_m(\theta))} \right) \bar{D}_{\text{Tm}}, \quad (B1)$$

$$Z_2 = \frac{1}{2} \text{Pr}_{\text{SUE}}(\theta) \bar{D}_{\text{Tsr}} \frac{1}{(1 - \bar{\xi}_s(\theta))^2} \frac{\partial \bar{\xi}_s(\theta)}{\partial \theta} - \frac{\lambda_s}{\lambda_g} \text{Pr}_{\text{SUE}}(\theta) \bar{D}_{\text{Tsb}} \left(1 + \frac{1}{2(1 - \mu_g(\theta))} \right), \quad (B2)$$

$$Z_3 = \frac{\lambda_s}{\lambda_g} \bar{D}_{\text{Tsb}}(1-\theta) \frac{\partial \text{Pr}_{\text{SUE}}(\theta)}{\partial \theta} \left(1 + \frac{1}{2(1-\mu_g(\theta))}\right) + \frac{1}{2} \frac{\lambda_s}{\lambda_g} \text{Pr}_{\text{SUE}}(\theta) \bar{D}_{\text{Tsb}}(1-\theta) \frac{1}{(1-\mu_g(\theta))^2} \frac{\partial \mu_g(\theta)}{\partial \theta}, \quad (\text{B3})$$

$$Z_4 = \frac{\partial \text{Pr}_{\text{SUE}}(\theta)}{\partial \theta} \left(1 + \frac{1}{2(1-\xi_s(\theta))}\right) \bar{D}_{\text{TSr}}, \quad (\text{B4})$$

$$\frac{\partial \text{Pr}_{\text{SUE}}(\theta)}{\partial \theta} = - \frac{\lambda_m \lambda_s (A_b P_{\text{st}} P_{\text{mt}})^{\frac{2}{\alpha}}}{((1-\theta) \lambda_s (A_b P_{\text{st}})^{\frac{2}{\alpha}} + \lambda_m P_{\text{mt}})^{\frac{2}{\alpha}}}, \quad (\text{B5})$$

$$\begin{aligned} \frac{\partial \xi_m(\theta)}{\partial \theta} = & \frac{\lambda_s}{\lambda \lambda_u} \left(\frac{A_b P_{\text{st}}}{P_{\text{mt}}} \right)^{\frac{2}{\alpha}} (1 + \xi_m(\theta) Z(\beta))^2 \left(Z(\beta) \int_0^{\overline{W_m \log(1+\beta)}} q_m(\theta) dt \right. \\ & \left. + (1 + \xi_m(\theta) Z(\beta)) \int_0^{\overline{W_m \log(1+\beta)}} \frac{y_m q_m(\theta)}{1 + \xi_m(\theta) y_m} dt \right)^{-1}, \end{aligned} \quad (\text{B6})$$

$$\frac{\partial \xi_s(\theta)}{\partial \theta} = \frac{\lambda_s}{\lambda \lambda_u} (1 + \xi_s(\theta) Z(\beta))^2 \left(Z(\beta) \int_0^{\overline{W_s \log(1+\beta)}} q_s(\theta) dt + (1 + \xi_s(\theta) Z(\beta)) \int_0^{\overline{W_s \log(1+\beta)}} \frac{y_s q_s(\theta)}{1 + \xi_s(\theta) y_s} dt \right)^{-1}, \quad (\text{B7})$$

$$q_s(\theta) = \frac{y_s - Z(\beta)}{1 + \xi_s(\theta) y_s}, \quad q_m(\theta) = \frac{y_m - Z(\beta)}{1 + \xi_m(\theta) y_m}. \quad (\text{B8})$$

We can obtain $\partial D(\theta)/\partial \theta = Z_1 + Z_2 + Z_3 + Z_4$, it can be observed that Z_1 and Z_2 increase with the increase of sleeping ratio θ . And $\partial Z_3/\partial \theta > 0$, $\partial Z_4/\partial \theta > 0$, thus $\partial \bar{D}(\theta)/\partial \theta$ is an increasing function of θ . On the other hand, it can be proven that system energy consumption is a decreasing function of sleeping ratio [17]. Therefore, the cost function is approximately convex for sleeping ratio in the feasible region [30].