

Article

An Efficient Energy Saving Scheme for Base Stations in 5G Networks with Separated Data and Control Planes Using Particle Swarm Optimization

Min Wook Kang¹ and Yun Won Chung^{2,*}

¹ Department of Information and Telecommunication Engineering, Graduate School, Soongsil University, Seoul 06978, Korea; goodlookmw@gmail.com

² School of Electronic Engineering, Soongsil University, Seoul 06978, Korea

* Correspondence: ywchung@ssu.ac.kr; Tel.: +82-02-820-0908

Received: 5 July 2017; Accepted: 11 September 2017; Published: 15 September 2017

Abstract: Reducing energy consumption of mobile communication networks has gained significant attentions since it takes a major part of the total energy consumption of information and communication technology (ICT). In this paper, we consider 5G networks with heterogeneous macro cells and small cells, where data and control planes are separated. We consider two types of data traffic, i.e., low rate data traffic and high rate data traffic. In basic separation architecture, a macro cell base station (MBS) manages control signals, while a small cell base station (SBS) manages both low rate data traffic and high rate data traffic. In the considered modified separation architecture, an MBS manages control signals and low rate data traffic, while an SBS manages high rate data traffic. Then, an efficient energy saving scheme for base stations (BSs) is proposed, where the state of a BS is determined depending on the number of user equipments (UEs) that request high rate data traffic and the number of UEs that exist under the overlapping areas commonly covered by the considered BS and the neighbor BSs. We formulate an optimization problem for the proposed energy saving scheme and obtain the solution using particle swarm optimization (PSO). Numerical results show that the proposed energy saving scheme in the modified separated network architecture has better energy efficiency compared to the conventional energy saving schemes in both basic and modified separated network architectures. Also, the proposed energy saving scheme has lower aggregate delay.

Keywords: 5G networks; separated data and control planes; energy saving; heterogeneous network; particle swarm optimization (PSO)

1. Introduction

Ever increasing energy consumption is one of the important issues related to the problem of global warming [1] and reducing the energy consumption of mobile communication networks has gained significant attentions since it takes a major part of the total energy consumption of information and communication technology (ICT). In future, the effect of the energy consumption of the mobile communication network will be more serious since more traffic load will be expected in forthcoming 5G networks [2,3]. Reference [4] considers the base station (BS) to be the major source of energy consumption of mobile communication networks and the energy consumption of the BS depends on the traffic load which varies depending on the geographical location.

To reduce the energy consumption of mobile communication networks, significant work has been carried out regarding the energy saving of a BS. The basic principle of the work to reduce the energy consumption of a BS is to turn off the components of the BS as much as possible, when they are no longer needed [4]. For example, the BS can be put into sleep mode by shutting down the energy to most of its components when they are not being actively used. Figure 1 shows that, furthermore,

the BS can be turned off to save energy more significantly, if there is little traffic of a BS and the traffic of the BS can be covered by neighbor BSs. Since user equipments (UEs) covered by the center BS can be also covered by neighbor BSs with overlapping coverage, the center BS can be turned off to save energy and the UEs are served by neighbor BSs [5].

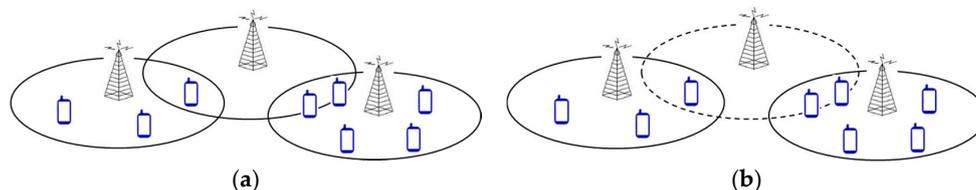


Figure 1. Activation and deactivation of base stations (BSs). (a) Activation; (b) deactivation.

The *on/off* state of BS is determined based on the context information such as the traffic demand and channel state in an environment where connections between the BS and UE are changing dynamically. In [6], greedy-on and greedy-off schemes were proposed, where the state of a BS is changed to either *on* or *off* state when a certain condition is met, and it was concluded that greedy-on scheme was more efficient in terms of energy efficiency. In [7], various ways of reducing the number of BSs with *on* state was introduced and a traffic-intensity-aware multi-cell cooperation scheme was introduced, where the state of a BS is changed to *off* state based on its traffic density, which is classified into peak hour traffic and off-peak traffic, depending on the traffic demand of the UE. Also, coverage hole is accommodated by neighbor BSs with *on* state.

In contrast to the activation and deactivation schemes of BSs [4–7], the cell zooming scheme was introduced in [8], where cells are zooming-in and zooming-out depending on the traffic load and channel condition of BSs by controlling the coverage area of BSs, as shown in Figure 2. In cell zooming, if a BS has a higher traffic load and is congested, the BS zooms in, while neighbor BSs zoom out to avoid any possible coverage hole. On the other hand, if a BS has a lower traffic load and is not congested, the BS zooms out, while neighbor BSs zoom in. In cell zooming, the zoomed-in BS can sleep in order to save energy and then neighbor BSs should cover the area that was previously covered by the sleeping BS. In the cell zooming scheme, the main objective is to balance the traffic load and increase the energy efficiency of the total network by reducing the energy consumption of BSs. Therefore, BSs with *on* state should accommodate the coverage hole and cooperate with nearby BSs to deal with the traffic of the turned off BSs [9]. In [10], as an extension of the original cell zooming scheme, the authors proposed an efficient energy saving scheme for BSs, where BSs are turned off when traffic load is very low and active BSs extend the cell coverage by increasing the transmission power. In [11], the authors proposed three cell zooming schemes, i.e., continuous, discrete, and fuzzy schemes. In continuous cell zooming scheme, a BS controls transmit power to cover its farthest user. In discrete cell zooming scheme, only discrete values of transmit power are allowed. Finally, in fuzzy cell zooming scheme, as an extension of the discrete scheme, a BS transmit power is increased with a small amount from each discrete level. In [12], the authors proposed an energy-saving small cell zooming scheme in a two-tier cellular network, where small cell BSs dynamically select turning on, turning off, zooming in, and zooming out based on varying traffic load, the speed of UE, and the location of UE, while satisfying data rate requirements of UEs.

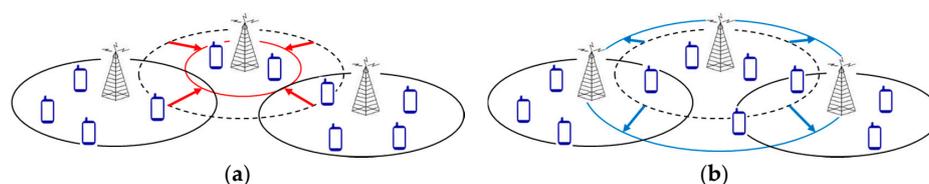


Figure 2. Operation of cell zooming: (a) zooming-in; (b) zooming-out.

In order to save the energy of BSs and efficiently serve UEs, work on state management has recently drawn a lot of attentions [13–15]. Although turning off BSs can save a significant amount of BSs' energy, it introduces long delay to respond to the users' traffic demand, since BSs should activate first from the *off* state to serve UEs and this takes a relatively longer time. In order to solve the long delay problem of a deactivated BS, the BS does not turn off all the components but turns off most of the components in order to save energy and quickly reacts to the users' demand by entering into *sleep* state instead of entering into *off* state. In [13], the authors classified the state of a BS as *on*, *standby*, *sleep*, and *off* states based on the activity and energy consumption of the BS. In the *on* state, a BS consumes the largest power since all of the operations of a BS are active. In *standby* state, the BS can move to *on* state very rapidly when it is needed, since only the radio frequency (RF) and temperature compensated crystal oscillators (TCXO) heater of a BS are inactive. A BS in *standby* state consumes less power than that in the *on* state but it consumes more power than that in the *sleep* or *off* state. In the *sleep* state, which is very similar to the low power mode in [14], a BS consumes less power than that in the *on* or *standby* state. A BS in *sleep* state only activates the power supply, backend connection, and central processing unit (CPU) core and thus the transition time to *on* from *sleep* state is longer than the transition time from *standby* state to *on* state, but it is much shorter than the transition time from *off* to *on* state. If the traffic load of a BS is very low, *sleep* state is more favorable. On the other hand, if the traffic load is high, the *on* state is more favorable. In the *off* state, a BS consumes no power since all of the operations of a BS are inactive. The transition time from *off* state to *on* state is the longest. In [15], the authors analyzed the relationship between the energy efficiency of a BS for varying traffic load and concluded that *sleep* state is more efficient if the traffic load is low. The authors also obtained an appropriate threshold value to make a transition to the *sleep* state. In [16], the authors analyzed the relationship between power efficiency and residence time in the *sleep* state.

In parallel to the study of the energy efficiency related to the BS states, there have been attempts to save the BS energy using new network architecture. In conventional heterogeneous network architecture, a number of small cells are overlaid on a macro cell. Small cells are introduced to accommodate the varying traffic distribution and increasing data traffic load. However, in the conventional heterogeneous network, energy efficiency is not significant since all the BSs in the macro cell and small cell should deal with both the control signal and data traffic, and thus should always be awake.

In a heterogeneous network architecture with overlapped macro cell BSs and small cell BSs, if small BSs are not used they are turned off to save energy [17]. In [18], the activation and deactivation of small cell BSs are controlled by a macro cell BS using the traffic load and position of UEs but it has high signaling overhead [18]. In [19], the authors propose distributed activation/deactivation schemes, where a small cell BS is turned off if there is no serving UE and periodically awakes to detect the activity of UEs. In device-assisted-networking for cellular greening (DANCE) schemes, small cell BSs are activated or deactivated for varying traffic loads to optimize the number of activated/deactivated small cell BSs [20]. In [21], the authors proposed a joint BS on/off switching and user association algorithm for energy efficiency maximization. In [22], the on/off states of small cell BSs are controlled based on two uniform and non-uniform distributions of UEs.

In [23–25], the authors analyze the energy efficiency of BSs in heterogeneous networks shared by multi-operators. In [23], the authors use distributed game theory to save power, where UEs of any mobile network operator are served by another mobile network operator and BSs of these UEs are turned off. In [24], the authors propose a cooperative BS switching off scheme, where BSs are turned off when its traffic load is low and UEs served by them can be covered by BSs operated by other mobile network operators. In [25], the authors propose pricing decision game and user association game, based on the relationship between roaming cost and user association.

To further reduce the energy consumption of BSs, a new concept of control split plane and user plane or the separation of data and control was proposed [26,27]. In heterogeneous network architecture with separated data and control, as shown in Figure 3, a macro cell BS (MBS) manages

the control signal and data traffic. On the other hand, small cell BS (SBS) manages only data traffic, since the control signal is provided by the MBS. The energy of SBSs can be saved significantly if there is little traffic within the coverage of SBSs, since they do not have to be awake to support the control signal, which is supported by a macro cell always and thus can be put into an inactive state to save power. A heterogeneous network can accommodate ever-increasing data traffic and can also reduce the energy consumption of BSs [28]. The performance of heterogeneous networks depends on the layout of SBSs when the density of data traffic is high or transmission power of a BS is low. Also, SBSs should awake in order to accommodate UEs from either macro cell or deactivated small cell and thus, energy consumption may increase [28,29]. In cloud cooperated heterogeneous network architecture, an SBS has high transmission power in a direction towards a highly concentrated traffic area and thus, data transmission efficiency is improved [30]. Also, the cloud-radio access network (C-RAN) reduces unnecessary resource usage. In [31], technical challenges of BS on/off switching schemes are discussed in 5G systems, and the state of the art on BS on/off switching, technical challenges, and open problems are presented. In [32], dynamic traffic BS switching on/off scheme was proposed, in separated network architecture with coverage BS and traffic BSs. In [32], random sleeping scheme is considered. In [33], the authors proposed random and repulsive schemes in a separated network architecture. In random scheme, every BS is switched off with probability p . On the other hand, in repulsive scheme, small cell BSs with a smaller distance than R from a macro cell BS is switched off.

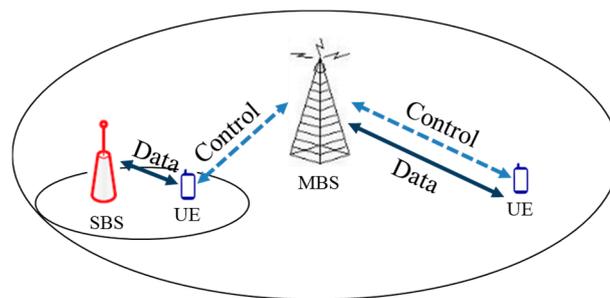


Figure 3. Separated data and control architecture.

Works on the optimization models and solution algorithms for the design and management of 5G networks have been carried out actively recently [34–38]. In [34], the authors design reusable functional block (RFB)-based 5G network architecture and consider the optimization problem of dynamically managing the RFBs by either maximizing user throughput or minimizing the number of used nodes subject to RFBs placement, 5G node capacity, user coverage and user data constraints. In [35], the authors develop a robust mixed integer optimization model for minimizing the power consumption of network function virtualization (NFV) infrastructure with uncertain resource demands by the virtual network functions (VNFs). Since it takes a very long time to solve the robust optimization problem using an exact solver, the authors develop a fast three-phase heuristic, i.e., green and robust VNFs placement. In [36], the authors propose novel optimization models to minimize the energy consumption of computing and network infrastructure of 5G networks by using both exact and heuristic methods. The authors show that the proposed heuristic converges very fast, compared to the exact solver. In [37], the authors design and implement a video streaming service using mobile edge computing functionalities in 5G networks based on virtual functions. In [38], the authors design an efficient PSO-based algorithm to deliver a high definition video to the users in 5G networks based on RFBs, with the objective of either the maximization of user throughput or the minimization of the number of used 5G nodes.

Wireless network design is considered as one of the important topics of wireless network areas and readers are referred to the works in [39–42] for better understanding. In [39], optimization issues in telecommunication are covered extensively, and the work in [40] deals with optimization problems

and models for the design of cellular networks. In [41], up-to-date challenges in the modern design of wireless networks are presented. In [42], an exhaustive introduction to the use of mathematical optimization models and algorithms for wireless network design is provided and also a hierarchy of design problems is discussed in detail.

Works on the optimization models and algorithms for wireless network design have been covered extensively in [43–52]. In [43], the authors investigate the use of accurate (non-finite precision) optimization solvers for linear and mixed-integer programs applied to signal-to-interference-based wireless coverage formulations. In [44], the authors present a mathematical formulation for wireless network design purely based on binary user-transmitter assignment variables to maximize network coverage. In [45], the authors propose a pure 0–1 formulation based on generalized upper bound cover inequalities for wireless network design for configuring a set of transmitters for coverage to a set of receivers based on a finite set of power values. In [46], the authors propose a genetic algorithm as a fast heuristic to solve large realistic instances of the power, frequency, and modulation assignment problem and show that the proposed genetic algorithm helps to improve the value of solutions found through a wider exploration of the power space. In [47], the authors propose a robust optimization model for tackling the uncertainty of wireless signal propagation in the design of fiber-based telecommunications access networks integrating wired and wireless technologies. Then, they propose a heuristic algorithm to find the solution to the proposed algorithm. In [48], the authors present a new model and solution for assigning users to mesh access points and for routing data within wireless mesh network to achieve max-min fair throughput by formulating a mixed-integer linear programming problem and propose a fast decomposition heuristic algorithm to solve the optimization problem. In [49], the authors propose effective ways to pursue interference minimization while maintaining the network coverage to improve cell capacity. In [50], the authors consider assignment of power level to access points and the assignment of user terminals to the opened access points in green wireless local area networks. Then they propose an effective branch-and-Benders-cut algorithm for power minimization in wireless networks. In [51], the authors introduce network packing problem in terrestrial broadcasting, where analog and digital networks share a common set of resources and propose a two-stage heuristic to optimize emission powers and frequencies. In [52], the authors provide an exact solution method for the universal mobile telecommunication systems/wideband code division multiple access (UMTS/WCDMA) network planning problem with signal-to-interference ratio constraints based on a nested Benders decomposition.

The multi-objective nature of wireless network design is another important issue and has been covered in [53–55]. In [53], the authors formulate WCDMA network planning as a multi-objective problem with the objective of maximization of system capacity and the minimization of installation cost. In [54], the authors develop the multi-objective problem with the objective functions of coverage optimization, transmitter minimization, exposure minimization, energy efficiency, carbon footprint, and green deployment for the green network planning of single frequency network based orthogonal frequency division multiplexing (OFDM) schemes. In [55], the authors propose a novel optimization model for resource assignment in heterogeneous wireless networks with the objective of maximization of the number of served users and the minimization of the granted utility at once.

The topic of data uncertainty in wireless network design and robust wireless network design has been considered important and significant works have been carried out [56–62]. In [56], the authors discuss how to apply robust optimization for tackling traffic uncertainty in telecommunication network design, considering capacitated network design problems. In [57], the authors formulate the robust counterpart of a linear program with an uncertain coefficient matrix, when the multi-band uncertainty set is considered. In [58], the authors propose a robust cutting-plane algorithm drawing inspiration from multiband robust optimization to deal with the uncertain nature of the jamming problem that optimally places and configures a set of jammers. In [59], the authors develop a robust optimization model to solve the traffic uncertainty in a multiperiod network design problem and obtain a solution using a hybrid heuristic based on the combination of ant colony optimization and an exact large

neighborhood search. In [60], the authors propose a robust optimization to save energy in wireless local area networks under both rate uncertainty and user mobility. In [61], the authors propose an optimal transmitter power allocation in a wireless communication network when the useful link coefficients and interference coefficients are random. In [62], the authors introduce a yield management model and develop a stochastic programming model for channel allocation to optimize revenue under demand uncertainty.

Although there have been numerous works regarding the energy saving of a BS in both homogeneous and heterogeneous networks, most of the works have not considered separated control and data planes. Also, works regarding the energy efficiency of a BS in 5G networks with separated control and data planes have focused on the optimization modeling and performance analysis of such separated network architecture. To the best of our knowledge, little work has been performed regarding the efficient way of energy saving of a BS in 5G networks with separated control and data planes, and mobility is not considered in the analysis of the proposed protocol [33].

In this paper, we propose an efficient energy saving scheme for 5G networks with separated control and data planes, where the state of a BS is determined. Then, an efficient energy saving scheme for BSs is proposed, where the state of a BS is determined depending on the number of UEs that request high rate data traffic and the number of UEs that exist under the overlapping areas commonly covered by the considered BS and the neighbor BSs. We formulate an optimization problem for the proposed energy saving scheme and the solution is obtained using particle swarm optimization. The basic idea of the proposed energy saving protocol is similar to our preliminary work [34] but the proposed work significantly extends the previous work in terms of the following aspects:

- The proposed energy saving scheme is further elaborated and detailed algorithms from the aspects of state management of BSs and session management of UEs are proposed.
- A thorough optimization problem for the proposed energy saving scheme is formulated for the performance analysis.
- Particle swarm optimization is applied to practically solve the formulated optimization problem.
- Extensive numerical examples are obtained through simulations developed by the authors and analyzed in detail.

The remaining part of this paper is organized as follows: Section 2 proposes an energy saving scheme. Section 3 analyzes the performance of the proposed scheme using particle swarm optimization and numerical examples are presented in Section 4. Finally, Section 5 summarizes this work and presents future work.

2. Proposed Energy Saving Scheme

In this paper, we consider 5G networks with heterogeneous macro cells and small cells, where data and control planes are separated, and we consider two types of data traffic, i.e., low rate data traffic and high rate data traffic. In basic separation architecture, an MBS manages the control signal and an SBS manages both low rate data traffic and high rate data traffic. In the considered modified separation architecture, which is similar to the work in [29], an MBS manages the control signal and low rate data traffic and an SBS manages high rate data traffic. Table 1 summarizes the basic and modified separation architectures. We note that data traffic of a UE which is not within the coverage of any SBS should be accommodated by an MBS, which is common to both basic and modified separation architectures. In the modified separation architecture, the state of a BS is defined as the *on*, *off*, and *sleep* states, and a BS in *sleep* state operates as in low power mode, as defined in [26]. If a BS is in *sleep* state, it consumes significantly lower power than that in *on* state. A BS in *sleep* state has smaller state transition time to/from *on* state than that between *on* and *off* states.

In the basic separation architecture, since both low rate data traffic and high rate data traffic are managed by an SBS only, if a UE is within the coverage of an SBS, the SBS can move to *off* state, only if there is no UE requiring either low or high rate data traffic. On the other hand, in the modified

separation architecture, since an SBS only manages high rate data traffic, it has a higher chance of moving into *off* state than that in the basic separation architecture. In basic separation architecture and modified separation architecture, energy saving of BSs can be performed by turning off unused BSs and this energy saving scheme is denoted as a conventional energy saving scheme in this paper.

Table 1. Basic and modified separation architectures.

Type		Basic Separation Architecture	Modified Separation Architecture
Control Signal		Macro Cell Base Station	Macro cell Base Station
Data Traffic	Low Rate Data	Small Cell Base Station	Macro cell Base Station
	High Rate Data	Small Cell Base Station	Small cell Base Station

In the proposed energy saving scheme, the state of a BS is determined depending on the number of UEs that request high rate data traffic and the number of UEs that exist under the overlapping areas commonly covered by the considered BS and the neighbor BSs. If all the UEs that request high rate data traffic under the coverage of a BS can be covered by neighbor BSs, the considered BS can be put into either sleep state or off state to save energy. In this paper, if a BS has a high probability of activation due to a higher number of UEs requiring high rate data traffic under the coverage of the BS or a higher number of UEs in the overlapping area commonly covered by neighbor BSs and the considered BS, the BS can be put into *sleep* state for fast transition to *on* state when needed. Otherwise, the BS moves to off state in order to save more energy.

Figure 4 shows the conventional energy saving operation of BSs in the modified separation architecture, where the state of an SBS can be either *on* or *off* state. The UE(HD) represents a UE with high rate data (HD) traffic and the UE(LD) represents a UE with either low rate data (LD) traffic or no traffic. Since an MBS can manage low rate data traffic, only SBSs with UEs requiring high rate data traffic needs to be awake and other SBSs can be turned off to save energy. In the left hand figure of Figure 4, the SBSs 2, 3, and 4 need to be in *on* state to serve UEs with high rate data traffic. On the other hand, the SBS 2 can be turned off if all the UE(HD)s covered only by the SBS2 finishes their high rate data traffic, as shown in the right hand figure of Figure 4.

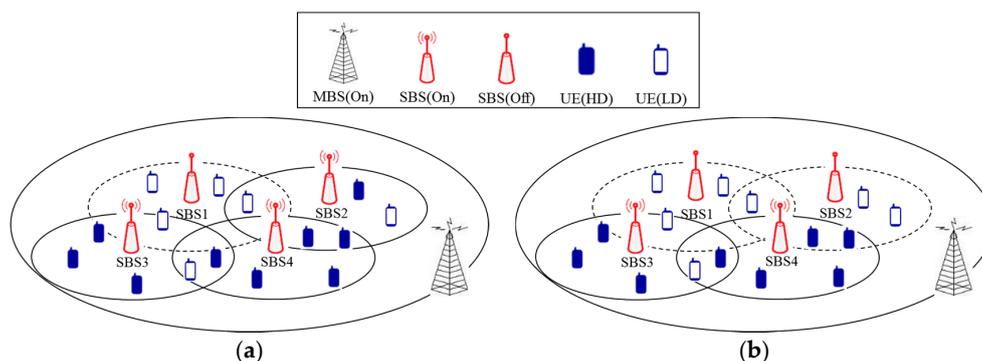


Figure 4. Conventional energy saving operation of BSs in the modified separation architecture (a) before energy saving; (b) after energy saving.

Figure 5 shows the operation of the proposed energy saving operation of BSs in the modified separation architecture, where the state of an SBS can be either *on*, *sleep*, or *off* state. In the left hand figure of Figure 5, the SBSs 2, 3, and 4 need to be in *on* state to serve UEs with high rate data traffic. On the other hand, the SBS 2 can be put into *sleep* state, if all UE(HD)s covered only by the SBS2 finishes their high rate data traffic, as shown in the right hand figure of Figure 5. In the proposed energy saving operation, an SBS with more UEs which request high rate data traffic and with UEs that exist under the

overlapping areas commonly covered by the BS and the neighbor BSs moves into *sleep* state, instead of the *off* state to react quickly on the change of traffic demands of UEs as well as to save energy, if the UEs located within the coverage area of the SBS can be covered by the neighboring SBSs.

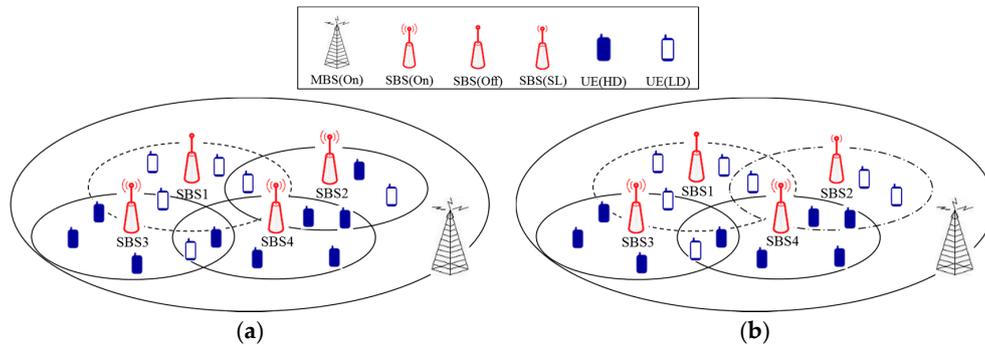


Figure 5. Proposed energy saving operation of BSs in the modified separation architecture (a) before energy saving; (b) after energy saving.

Figure 6 shows the flowchart of the state management of BSs in the proposed energy saving scheme. In Figure 6, $n(u_{r,j})$ is the number of UEs that request high rate data traffic and $n(u_{l,j})$ is the number of UEs that exist in the overlapping areas commonly covered by the BS and the neighbor BSs. An MBS manages all of the related information about SBSs within the coverage of the MBS. If $n(u_{r,j})$ of SBS j is larger than a threshold value of P_{thr} , or $n(u_{l,j})$ of SBS j is larger than a threshold value of I_{thr} , SBS j switches to sleep state to save energy. Otherwise, SBS j switches to off state.

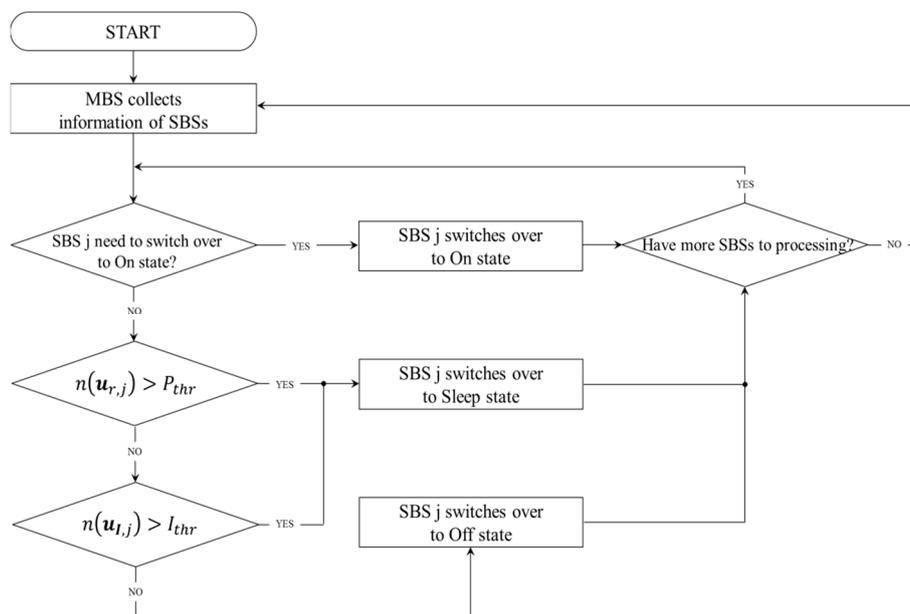


Figure 6. State management of BSs.

Figure 7 shows a flowchart of the session management of UEs in the proposed energy saving scheme. If a UE requests a service, the MBS checks the traffic type of the UE. If a UE requires high rate data traffic, the MBS requests the nearest SBS to accommodate the data traffic. On the other hand, the MBS itself accommodates the data traffic of the UE.

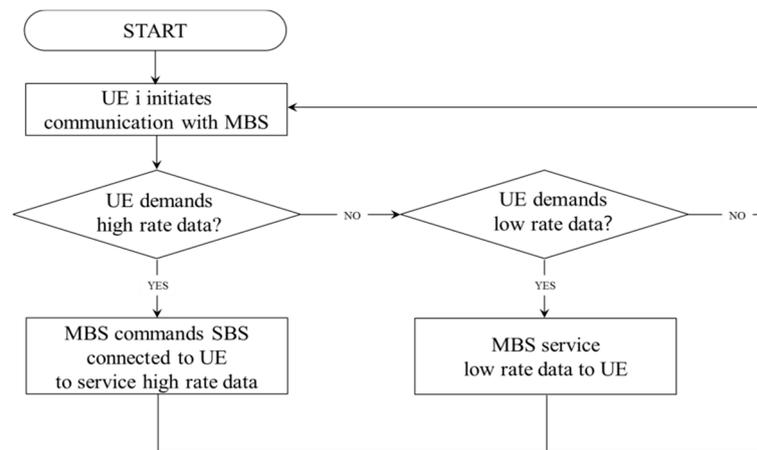


Figure 7. Session management of user equipments (UEs).

3. Performance Analysis

In this paper, a heterogeneous network architecture with a macro cell with overlaid small cells overlapped with each other is assumed for performance evaluation, as shown in Figure 8. Within a macro cell, small cells are densely distributed at a few hot spot areas, which are generated randomly. Small cells are sparsely distributed outside the hot spot areas. We note that we only consider a single macro cell in this paper, similar to the work in [63]. Work on multi-macro cell environment will be carried out in our future work. The mobility of a UE follows a random movement model with locality within hot spot areas and it follows a random movement outside hot spot areas. The speed of a UE follows a uniform distribution. Data traffic generated at a UE is defined as either low rate data traffic or high rate data traffic, depending on a predefined value of the ratio of high rate data traffic.

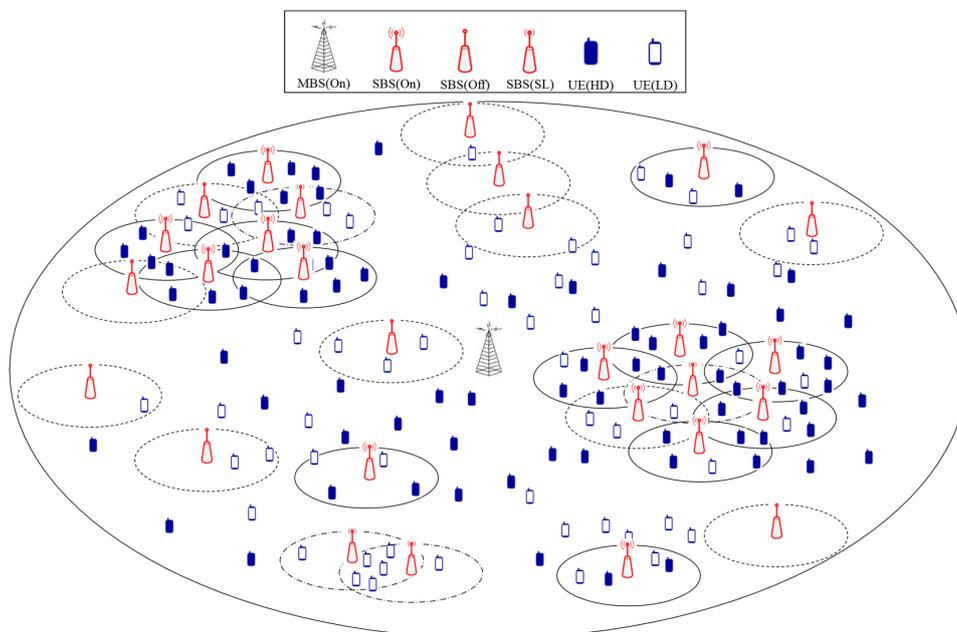


Figure 8. Network configuration.

In the proposed energy saving scheme, the main objective is to associate UEs with appropriate BSs with maximum energy efficiency. To this end, we define an optimization problem, where the total sum of the power consumption of BS is set as an objective function and the maximum power consumption

of BSs and the context information of BSs are set as constraints. The optimization problem in this paper is Non-Deterministic Polynomial-time hard (NP-hard) problem and thus the solution is obtained using the particle swarm optimization algorithm, which is one of the widely used bio-inspired optimization algorithms. This is because it is generally accepted that the particle swarm optimization has merit of easy implementation, very simple calculation, and fast calculation speed within a short time. Also, it is less dependent of initial points than other optimization algorithms and there exist techniques which ensure convergence. Therefore, it has been applied to network issues such as the resource allocation problem [64] and routing algorithms [65].

Table 2 shows the notations used in the development of the optimization problem.

Table 2. Notations.

Parameter	Description
n	number of UEs
m	number of SBSs
u_i	a UE i
B_j	a BS j
C_h	capacity needed to serve a high rate data traffic service
C_l	capacity needed to serve a low rate data traffic service
C_S^{max}	maximum capacity of an SBS
C_M^{max}	maximum capacity of an MBS
ρ_S	slope of the load-dependent power consumption in an SBS
ρ_M	slope of the load-dependent power consumption in an MBS
P_S^f	fixed power consumption of an SBS in <i>on</i> state
P_S^e	fixed power consumption of an SBS in <i>sleep</i> state
P_M^f	fixed power consumption of an MBS in <i>on</i> state
P_S^{max}	maximum power consumption of an SBS
P_M^{max}	maximum power consumption of an MBS
P_{on-off}	power consumption for a switching between <i>on</i> and <i>off</i> states in an SBS
$P_{on-sleep}$	power consumption for a switching between <i>on</i> and <i>sleep</i> states in an SBS
T_{on-off}	delay for a switching between <i>on</i> state and <i>off</i> state in an SBS
$T_{on-sleep}$	delay for a switching between <i>on</i> state and <i>sleep</i> state in an SBS
T_c	delay for a new connection initiation between a UE and a BS

The objective function is the sum of power consumption by an MBS, P_{macro} and power consumption by SBSs under the coverage of an MBS, $\sum_{j=0}^{m-1} P(j)_{small}$, and power consumption due to switching between different states, $P_{switching}$ and it is denoted as P_{total} as in Equation (1):

$$P_{total} = P_{macro} + \sum_{j=0}^{m-1} P(j)_{small} + P_{switching}, \quad (1)$$

The connection between BS_j and UE i is denoted as $a_{j,i}$, and the value of it is defined as in Equation (2), where we note that BS_m corresponds to an MBS and BS_j ($0 \leq j \leq m - 1$) corresponds to an SBS. We assume that all UEs under the coverage area of an MBS are associated with the MBS.

$$a_{j,i} = \begin{cases} 1, & \text{If } u_i \text{ is connected with } B_j, (0 \leq j < m + 1, 0 \leq i < n), \\ 0, & \text{otherwise} \end{cases}, \quad (2)$$

The data traffic demand of a UE i is denoted as d_i , and the value of it is defined as in Equation (3):

$$d_i = \begin{cases} 1, & \text{If } u_i \text{ required data service}, (0 \leq i < n), \\ 0, & \text{otherwise} \end{cases}, \quad (3)$$

The type of required data traffic of UE i is denoted as r_i and the value of it is defined as in Equation (4):

$$r_i = \begin{cases} 1, & \text{If } u_i \text{ required high rate data service} \\ 0, & \text{If } u_i \text{ required high low data service} \end{cases}, (0 \leq i < n), \quad (4)$$

In Equation (5), w_j is related with whether the BS j is in *on* state or not and e_j is related with whether the BS j is in *sleep* state or not in Equation (6):

$$w_j = \begin{cases} 1, & \text{If } B_j \text{ is in on state} \\ 0, & \text{otherwise} \end{cases}, (0 \leq j < m + 1), \quad (5)$$

$$e_j = \begin{cases} 1, & \text{If } B_j \text{ is in sleep state} \\ 0, & \text{otherwise} \end{cases}, (0 \leq j < m), \quad (6)$$

The notations related with state transitions of an SBS are f_j and s_j , and the values of them are defined as in Equations (7) and (8):

$$f_j = \begin{cases} 1, & \text{If } B_j \text{ switches between on state and off state} \\ 0, & \text{otherwise} \end{cases}, (0 \leq j < m), \quad (7)$$

$$s_j = \begin{cases} 1, & \text{If } B_j \text{ switches between on state and sleep state} \\ 0, & \text{otherwise} \end{cases}, (0 \leq j < m), \quad (8)$$

Using the above equations, the power consumption of an SBS j , $P(j)_{small}$, is obtained as in Equation (9):

$$P(j)_{small} = w_j \times (1 - e_j) \times (P_S^f + \rho_S \times P_S^{tx}(j)) + e_j \times (1 - w_j) \times P_S^e, \quad (9)$$

where if an SBS is in *on* state, the power consumption is calculated as the sum of fixed power consumption and load-dependent power consumption, and if an SBS is in *sleep* state, then the power consumption for *sleep* state is used for the calculation. Regarding load dependent power consumption, $P_S^{tx}(j)$ is defined as in Equation (10):

$$P_S^{tx}(j) = P_S^{tx,max} \times \sum_{i=0}^{n-1} \left(d_i \times a_{j,i} \times \frac{(r_i \times C_h + (1 - r_i) \times C_l)}{C_S^{max}} \right) \quad (10)$$

Likewise, the power consumption of MBS, P_{macro} , is obtained as in Equation (11) as the sum of fixed power consumption and load-dependent power consumption.

$$P_{macro} = P_M^f + \rho_M \times P_M^{tx}, \quad (11)$$

where P_M^{tx} is defined as as in Equation (12):

$$P_M^{tx} = P_M^{tx,max} \times \sum_{i=0}^{n-1} \left(d_i \times a_{j,i} \times \frac{(r_i \times C_h + (1 - r_i) \times C_l)}{C_M^{max}} \right), \quad (12)$$

Also, the power consumption of state transitions of SBS, $P_{switching}$, is obtained as in Equation (13):

$$P_{switching} = \sum_{j=0}^{m-1} \left(f_j \times (1 - s_j) \times P_{on-off} + s_j \times (1 - f_j) \times P_{on-sleep} \right), \quad (13)$$

It is assumed that the maximum power consumption of an SBS is P_S^{max} , and thus, the power consumption of an SBS j should be smaller than or equal to P_S^{max} :

$$P(j)_{small} \leq P_S^{max}, (0 \leq j \leq m - 1), \quad (14)$$

Likewise, it is assumed that the maximum power consumption of an MBS is P_M^{max} , and thus, the power consumption of an MBS should be smaller than or equal to P_M^{max} :

$$P_{macro} \leq P_M^{max}, \quad (15)$$

If the SBS j is on state, UE i can have the connection with SBS j . However, if the SBS j is not on state, UE i cannot have the connection with SBS j . This relationship is constrained as follows:

$$a_{j,i} \leq w_j, \quad (16)$$

The type of required data traffic of UE i is defined as either 1 or 0 depending on the type of the required data service and this is smaller than or equal to the existence of data traffic demand of UE i :

$$r_i \leq d_i, \quad (17)$$

Since UE i can have at most one connection with either an MBS or any of SBSs within an MBS, the constraint is expressed as follows:

$$\sum_{j=0}^m a_{j,i} \leq 1, \quad (18)$$

Since an SBS j can be in one of on, sleep, and off states, the sum of w_j for on state and e_j for off state should be smaller than or equal to 1:

$$w_j + e_j \leq 1, (0 \leq j \leq m - 1), \quad (19)$$

Since an SBS j can have only one state transition at any time, i.e., either between on state and off state or between on state and *sleep* state, the sum of f_j for the state transition between on state and off state and s_j for the state transition between on state and off state should be smaller than or equal to 1:

$$s_j + f_j \leq 1, (0 \leq j \leq m - 1), \quad (20)$$

The maximum number of UEs in an MBS is n , the maximum number of SBSs in an MBS is m , the number of MBS is 1, and they are constrained as follows:

$$0 \leq i < n, 0 \leq j < m + 1, \quad (21)$$

Using the above constraints, the optimization problem of the proposed energy saving scheme can be formulated as follows:

$$\min P_{total}, \quad (22)$$

Subject to: (14)–(21).

In this paper, we also analyze the aggregate delay of the proposed energy saving scheme. To this end, the $T_{aggregate}$, which is defined as the time needed to get service from a BS by a UE, is additionally introduced as a performance measure as in Equation (23), where the $T_{switching}$ is defined as average time needed to make a transition between states of an SBS and the $T_{connect}$ is defined as average connection initiation time between a UE and a BS.

$$T_{aggregate} = T_{switching} + T_{connect}, \quad (23)$$

The $T_{switching}$ can be obtained as in Equation (24), where the T_{on-off} is the time needed to make a transition from *off* state to *on* state by an SBS and the $T_{on-sleep}$ is the time needed to make a transition from *sleep* state to *on* state.

$$T_{switching} = \sum_{j=0}^{m-1} (1 - w_j) \times \left((1 - e_j) \times f_j \times T_{on-off} + e_j \times s_j \times T_{on-sleep} \right), \quad (24)$$

The value of $T_{connect}$ is obtained as in Equation (25), where T_c is defined as the time needed to initiate a connection between a UE and a BS.

$$T_{connect} = \sum_{i=0}^{n-1} (d_i \times T_c), \quad (25)$$

The optimization problem defined in this paper is a mixed integer linear programming since some of the variables in the proposed optimization are restricted to be discrete integer values [66] and the objective function is linear function. The mixed integer linear programming problem considered in this paper does not scale well with the number of binary variables and it has been proven in the literature to be NP-hard [67–71]. Therefore, the solution of the optimization is obtained by particle swarm optimization, which is one of the most widely-used bio-inspired optimization algorithms [72]. In particle swarm optimization, if the search space is D dimension, the i -th particle of a swarm is denoted as $\mathbf{X}_i = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{iD})$. The velocity of a particle is denoted as $\mathbf{V}_i = (v_{i1}, v_{i2}, v_{i3}, \dots, v_{iD})$. The best position visited by the i -th particle is denoted by $\mathbf{P}_i = (p_{i1}, p_{i2}, p_{i3}, \dots, p_{iD})$. In each swarm, the initial value of position \mathbf{X} is denoted as \mathbf{Pb} and the value of the objective function applying \mathbf{Pb} is defined as P_{best} . In each P_{best} , the next P_{best} is obtained by using velocity $\mathbf{V}_i = (v_{i1}, v_{i2}, v_{i3}, \dots, v_{iD})$, and the value of P_{best} is updated and stored at \mathbf{Pb} . In each swarm, the best value of P_{best} is defined as G_{best} , and it is stored in Gb . Equation (26) shows the update of velocity.

$$v_k^{nm}(t+1) = wg \times v_k^{nm}(t) + c_1 r_1 (\mathbf{Pb}_k^{nm}(t) - \mathbf{X}_k^{nm}(t)) + c_2 r_2 (Gb_k^{nm}(t) - \mathbf{X}_k^{nm}(t)) \quad (26)$$

Since the optimization problem defined in this paper is mixed integer linear programming, discrete particle swarm optimization [69] should be used to obtain the solution. To find new \mathbf{Pb} , the values of $v_k^{nm}(t+1)$ are normalized using Equation (27):

$$S(v_k^{nm}(t+1)) = [1 + \exp(v_k^{nm}(t+1))]^{-1}, \quad (27)$$

Then, $\mathbf{X}_k^{nm}(t+1)$ is defined as in Equation (28):

$$\mathbf{X}_k^{nm}(t+1) = \begin{cases} 1, & S(v_k^{nm}(t+1)) > rd \\ 0, & \text{otherwise} \end{cases}, \quad (28)$$

Since the optimization problem of this paper is mixed integer linear programming, Equation (28) is modified as Equation (29):

$$\mathbf{X}_k^{nm}(t+1) = \begin{cases} 1, & S(v_k^{nm}(t+1)) > rd \text{ and } \max(S(v_k^{nm}(t+1))) \\ 0, & \text{otherwise} \end{cases}, \quad (29)$$

Since the formulated optimization is NP-hard, we propose a PSO-based heuristic, as shown in the following Algorithms 1–3. In Algorithm 1, PSO algorithm is applied to the formulated optimization and Algorithms 2 and 3 are used within the operation of Algorithm 1. In Algorithm 1, initially, a number of matrices, i.e., a , z , v , p , Pb , Gb , $Pbest$, $Gbest$, d , r , w , e , f , and s are defined and initialized, where the

definition and the size of each matrix is given in the Algorithm 1. Then, for a given PSO iteration, optimal BS state information is obtained for each swarm. To this end, the velocity of the particle is calculated using Equations (26) and (27). Then, Algorithm 2 for finding particle and Algorithm 3 for finding state information of BS are carried out, which will be explained later. Then, the power consumption of BSs is calculated, where the power consumption due to traffic processing is obtained firstly and the power consumption due to state switching is obtained later. The traffic load dependent power consumption of MBS is calculated using Equation (12). Then, the power consumption of MBS is calculated using Equation (11). After that, the traffic load dependent power consumption of SBS is calculated using Equation (10). Then, the power consumption of SBS is obtained using Equation (9). The power consumption due to state switching is calculated using Equation (13). Finally, the total power consumption of BSs is calculated using Equation (1). By comparing the obtained total power consumption with P_{best} , the values of P_b , P_{best} , and v are updated. After that, P_b with the best P_{best} of all the particles are chosen as the G_{best} , and then the BS state information is updated. In Algorithm 2, after applying the Equations (16)–(18), association is set depending on the demanded data service type. Then, the best index of BS is extracted using Equation (29). Then, the UE's demand is served by either MBS or SBS. In Algorithm 3, after applying Equations (19) and (20), algorithm shown in Figure 6 is carried out to find the state information of BS.

Algorithm 1. PSO Process

```

1:   Initialize  $a, z, v, p, P_b, G_b, P_{best}, G_{best}, d, r, w, e, f, s$ 
2:   //  $a$  is  $n \times m + 1$  matrix of current association between UE and BS.
3:   //  $z$  is swarm size.
4:   //  $v$  is  $z \times n \times m + 1$  matrix of velocity of particle.
5:   //  $p$  is  $z \times n \times m + 1$  matrix of particle of selected association between UE and BS.
6:   //  $P_b$  is  $z \times n \times m + 1$  of local optimum.
7:   //  $G_b$  is  $n \times m + 1$  matrix of global optimum.
8:   //  $P_{best}$  is  $1 \times z$  matrix of fitness value of  $P_b$ .
9:   //  $G_{best}$  is the best fitness value of  $P_{best}$ .
10:  //  $d$  is  $1 \times n$  matrix of the data traffic demand of UE.
11:  //  $r$  is  $1 \times n$  matrix of the type of required data traffic of UE.
12:  //  $w$  is  $1 \times m + 1$  matrix of on state of BS.
13:  //  $e$  is  $1 \times m$  matrix of sleep state of BS.
14:  //  $f$  is  $1 \times m$  matrix of state transition of SBS between on and off.
15:  //  $s$  is  $1 \times m$  matrix of state transition of SBS between on(off) and sleep.
16:  for  $iteration < PSO\_iteration$  do
17:    for  $p\_index < z$  do
18:      // Calculate velocity using the Equations (26) and (27).
19:       $v' = 1/(1 + \exp(wg \times v[p\_index,;])) + c1 \times r1 \times (P_b[p\_index,;] - a) + c2 \times r2 \times (G_b - a));$ 
20:      // Algorithm 2: Find particle
21:       $[p[p\_index,;], w'] = FindParticle(v', d, r);$ 
22:      // Algorithm 3: Find state information of BS
23:       $[e', f, s] = FindBSStateInformation(w, e, w');$ 
24:      // Calculate power consumption
25:       $P_{small} = 0, SumP_{small} = 0, P_{macro} = 0, P_{switching} = 0, P_{processing} = 0;$ 
26:      for  $j < m+1$  do
27:        // Calculate power consumption of BS due to traffic processing.
28:        if  $j == m$ 
29:          // Calculate the traffic load dependent power consumption of MBS
30:          // using the Equation (12).
31:          for  $i < n$  do
32:             $P_{macro} += P_{tx\_m} \times d[i] \times p[p\_index,i,j] \times (r[i] \times Ch + (1 - r[i]) \times Cl)/C_{max\_m};$ 
33:          end

```

```

34:         // Calculate the power consumption of MBS using the Equation (11).
35:         Pmacro = Pf_m + rho_m × Pmacro
36:     else
37:         // Calculate the traffic load dependent power consumption of SBS
38:         // using the Equation (10).
39:         Psmall = 0;
40:         for i < n do
41:             Psmall += Ptx_s × d[i] × p[p_index,i,j] × (r[i] × Ch + (1-r[i]) × Cl) / Cmax_s;
42:         end
43:         // Calculate the power consumption of SBS using the Equation (9).
44:         Psmall = w'[j] × (1 - e'[j]) × (Pf_s + rho_s × Psmall) + (1 - w'[j]) + e'[j] × Pe;
45:         SumPsmall += Psmall;
46:     end
47:     // Calculate power consumption due to state switching using the Equation (13).
48:     Pswitching += f[j] × (1 - s[j]) × Pon_off + (1 - f[j]) × s[j] × Pon_sleep
49: end
50: // Calculate total power consumption using the Equation (1).
51: Ptotal = Pmacro + SumPsmall + Pswitching;
52: // Set Pb and Pbest if the fitness value is lower than the Pbest and update v.
53: if Ptotal < Pbest
54:     Pb[p_index,;,] = p[p_index,;,];
55:     Pbest[p_index] = Ptotal;
56:     v[p_index,;,] = v';
57: end
58: // Choose the Pb with the best Pbest of all the particles as the Gbest.
59: // Update BS state information.
60: if Pbest < Gbest
61:     Gb = Pb[p_index];
62:     Gbest = Pbest[p_index];
63:     BSstate = [w', e'];
64: end
65: end
66: end
67: [w, e] = BSstate;
68: a = Gb;

```

Algorithm 2. Find Particle

```

1: FindParticle(v', d, r)
2: Initialize w, p
3: // Apply the Equations (16)–(18).
4: for i < n do
5:     // Set association if a UE demands low rate data service.
6:     if d[i] == 1 and r[i] == 0
7:         p[i, m] = 1;
8:         w[m] = 1;
9:         continue;
10:    end
11:    for j < m + 1 do
12:        // Set association if a UE demands high rate data service.
13:        if d[i] == 1 and r[i] == 1
14:            // Extract the k of the best index of BS using the Equation (29).

```

```

15:         extract  $k$ 
16:     end
17: end
18: // Serve UE's demand using MBS.
19: if  $k$  does not exist or does not meet the condition of the Equation (14).
20:      $p[i, m] = 1$ ;
21:      $w[m] = 1$ ;
22: // Serve UE's demand using SBS.
23: else
24:      $p[i, k] = 1$ ;
25:      $w[k] = 1$ ;
26: end
27: end
28: return  $[p, w]$ ;

```

Algorithm 3. Find State Information of BS

```

1: FindBSStateInformation( $w, e, w'$ )
2: Initialize  $e, f, s$ 
3: // Apply the Equations (19) and (20).
4: for  $j < m$  do
5:     //  $u_r$  is the number of UEs that request high rate data traffic.
6:     //  $u_l$  is the number of UEs that exist in the overlapping areas commonly covered
7:     // by the BS and the neighbor BSs.
8:     if  $u_r[j] > Pthr$  and  $u_l[j] > lthr$  and  $w'[j] \neq 1$ 
9:          $e[j] = 1$ ;
10:    end
11:    // Find state switching of BS.
12:    if  $w[j] \neq w'[j]$  and  $e[j] == e[j]$ 
13:         $f[j] = 1$ ;
14:        continue;
15:    end
16:    if  $w[j] == w'[j]$  and  $e[j] \neq e[j]$ 
17:         $s[j] = 1$ ;
18:        continue;
19:    end
20: end
21: return  $[e, f, s]$ ;

```

4. Numerical Results

In this section, we compare the performance of the proposed energy saving scheme with that of conventional energy saving schemes in both basic and modified separation architecture. Table 3 shows simulation parameters assumed in numerical results. According to the literature, various values are assumed as the power consumption of sleep mode, such as 1% [73], 15% [13], and 50% [74] of fixed power consumption. In Table 3, we assumed this value as 25% by approximately averaging the reference values. Also, we assumed the switching power between on and off states as the 2.5 times of fixed power consumption using the similar rationale of that in [75] and the switching power between on and *sleep* states as the 2.5 times of power consumption of *sleep* state similarly.

Table 3. Simulation parameters.

Parameter	Value
simulation time	24 h
number of SBSs	50
number of UEs	250
service time of traffic	U[0, 2.5] min
inter-arrival time of traffic	U[5, 10] min
ratio of high data	0.5
speed of UE	U[0, 100] km/h
P_S^f	21.6 W
P_S^e	5.4 W
P_S^{max}	27 W
P_M^f	780 W
P_M^{max}	1350 W
P_{on-off}	54 W
$P_{on-sleep}$	13.5 W
C_h	1 Mbps
C_l	0.01 Mbps
C_S^{max}	25 Mbps
C_M^{max}	500 Mbps
$P_S^{tx,max}$	1.3 W
$P_M^{tx,max}$	120 W
ρ_S	4.15 W/Mbps
ρ_M	4.5 W/Mbps
T_{on-off}	0.532 s
$T_{on-sleep}$	0.00216 s
T_c	0.05 s

The parameter values assumed in the simulation of particle swarm optimization is summarized as in Table 4.

Table 4. Parameters for particle swarm optimization (PSO).

Parameter	Value
PSO iteration	200
number of swarm	5
v_{max}	4
v_{min}	-4
c_1	0.5
c_2	0.5
r_1	0.5
r_2	0.5
wg	0.25

For numerical examples, we obtained energy consumption due to traffic processing, energy consumption due to state switching, total energy consumption, average delay for varying number of UEs, ratio of high rate data, and service time of traffic. There are three schemes, i.e., basic separation, modified separation, and proposed energy saving schemes. In the basic separation scheme, the conventional energy saving scheme is applied in the basic separation architecture. In the modified separation scheme, conventional energy saving scheme is applied in the modified separation architecture. In the proposed energy saving scheme, the proposed energy saving scheme is applied in the modified separation architecture.

Figures 9–11 show the energy consumption of BSs due to traffic processing, energy consumption of BSs due to state switching, and total energy consumption of BS by varying the number of UEs. As shown in Figure 9, the energy consumptions of three schemes increase as the number of UEs increases, since traffic load and the number of SBSs in *on* state increase. The basic separation scheme has the largest energy consumption since low rate data traffic is managed by SBSs and thus more SBSs should be awake. The proposed energy saving scheme has slightly lower energy consumption than the modified separation scheme, since the proposed energy saving scheme can reduce the number of SBSs in *on* state by efficiently putting SBSs into *sleep* state. As shown in Figure 10, the proposed energy saving scheme has the smallest energy consumption for most values of the number of UEs

since the state of SBSs in the proposed scheme changes between *on* and *sleep* states, which has the smaller energy consumption due to state switching than the transition between *on* and *off* states. In the basic separation scheme, where SBSs manage both low and high rate data traffic, if the number of UEs is small, the energy consumption due to state switching increases if the number of UEs increases, because the state of SBSs changes from *off* to *on* state. However, if the number of UEs is sufficient, the number of SBSs in *on* state is sufficient already, and thus, the effect of energy consumption increase due to state switching decreases. In the modified separation scheme, the SBSs only manage high rate data traffic and the increase of energy consumption due to state switching occurs more slowly than that in the basic separation scheme. Also, the decrease of energy consumption due to state switching also occurs more slowly. Figure 11 shows total energy consumption of BSs, which is the sum of the results in Figures 9 and 10. As expected from Figures 9 and 10, the proposed energy saving scheme has the smallest energy consumption.

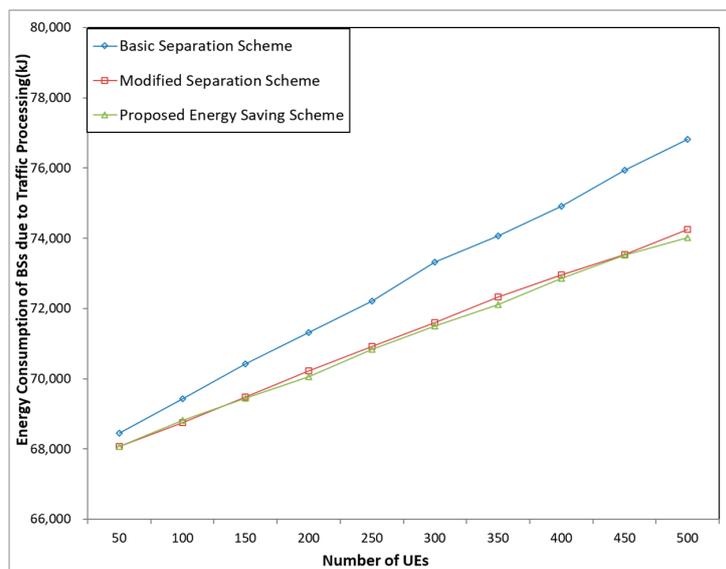


Figure 9. Energy consumption of BSs due to traffic processing.

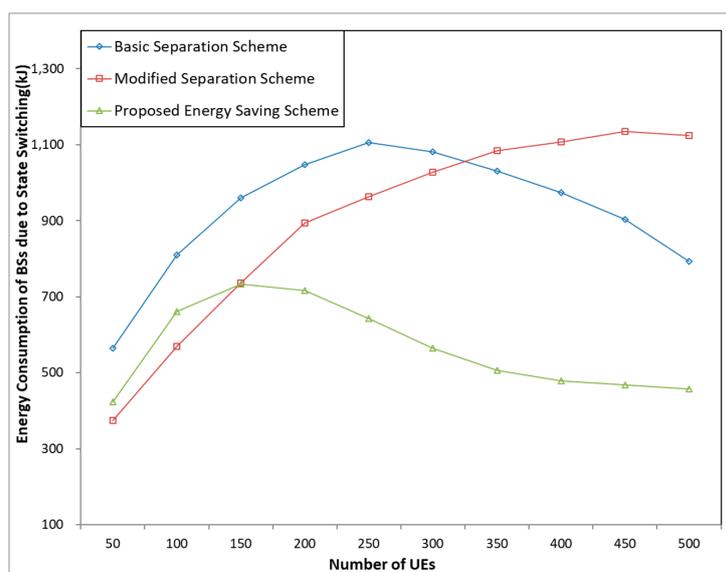


Figure 10. Energy consumption of BSs due to state switching.

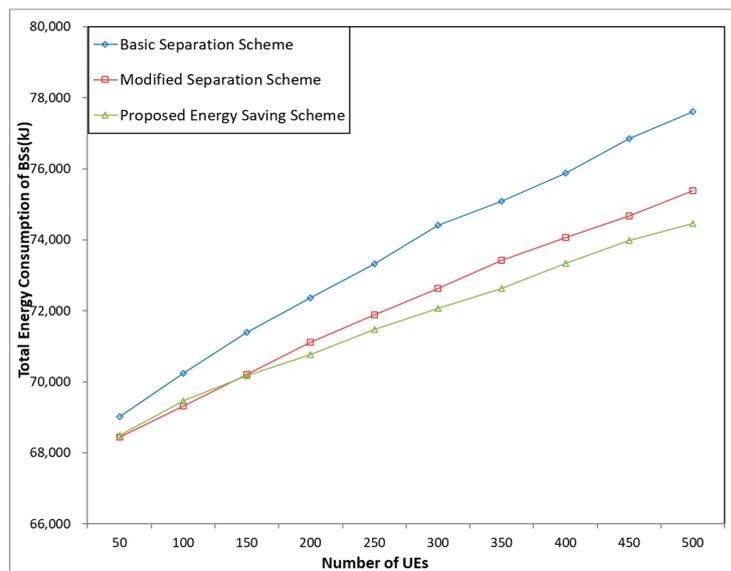


Figure 11. Total energy consumption of BSs.

Figure 12 shows aggregate delay and reveals that the proposed energy saving scheme has the smallest delay, since the other two schemes have to move from *off* state to *on* state to accommodate traffic load, which takes relatively longer delay than the delay from *sleep* state to *on* state in the proposed energy saving scheme. There is a tradeoff between the basic separation scheme and the modified separation scheme. For small values of UE, the basic separation scheme has more aggregate delay since both low rate data traffic and high rate data traffic are managed by small cells and thus, SBSs have to awake to accommodate data traffic. However, in the modified separation architecture, since low rate data traffic is managed by an MBS, SBSs only need to awake to manage high rate data traffic. For large values of UE, most SBSs are already awake to accommodate both low rate data and high rate data traffic from UEs and have less delay compared to the modified separation scheme, where a part of SBSs should still wake up to manage high rate data traffic.

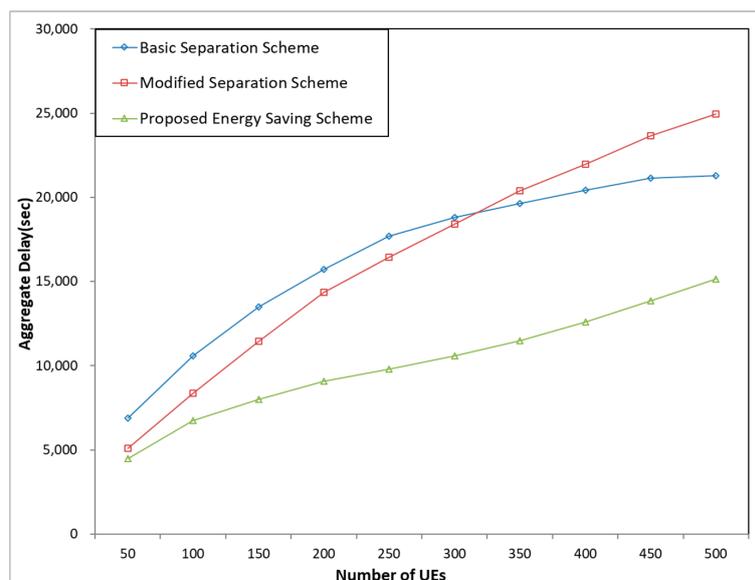


Figure 12. Aggregate delay.

Figures 13–15 show the energy consumption of BSs due to traffic processing, energy consumption of BSs due to state switching, and total energy consumption of BS by varying the ratio of high data. As shown in Figure 13, the energy consumptions of BSs due to traffic processing of the three schemes increase as the ratio of high rate data traffic increases, since more power is consumed to process high rate data traffic than low rate data traffic. The modified separation scheme and the proposed power saving scheme have lower energy consumption since low rate data traffic can be managed by a MBS and thus more SBSs can be turned off or put into sleep mode to save more energy. The difference is more significant for smaller values of ratio of high data traffic. As shown in Figure 14, the basic separation scheme has the largest energy consumption due to state switching and is not sensitive to the change of ratio of high data traffic since SBSs manage both low and high rate data traffic and thus more SBSs should change states. In the modified separation scheme and the proposed energy saving scheme, the energy consumption increases as the ratio of high data traffic increases since more SBSs should be turned on. For large values of ratio of high rate data traffic, the energy consumption of BSs due to state switching saturates because most SBSs are already turned on. The proposed energy saving scheme has the smallest energy consumption of BSs because of smaller power consumption due to switching between on and sleep states. Figure 15 shows the total energy consumption of BSs. As expected from Figures 13 and 14, the proposed energy saving scheme has the smallest energy consumption.

Figure 16 shows aggregate delay. The proposed energy saving scheme has the smallest delay, as expected. There is a tradeoff between the basic separation scheme and the modified separation scheme. Except for large values of UE, the basic separation scheme has more aggregate delay since both low rate data traffic and high rate data traffic are managed by small cells and thus SBSs have to wake up to accommodate data traffic. For large values of UE, most SBSs are already awake and have less delay compared to the modified separation scheme, where a part of SBSs should still wake up to manage high rate data traffic.

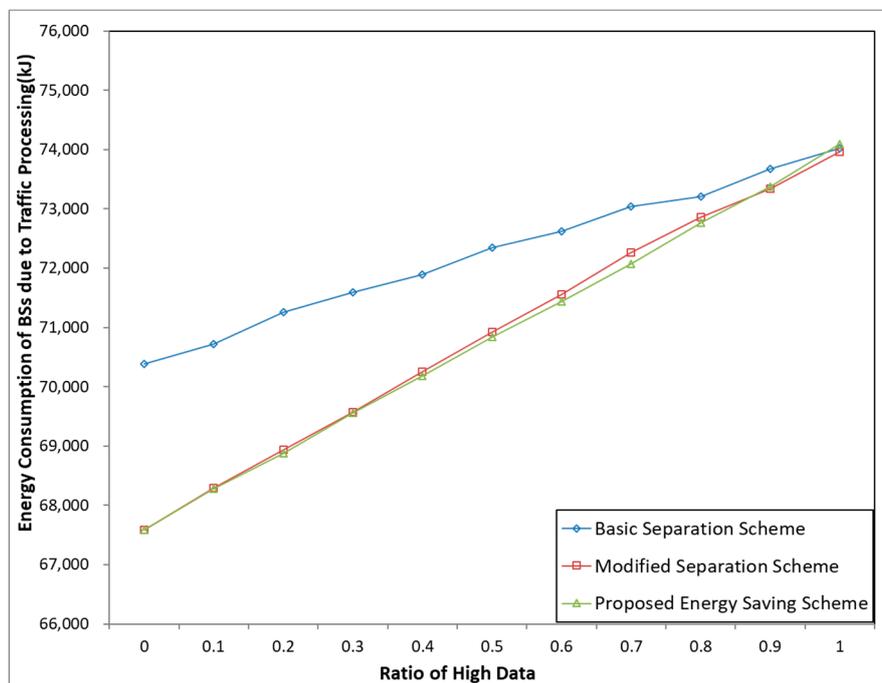


Figure 13. Energy consumption of BSs due to traffic processing.

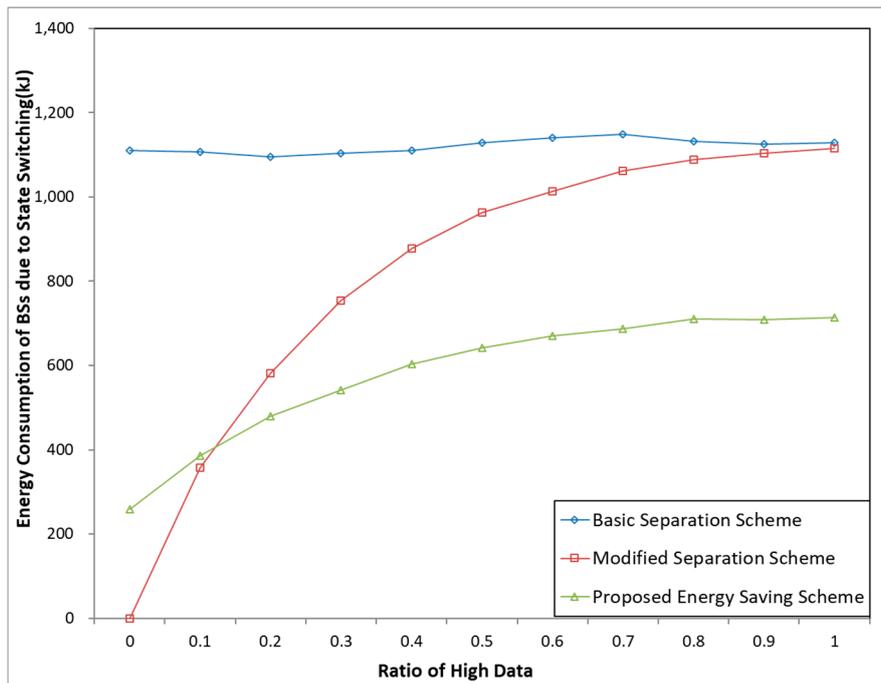


Figure 14. Energy consumption of BSs due to state switching.

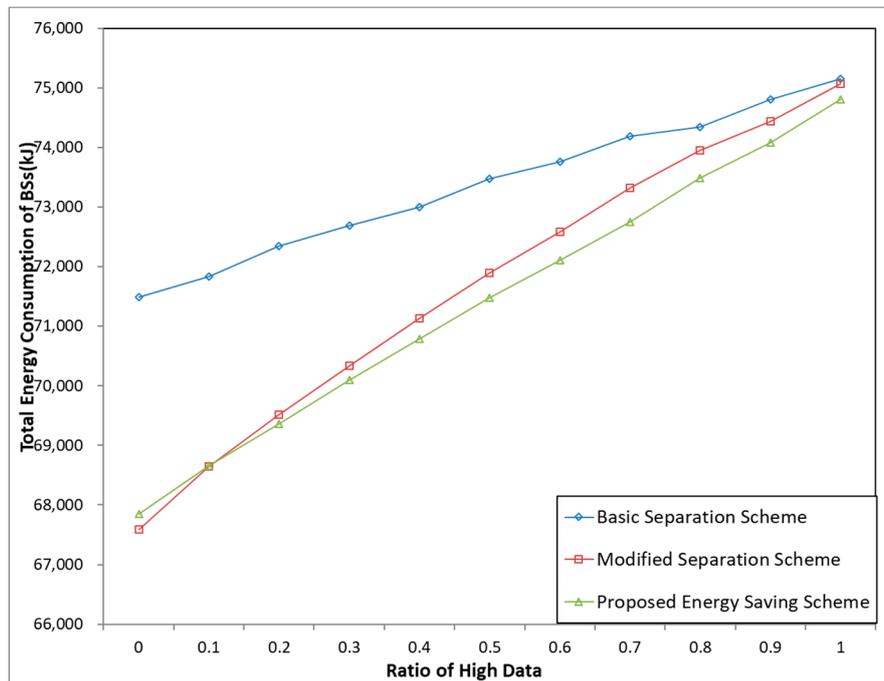


Figure 15. Total energy consumption of BSs.

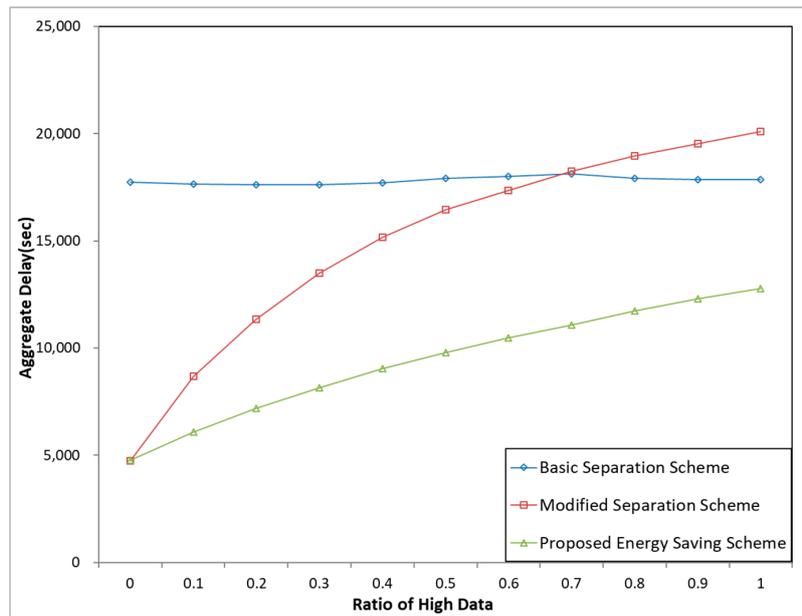


Figure 16. Aggregate delay.

Figures 17–19 show the energy consumption of BSs due to traffic processing, energy consumption of BSs due to state switching, and total energy consumption of BS by varying service time of traffic. Since increasing service time of traffic has a similar effect of increasing the number of UEs, the shapes of the results in Figures 17–19 are very similar to those in Figures 9–11, as expected. Figure 20 shows the aggregate delay, which follows similar shapes with the results in Figure 18, and the proposed energy saving scheme has the smallest aggregate delay, similar to that in Figures 12 and 16 based on similar rationale.

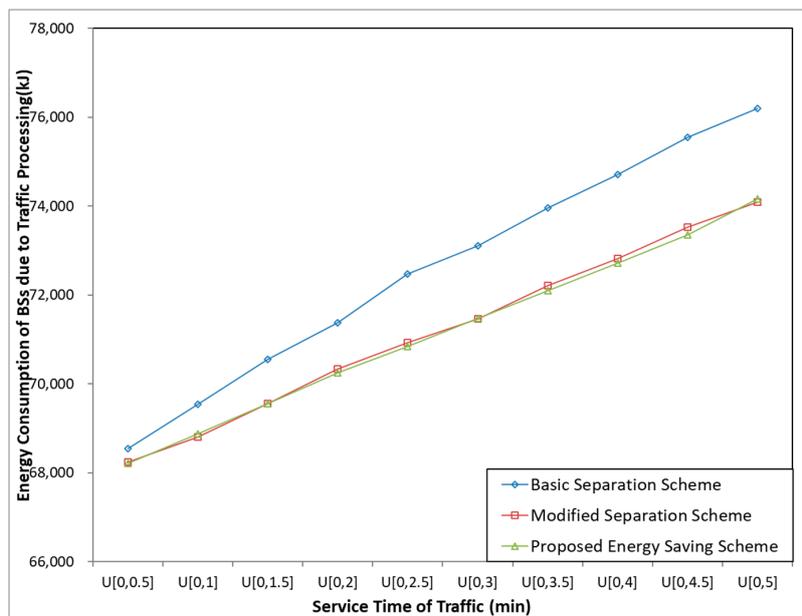


Figure 17. Energy consumption of BSs due to traffic processing.

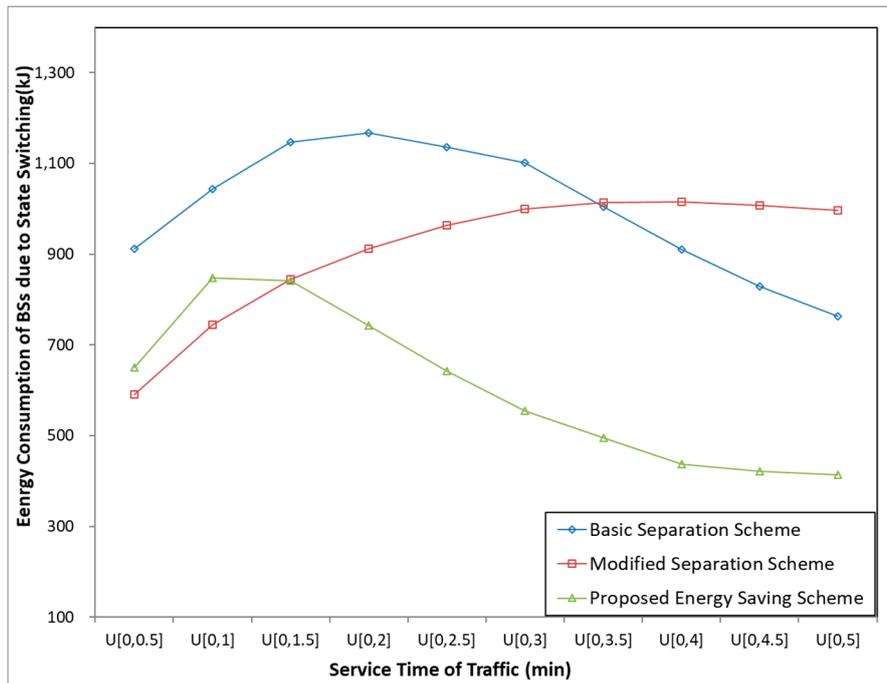


Figure 18. Energy consumption of BSs due to state switching.

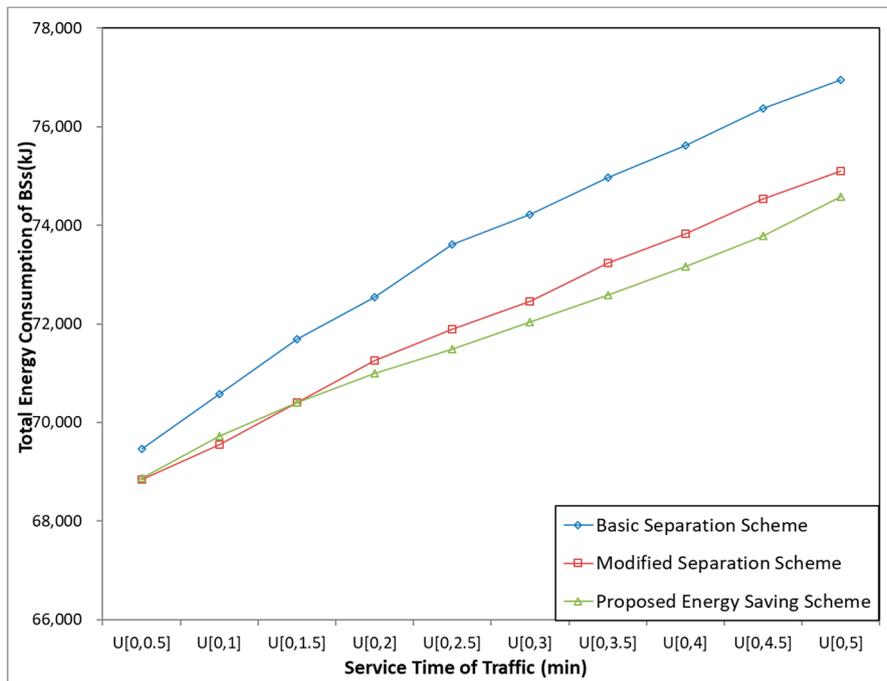


Figure 19. Total energy consumption of BSs.

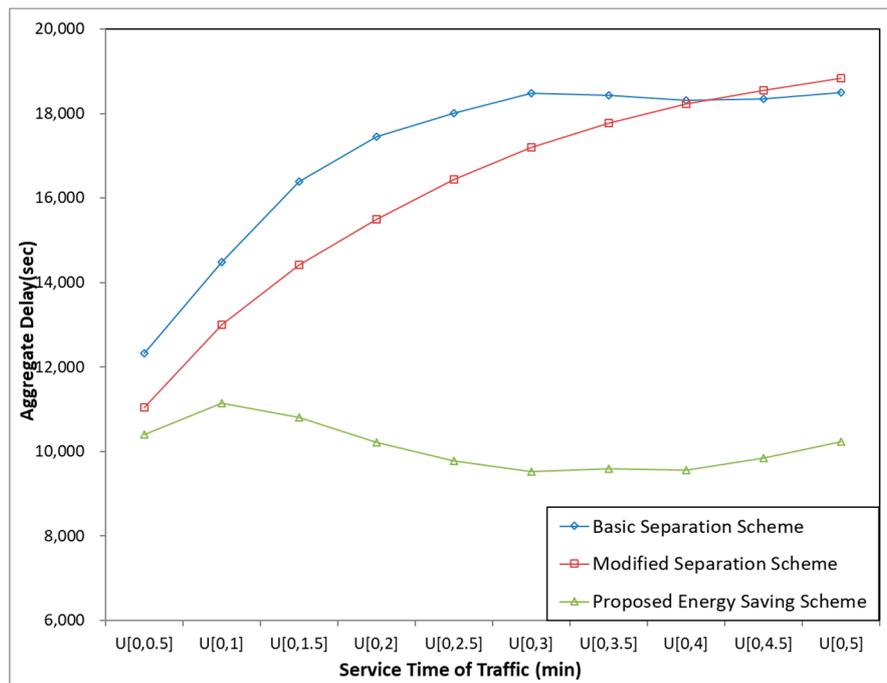


Figure 20. Aggregate delay.

5. Conclusions and Future Work

Ever-increasing energy consumption is one of the most important issues related to the problem of global warming and thus the energy consumption of mobile communication networks has drawn much attention since it takes a significant part of the total energy consumption of ICT. In this paper, we considered 5G networks with heterogeneous macro cells and small cells, where data and control planes are separated, and an MBS manages the control and low rate data traffic, while an SBS manages high rate data traffic, in contrast to the basic separation architecture where an MBS manages the control while an SBS manages data traffic. Then, an energy saving scheme for BSs was proposed, based on the state management of BSs and session management of UEs, where the state of BS is determined depending on the number of UEs that request high rate data traffic and the number of UEs that exist under the overlapping areas commonly covered by the considered BS and the neighbor BSs.

In the proposed energy saving scheme, if a BS has a high probability of activation due to higher number of UEs that request high rate data under the coverage of the BS or higher number of UEs in the overlapping area commonly covered by neighbor BSs and the considered BS, the BS moves to *sleep* state. Otherwise, the BS moves to *off* state in order to save more energy. We formulated an optimization problem for the proposed energy saving scheme, where the objective function is the sum of power consumption by an MBS, power consumption by SBSs under the coverage of a BS, and energy consumption due to switching between different states of BSs. Since the developed optimization problem is a NP-hard problem, the solution was obtained using particle swarm optimization. Numerical results showed that the proposed energy saving scheme in the modified separation architecture had better energy efficiency compared to the basic and modified separation schemes. Also, the proposed energy saving scheme had lower aggregate delay than other schemes. These results show that the proposed scheme is energy-efficient with less aggregate delay.

In our future work, we will consider the energy consumption of UEs together in a multi-macro cell environment. Since UEs consume more energy when they communicate with an MBS, since the distance between a UE and an MBS is generally much longer than that between a UE and a nearby SBS, the proposed energy saving scheme has more energy consumption of UEs. This may degrade the quality of experience (QoE) of UEs. Therefore, we will consider the energy consumption of BSs

and UEs together and draw an optimal scheme from the aspect of the sum of energy consumption of BSs and UEs, while satisfying the QoE of UEs. Also, work on how to apply the concept of separated architecture to heterogeneous networks owned by multi-operator mobile networks will be carried out in our future works.

Acknowledgments: The authors would like to thank anonymous reviewers, who greatly improved the quality of this paper. This research was supported in part by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2014R1A1A1037728).

Author Contributions: Min Wook Kang and Yun Won Chung conceived and designed the proposed scheme. Min Wook Kang and Yun Won Chung conceived and designed the simulations; Min Wook Kang performed the simulations; Min Wook Kang and Yun Won Chung analyzed the data; Yun Won Chung wrote the paper. Both authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wu, J.; Zhang, Y.; Zukerman, M. Energy-efficient base-station sleep-mode techniques in green cellular networks: A survey. *IEEE Commun. Surv. Tutor.* **2015**, *17*, 803–826. [CrossRef]
2. Webb, M.; GeSI (Global e-Sustainability Initiative). *SMART 2020: Enabling the Low Carbon Economy in the Information Age*; The Climate Group: Lambeth, London, 2008; Available online: <http://gesi.org/files/Reports/Smart%202020%20report%20in%20English.pdf> (accessed on 12 February 2017).
3. Fehske, A.; Fettweis, G.; Malmudin, J.; Biczok, G. The global footprint of mobile communications: The ecological and economic perspective. *IEEE Commun. Mag.* **2011**, *49*, 55–62. [CrossRef]
4. Hasan, Z.; Boostanimehr, H.; Bhargava, V.K. Green cellular networks: A survey, some research issues and challenges. *IEEE Commun. Surv. Tutor.* **2011**, *13*, 524–540. [CrossRef]
5. Oh, E.; Son, K.; Krishnamachari, B. Dynamic base station switching-on/off strategies for green cellular networks. *IEEE Trans. Wirel. Commun.* **2013**, *12*, 2126–2136. [CrossRef]
6. Son, K.; Kim, H.; Yi, Y.; Krishnamachari, B. Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks. *IEEE J. Sel. Areas Commun.* **2011**, *29*, 1525–1536. [CrossRef]
7. Han, T.; Ansari, N. On greening cellular networks via multicell cooperation. *IEEE Wirel. Commun.* **2013**, *20*, 82–89. [CrossRef]
8. Niu, Z.; Wu, Y.; Gong, J.; Yang, Z. Cell zooming for cost-efficient green cellular networks. *IEEE Commun. Mag.* **2010**, *48*, 74–79. [CrossRef]
9. Chang, C.Y.; Liao, W.; Hsieh, H.Y.; Shiu, D.S. On optimal cell activation for coverage preservation in green cellular networks. *IEEE Trans. Mob. Comput.* **2014**, *13*, 2580–2591. [CrossRef]
10. Niu, Z. Tango: Traffic-aware network planning and green operation. *IEEE Wirel. Commun. Mag.* **2011**, *18*, 25–29. [CrossRef]
11. Balasubramaniam, R.; Nagaraj, S.; Sarkar, M.; Paolini, C.; Khaitan, P. Cell zooming for power efficient base station operation. In Proceedings of the IEEE 9th International Wireless Communications and Mobile Computing Conference, Sardinia, Italy, 1–5 July 2013.
12. Chung, Y.L. An energy-saving small-cell zooming scheme for two-tier hybrid cellular networks. In Proceedings of the International Conference on Information Networking, Siem Reap, Cambodia, 12–14 January 2015.
13. Liu, C.; Natarajan, B.; Xia, H. Small cell base station sleep strategies for energy efficiency. *IEEE Trans. Veh. Technol.* **2016**, *65*, 1652–1661. [CrossRef]
14. Ashraf, I.; Boccardi, F.; Ho, L. SLEEP mode techniques for small cell deployments. *IEEE Commun. Mag.* **2011**, *49*, 72–79. [CrossRef]
15. Ghosh, P.; Das, S.S.; Naravaram, S.; Chandhar, P. Energy saving in OFDMA cellular systems using base-station sleep mode: 3GPP-LTE a case study. In Proceedings of the National Conference on Communications (NCC), Kharagpur, India, 3–5 February 2012; pp. 1–5.
16. Dini, P.; Miozzo, M.; Bui, N.; Baldo, N. A model to analyze the energy savings of base station sleep mode in LTE HetNets. In Proceedings of the IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing, Beijing, China, 20–23 August 2013; pp. 1375–1380.

17. Darnjanovic, A.; Montojo, J.; Wei, Y.; Ji, T.; Luo, T.; Vajapeyam, M.; Yoo, T.; Song, O.; MaUadi, D. A survey on 3GPP heterogeneous networks. *IEEE Wirel. Commun. Mag.* **2011**, *18*, 10–21. [[CrossRef](#)]
18. Saker, L.; Elayoubi, S.E.; Combes, R.; Chahed, T. Optimal control of wake up mechanisms of femtocells in heterogeneous networks. *IEEE J. Sel. Areas Commun.* **2012**, *30*, 664–672. [[CrossRef](#)]
19. Wildemeersch, M.; Quek, T.Q.S.; Slump, C.H.; Rabbachin, A. Cognitive small cell networks: Energy efficiency and trade-offs. *IEEE Trans. Commun.* **2013**, *61*, 4016–4029. [[CrossRef](#)]
20. Lee, K.S.; Kim, H.S.; Kim, Y.T.; Kim, B.H. DANCE: Small AP on/off algorithms in ultra dense wireless network. *J. Korean Inst. Commun. Inf. Sci.* **2013**, *38*, 1135–1144. [[CrossRef](#)]
21. Feng, M.; Mao, S.; Jiang, T. BOOST: Base station on-off switching strategy for energy efficient massive mimo hetnets. In Proceedings of the 35th Annual International Conference on Computer Communications, IEEE INFOCOM 2016, San Francisco, CA, USA, 10–14 April 2016; pp. 1395–1403.
22. Cai, S.; Che, Y.; Duan, L.; Wang, J.; Zhou, S.; Zhang, R. Green 5G heterogeneous networks through dynamic small-cell operation. *IEEE J. Sel. Areas Commun.* **2016**, *34*, 1103–1115. [[CrossRef](#)]
23. Antonopoulos, A.; Kartsakli, E.; Bousia, A.; Alonso, L.; Verikoukis, C. Energy-efficient infrastructure sharing in multi-operator mobile networks. *IEEE Commun. Mag.* **2015**, *53*, 242–249. [[CrossRef](#)]
24. Oikonomakou, M.; Antonopoulos, A.; Alonso, L.; Verikoukis, C. Cooperative base station switching off in multi-operator shared heterogeneous network. In Proceedings of the 2015 IEEE Global Communications Conference, San Diego, CA, USA, 6–10 December 2015; pp. 1–6.
25. Lee, S.; Moon, S.; Yi, Y. On greening cellular networks by sharing base stations: A game-theoretic approach. In Proceedings of the 9th EAI International Conference on Performance Evaluation Methodologies and Tools, Berlin, Germany, 14–16 December 2015; pp. 87–94.
26. Ishii, H.; Kishiyama, Y.; Takahashi, H. A novel architecture for LTE-B: C-plane/U-plane split and phantom cell concept. In Proceedings of the IEEE GLOBECOM Workshops, Anaheim, CA, USA, 3–7 December 2012; pp. 624–630.
27. Astely, D.; Dahlman, E.; Fodor, G.; Parkvall, S.; Sachs, J. LTE release 12 and beyond. *IEEE Commun. Mag.* **2013**, *51*, 154–160. [[CrossRef](#)]
28. Mukherjee, S.; Ishii, H. Energy efficiency in the phantom cell enhanced local area architecture. In Proceedings of the 2013 IEEE Wireless Communications and Networking Conference (WCNC), Shanghai, China, 7–10 April 2013; pp. 1267–1272.
29. Xu, X.; He, G.; Zhang, S.; Chen, Y.; Xu, S. On functionality separation for green mobile networks: Concept study over LTE. *IEEE Commun. Mag.* **2013**, *51*, 82–90. [[CrossRef](#)]
30. Sakaguchi, K.; Sampei, S.; Shimodaira, H.; Rezagah, R.; Tran, G.K.; Araki, K. Cloud cooperated heterogeneous cellular networks. In Proceedings of the 2013 International Symposium on Intelligent Signal Processing and Communication Systems, Naha, Japan, 12–15 November 2013; pp. 787–791.
31. Feng, M.; Mao, S.; Jiang, T. Base station on-off switching in 5G wireless networks: Approaches and challenges. *IEEE Wirel. Commun.* **2017**, *24*, 46–54. [[CrossRef](#)]
32. Wang, Z.; Zhang, W. A separation architecture for achieving energy-efficient cellular networking. *IEEE Trans. Wirel. Commun.* **2014**, *13*, 3113–3123. [[CrossRef](#)]
33. Zhang, S.; Gong, J.; Zhou, S.; Niu, Z. How many small cells can be turned off via vertical offloading under a separation architecture? *IEEE Trans. Wirel. Commun.* **2015**, *14*, 5440–5453. [[CrossRef](#)]
34. Chiaraviglio, L.; Amorosi, L.; Cartolano, S.; Blefari-Melazzi, N.; Dell’Olmo, P.; Shojafar, M.; Salsano, S. Optimal superfluid management of 5G networks. In Proceedings of the 3rd IEEE Conference on Network Softwarization, Bologna, Italy, 3–7 July 2017.
35. Marotta, A.; Zola, E.; D’Andreagiovanni, F.; Kassler, A. A fast robust optimization-based heuristic for the deployment of green virtual network functions. *J. Netw. Comput. Appl.* **2017**, *95*, 42–53. [[CrossRef](#)]
36. Marotta, A.; D’Andreagiovanni, F.; Kassler, A.; Zola, E. On the energy cost of robustness for green virtual network function placement in 5G virtualized infrastructures. *Comput. Netw.* (in press). [[CrossRef](#)]
37. Salsano, S.; Chiaraviglio, L.; Blefari-Melazzi, N.; Parada, C.; Fontes, F.; Mekuria, R.; Griffioen, D. Toward superfluid deployment of virtual functions: Exploiting mobile edge computing for video streaming. In Proceedings of the First International Workshop on Softwarized Infrastructures for 5G and Fog Computing, Genoa, Italy, 4–8 September 2017.
38. Shojafar, M.; Chiaraviglio, L.; Blefari-Melazzi, N.; Salsano, S. P5G: A bio-inspired algorithm for the superfluid management of 5G Networks. In Proceedings of the 18th IEEE GLOBECOM, Singapore, 4–8 December 2017.

39. Amaldi, E.; Capone, A.; Malucelli, F.; Mannino, C. Optimization problems and models for planning cellular networks. In *Handbook of Optimization in Telecommunication*; Resende, M., Pardalos, P., Eds.; Springer: New York, NY, USA, 2006; Volume 31, pp. 917–939.
40. Amaldi, E.; Capone, A.; Malucelli, F.; Mannino, C. *Optimization Problems and Models for Planning Cellular Networks*; Springer: New York, NY, USA, 2006; Volume 31, pp. 879–901.
41. Kennington, J.; Olinick, E.; Rajan, D. *Wireless Network Design: Optimization Models and Solution Procedures*; Springer: New York, NY, USA, 2010.
42. D’Andreagiovanni, F. Pure 0–1 programming approaches to wireless network design. *4OR Q. J. Oper. Res.* **2012**, *10*, 211–212. [[CrossRef](#)]
43. D’Andreagiovanni, F.; Gleixner, A. Towards an accurate solution of wireless network design problems. In Proceedings of the International Symposium on Combinatorial Optimization, LNCS, Vietri sul Mare, Italy, 16–18 May 2016; Volume 9849, pp. 135–147.
44. D’Andreagiovanni, F.; Mannino, C.; Sassano, A. Negative cycle separation in wireless network design. In Proceedings of the 5th International Conference on Network Optimization, LNCS, Hamburg, Germany, 13–16 June 2011; Volume 6701, pp. 51–56.
45. D’Andreagiovanni, F.; Mannino, C.; Sassano, A. GUB covers and power-indexed formulations for wireless network design. *Manag. Sci.* **2012**, *59*, 142–156. [[CrossRef](#)]
46. D’Andreagiovanni, F. On improving the capacity of solving large-scale wireless network design problems by genetic algorithms. In Proceedings of the Applications of Evolutionary Computation, LNCS, Torine, Italy, 27–29 April 2011; Volume 6625, pp. 11–20.
47. D’Andreagiovanni, F.; Mett, F.; Nardin, A.; Pulaj, J. Integrating LP-guided variable fixing with MIP heuristics in the robust design of hybrid wired-wireless FTTx access networks. *Appl. Soft Comput.* **2017**. [[CrossRef](#)]
48. Dely, P.; D’Andreagiovanni, F.; Kassler, A. Fair optimization of mesh-connected wlan hotspots. *Wirel. Commun. Mob. Comput.* **2013**, *15*, 924–946. [[CrossRef](#)]
49. Eisenblatter, A.; Geerdes, H. Capacity optimization for UMTS: Bounds and benchmarks for interference reduction. In Proceedings of the IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications, Cannes, France, 15–18 September 2008.
50. Gendron, B.; Scutellà, M.G.; Garroppo, R.G.; Nencioni, G.; Tavanti, L. A branch-and-Benders-cut method for nonlinear power design in green wireless local area networks. *Eur. J. Oper. Res.* **2016**, *255*, 151–162. [[CrossRef](#)]
51. Mannino, C.; Rossi, F.; Smriglio, S. The network packing problem in terrestrial broadcasting. *Oper. Res.* **2016**, *54*, 611–626. [[CrossRef](#)]
52. Sawaya, N.J.; Elhedhli, S. A nested benders decomposition approach for optimal W-CDMA telecommunication network planning. *Nav. Res. Logist.* **2010**, *57*, 519–539. [[CrossRef](#)]
53. Gu, F.; Liu, L.H.; Cheung, M.Y.; Xie, S. Optimal WCDMA network planning by multi objective evolutionary algorithm with problem-specific genetic operation. *Knowl. Inf. Syst.* **2015**, *45*, 679–703. [[CrossRef](#)]
54. Koutitas, G. Green network planning of single frequency networks. *IEEE Trans. Broadcast.* **2010**, *56*, 541–550. [[CrossRef](#)]
55. Zakrzewska, A.; D’Andreagiovanni, F.; Ruepp, S.; Berger, M.S. Biobjective optimization of radio access technology selection and resource allocation in heterogeneous wireless networks. In Proceedings of the 2013 11th International Symposium on Modeling & Optimization in Mobile, Ad Hoc & Wireless Networks, Tsukuba Science City, Japan, 13–17 May 2013.
56. Bauschert, T.; Büsing, C.; D’Andreagiovanni, F. Network planning under demand uncertainty with robust optimization. *IEEE Commun. Mag.* **2014**, *52*, 178–185. [[CrossRef](#)]
57. Büsing, C.; D’Andreagiovanni, F. New results about multi-band uncertainty in robust optimization. In Proceedings of the SEA 2012, Springer Lecture Notes in Computer Science 7276, Bordeaux, France, 7–9 June 2012.
58. D’Andreagiovanni, F. Revisiting wireless network jamming by SIR-based considerations and multiband robust optimization. *Optim. Lett.* **2015**, *9*, 1495–1510. [[CrossRef](#)]
59. D’Andreagiovanni, F.; Krolikowski, J.; Pulaj, J. A fast hybrid primal heuristic for multiband robust capacitated network design with multiple time periods. *Appl. Soft Comput.* **2015**, *26*, 497–507. [[CrossRef](#)]
60. Garroppo, R.G.; Nencioni, G.; Scutellà, M.G.; Tavanti, L. Robust optimisation of green wireless LANs under rate uncertainty and user mobility. *Electr. Notes Discret. Math.* **2016**, *52*, 221–228. [[CrossRef](#)]

61. Heikkinen, T.; Prekopa, A. Optimal power control in a wireless network using stochastic link coefficients. *Nav. Res. Logist.* **2004**, *52*, 178–192. [[CrossRef](#)]
62. Olinick, E.; Rosenberger, J. Optimizing revenue in CDMA networks under demand uncertainty. *Eur. J. Oper. Res.* **2008**, *186*, 812–825. [[CrossRef](#)]
63. Kang, M.W.; Chung, Y.W. A novel power saving scheme for base stations in 5G network. *Int. J. Mob. Device Eng.* **2017**, *1*, 29–36. [[CrossRef](#)]
64. Shahid, A.; Aslam, S.; Lee, L.G. A decentralized heuristic approach towards resource allocation in femtocell networks. *Entropy* **2013**, *15*, 2524–2547. [[CrossRef](#)]
65. Hu, W.; Liang, H.; Peng, C.; Du, B.; Hu, Q. A hybrid chaos-particle swarm optimization algorithm for the vehicle routing problem with time window. *Entropy* **2013**, *15*, 1247–1270. [[CrossRef](#)]
66. Conforti, M.; Cornuejols, G.; Zambelli, G. *Integer Programming*; Springer: New York, NY, USA, 2014.
67. Garey, M.R.; Johnson, D.S. *Computers and Intractability: A Guide to the Theory of NP-Completeness*; W.H. Freeman and Co.: London, UK, 1979.
68. Papadimitriou, C.H.; Steiglitz, K. *Combinatorial Optimization: Algorithms and Complexity*; Dover Publications: New York, NY, USA, 1998.
69. Campbell, S.E. Multi-Scale Path Planning for Reduced Environmental Impact of Aviation. Ph.D. Thesis, University of Illinois at Urbana-Champaign, Champaign, IL, USA, 2010.
70. Melo, G. Not quite the same: Identity constraints for the web of linked data. In Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence, Berkeley, CA, USA, 14–18 July 2013.
71. Gavanelli, M.; Milano, M.; Bragaglia, S.; Chesani, F.; Marengo, E.; Cagnoli, P. Multi-criteria optimal planning for energy policies in CLP. In Proceedings of the 29th Italian Conference on Computational Logic, Torino, Italy, 16–18 June 2014.
72. Zhang, W.; Kwak, K.S.; Feng, C. Network selection algorithm for heterogeneous wireless networks based on multi-objective discrete particle swarm optimization. *KSII Trans. Internet Inf. Syst.* **2012**, *6*, 1802–1814. [[CrossRef](#)]
73. Wu, J.; Wong, E.W.M.; Guo, J.; Zukerman, M. Performance analysis of green cellular networks with selective base-station sleeping. *Perform. Eval.* **2017**, *111*, 17–36. [[CrossRef](#)]
74. Auer, G.; Giannini, V.; Dessel, C.; Godor, I.; Skillermark, P.; Olsson, M.; Imran, M.; Sabella, D.; Gonzalez, M.; Blume, O.; et al. How much energy is needed to run a wireless network? *IEEE Wirel. Commun.* **2011**, *18*, 40–49. [[CrossRef](#)]
75. Guo, X.; Niu, Z.; Zhou, S.; Kumar, P.R. Delay-constrained energy-optimal base station sleeping control. *IEEE J. Sel. Areas Commun.* **2016**, *34*, 1073–1085. [[CrossRef](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).