IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received April 9, 2019, accepted May 7, 2019, date of publication May 20, 2019, date of current version June 3, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2917722

Density-Aware, Energy- and Spectrum-Efficient Small Cell Scheduling

SHAHRAM MOLLAHASANI, (Student Member, IEEE), AND ERTAN ONUR, (Member, IEEE)
Department of Computer Engineering, Middle East Technical University, 06800 Ankara, Turkey

Corresponding author: Shahram Mollahasani (shahram.mollahasani@ceng.metu.edu.tr)

This work was supported in part by the TÜBİTAK under Grant 215E127.

ABSTRACT Future mobile networks have to be densified by employing small cells to handle the upsurge in traffic load. Although the amount of energy each small cell consumes is low, the total energy consumption of a large-scale network may be enormous. To enhance energy efficiency, we have to adapt the number of active base stations to the offered traffic load. Deactivating base stations may cause coverage holes, degrade the quality of service and throughput while redundant base stations waste energy. That is why we have to adapt the network to an effective density. In this paper, we show that achieving an optimal solution for adapting the density of base stations to the demand is NP-hard. We propose a solution that consists of two heuristic algorithms: a base station density adaptation algorithm and a cell-zooming algorithm that determines which base stations must be kept active and adapts transmit power of base stations to enhance throughput, energy, and spectral efficiency. We employ a multi-access edge cloud for taking a snapshot of the network state in nearly real time with a wider perspective and for collecting network state over a large area. We show that the proposed algorithm conserves energy up to 12% while the spectral efficiency and network throughput can be enhanced up to 30% and 26% in comparison with recent works, respectively.

INDEX TERMS 5G mobile networks, densification, density-aware networking, energy-efficiency, green networks, multi-access edge cloud (MEC), self-organizing networks.

I. INTRODUCTION

With the emergence of multimedia applications such as online games, video streaming, and social networks, users expect a faster network with a higher throughput. To satisfy users and their increasing traffic demands, we need to enhance the network throughput by 1000 times in the next 15 years [1]. Among several solutions for this purpose, it is recommended to enhance the spectral efficiency [2]. Although spectral efficiency and capacity can be improved by increasing density of base stations (BSs), energy consumption will also increase [3]. To keep the energy consumption at the same level, energy efficiency is also needed to be increased by 1000 times in future mobile networks. Jointly satisfying these conflicting goals is one of the most significant challenges of future networks.

Until 2016, over 14 million small cells are deployed in Long Term Evolution (LTE) networks and this number has increased for about 270 percent in 2017 [4]. In heterogeneous networks (HetNets), by employing small cells in addition to macro cells, spectral and energy efficiency (EE) can be enhanced simultaneously [5]. By increasing density of small cells, the distance between base stations and user equipment (UE) can be reduced, which can enhance network area throughput and reduce power consumption. However, increasing density of small cells implies more infrastructure and hardware deployment, which increases the overall circuit energy consumption in the network [6].

Because of mobile/nomadic BSs such as drone cells, user-controlled based BSs (indoor small cells) and sleep scheduling of BSs, the network density may dynamically change in time and space. On the one hand, by increasing the density of small cells, we need to adapt BS density to the network condition for maximizing energy efficiency in the network. On the other hand, future networks need to be smart to adapt themselves to BS density for maintaining the quality of service (QoS) under various conditions when the density of BSs dynamically changes. For instance, when a BS is turned off in a HetNet, the associated traffic load of the deactivated BS needs to be transferred over to other cells. The demand can

The associate editor coordinating the review of this manuscript and approving it for publication was Oussama Habachi.

be different at each cell because of heterogeneity. Therefore, the impact of turning off a BS on the neighboring cells have to be considered in an adaptive fashion. Turning off a BS may also cause coverage holes that requires the assistance of neighboring cells for providing service to blind spots. In case of network expansion or occasional events (such as a football match), the number (density) of BSs may need to be increased permanently or temporarily. Human intervention may not be possible in a large-scale HetNet. Therefore, future networks need to be equipped with flexible algorithms that can adapt network parameters such as coverage or transmit power to the density of BSs to dynamically react over any changes in the network state.

The main idea behind our approach is adapting density of BSs to the network parameters such as coverage and transmit power to provide a highly flexible scheduling model which can enhance energy and spectral efficiency and maintain QoS continuously in the whole network for different conditions. We assume the QoS can be maintained when the minimum throughput by each UE can be satisfied by the network. In this paper, we define a joint heuristic sleep scheduling and power allocation algorithm for saving energy while satisfying the throughput requirement of users and maintaining the coverage in dynamic heterogeneous networks by considering the density of BSs and employing MEC in the network architecture. The main advantages of this work with respect to old energy efficient techniques are employing MECs in the network architecture and providing a self-organized network which can adapt network parameters including BS transmit power and density of BSs to the network state, that can be categorized as one of the basic requirements of 5G networks. Employing MEC platforms provide higher flexibility, processing power and support multi-tenancy in cellular networks, and can provide self-adaptability in dynamic networks where in addition of density of UEs, density of BSs can be also changed. The overall contributions of this paper are as follows.

- We formulate density-, energy- and spectrum-aware base station scheduling problem (DESAS) in Section II and we show that providing an optimal solution to this problem for adapting density of BSs to network parameters such as coverage and transmit power to reduce energy consumption and enhance spectral efficiency in the network is NP-hard.
- We propose a heuristic solution named as BS density and power adaptation algorithm (BDPA) that consists of two sub-algorithms named by BS density adaptation (BDA) and power adaptation (PA) in Section III. These algorithms can jointly reduce the amount of power consumption in the network through minimizing the number of active BSs with respect to the cells' load and can enhance the throughput and the coverage in the network by applying a cell-zooming technique in each cell in a distributed manner by adapting transmit power of BSs based on channel conditions and effective density.

TABLE 1. List of symbols.

Symbol	Definition	Unit
α	A fit adjustment on a curve	
α_P	A normalized coefficient for power allocation	
α_S	A normalized coefficient for spectral efficiency	
β	A shift to adjust peak time	
Δ	The power for adjusting transmission power	W
δ_j	Load factor of BS j	%
λ_b	Density of small BSs	$\frac{BS}{Km^2}$
σ^2	Noise power	dBm
τ_j	Total number of available RBs in BS j	
τ_{ij}^*	The required number of RBs by UE i from BS j	
ϕ	Spectrum efficiency	b/s/Hz
A_j	Cell's activity factor	%
A_{mj}	Activity ratio of the macro cell responsible for BS j	%
b_τ	Bandwidth of a RB	Hz
B_j	Bandwidth of BS j	Hz
B_{ij}^*	Min required BW by UE i from BS j	Hz
B_j^*	Min required BW of BS j	Hz
C-RAN	Cloud radio access network	
CQI	Channel quality indicator	
D	UEs' demand	Mbps
D_{avg}	Average bitrate in the predefined time interval	Mbps
g_{ij}	Channel gain between UE i and its associated BS j	dBm
i	Index of UE	
ICI	Inter cell interference	
I_{ik}	Interference received by UE i from BS k	dB
j	Index of BS	
MCS	Modulation and coding scheme	
MEC	Multi access edge cloud	
N_{macro}	The number of macro cells	
N_{small}	The number of small cells	
P	Current allocated power by a BS	W
P_{avg}	Average of allocated power to a BS	W
P_{macro}	Power consumed by a macro BS	W
P_{min}	Minimum required transmission power by a BS	W
P_{max}	Maximum transmission power of a BS	W
P_{small}	Power consumed by a small BS	W
P_{ij}	Downlink transmit power assigned by BS j to UE i	dBm
P_{ij}^*	Minimum required power by UE i from BS j	W
P_{τ_j}	The power allocated by BS j to a resource block	dBm
P_j	Maximum transmit power of BS j	dBm
P_0	Fixed operational power	W
$P_C(T)$	Coverage probability when received SINR is T	%
QoS	Quality of service	
R	Average rate per UE	bps
R_{ij}^*	The required throughput by UE i from BS j	bps
R_j	Total transmission rate in BS j	bps
$SINR_{ij}$	SINR received by UE i from BS j	dB
TTI	Transmission time interval	ms
T_{CQI}	Threshold value for minimum required CQI level	
T_{SINR}	Threshold value for minimum received SINR	dB
T_{cov}	Threshold value for minimum coverage probability	%
U_j	Utility function	%

- We define a framework for future network infrastructures that employs MECs that facilitate reduction of latency and access a larger amount of data about the state of the network for management and control purposes. On one side, a global optimization is nearly impossible since the DESAS problem is an NP-hard problem. On the other side, solutions carried out individually by base stations lack the required information for achieving an optimal solution. Therefore, we assert that solutions designed on MECs by employing a divide-and-conquer approach are appropriate candidates as we discuss in Section III-B.

We validate the BDPA algorithm and compare it with the low-power wake-up radio (LP-WUR) algorithm [7] where

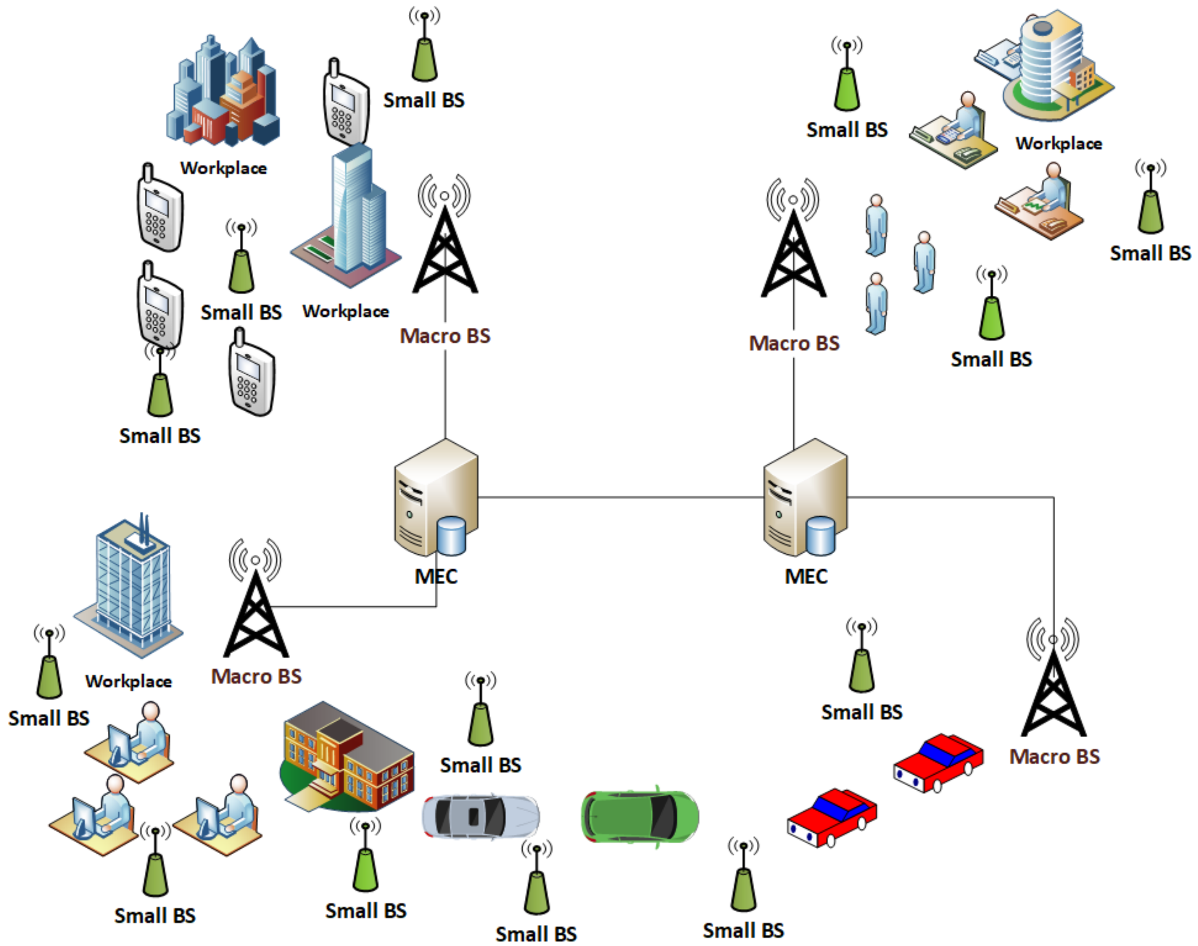

FIGURE 1. Illustration of a network architecture where multi-access edge clouds are employed for self organization.

base stations activated or deactivated by UEs' request and the cooperative inter-cell interference control (C-ICIC) algorithm [8], which energy consumption is optimized by collaborating among neighbor cells, in Section IV. We evaluate and categorized other power management and energy efficient models in Section V to highlight the advantages of the BDPA algorithm with respect to them. We conclude our paper in section VI.

II. DENSITY-, ENERGY- AND SPECTRUM-AWARE BASE STATION SCHEDULING PROBLEM (DESAS)

We define the system model and the density-, energy- and spectrum-aware base station scheduling problem (DESAS) based on the system model presented in this section. The nomenclature used in this paper is presented in Table 1.

The set $\mathbb{M} = \{1, 2, \ldots, T, \ldots, M\}$ represents M BSs including T macro BSs and S small cells. The set $\mathbb{N}$ represents N UEs. BSs and UEs are uniform randomly distributed in a two-tier LTE-like environment as shown in Fig.1. We assume that no power control is employed in the system and we only focus on the downlink transmission. In LTE-systems [2], BSs schedule their users in $1\,ms$ subframes. A UE is associated to the BS from which it receives the strongest Signal-to-Interference-plus-Noise Ratio (SINR). The experienced SINR by UE i from its serving BS j is

$$SINR_{ij} = \frac{g_{ij}P_{ij}}{\sigma^2 + \sum_{k=1,k\neq j}^{M} g_{ik}P_{ik}} \text{dB}, \tag{1}$$

where P_{ij} and g_{ij} are the downlink transmit power assigned by BS j to UE i and the channel gain between UE i and BS j, respectively. The amount of noise power is σ^2 and $\sum_{k=1,k\neq j}^{M} g_{ik}P_{ik} = \sum_{k\neq j} I_{ik}$ is the total received interference by UE i from other BSs.

By considering power is distributed uniformly among resource blocks (RBs) [9], the amount of allocated power in W to each RB can be obtained as:

$$P_{\tau_j} = \frac{P_j}{\tau_j}, \tag{2}$$

where P_j is the maximum transmit power of BS j and τ_j is the total number of available RBs in BS j. Then, the number of RBs that has to be allocated to UE i from BS j to satisfy the user's required throughput R^*_{ij} is

$$\tau^*_{ij} = \frac{R^*_{ij}}{b_\tau log_2(1 + SINR_{ij})}, \tag{3}$$

where the denominator provides the maximum achievable data rate and b_τ is the bandwidth of a RB in the network, which is 180 kHz in LTE [10]. By obtaining P_{τ_j} and τ^*_{ij} from (2) and (3) respectively, we can calculate the minimum amount of power in W that has to be allocated by BS j to UE i as

$$P^*_{ij} = P_{\tau_j}\tau^*_{ij}. \tag{4}$$

The dynamic power consumption of BS j can then be computed as

$$P_j = \beta \sum_{i\in\mathbb{N}} P_{ij} + P_0, \tag{5}$$

where the constant $\beta \geq 1$ is the inverse of power amplifier efficiency to evaluate amplifier losses and P_0 is the fixed operational power consumed for backhaul signaling and cooling including losses. Although the most influential component of energy consumption in macro cells is the static power, it is not the case for small cells. In small cells, the main part of energy is consumed by the radio transceiver unit due to the absence of cooling system and low-power amplifier [11]. Therefore, we did not analyze the effect of P_0 in this paper. By considering this fact, we do not need to turn off the small cells completely. While the transmitter of BSs (which consume the major part of energy in small cells) will be turned off completely, other small cells' modules remain active for rapid response to the network condition (we do it per TTI which implicitly sets the timescale to 1 ms). That is why P_0 is considered in the problem formulations. Generally, P_0 is around 500 W and 15 W for macro and small cells, respectively [10]. This model can be expanded by including switching power consumption which is out of scope of this work [12], [13]. Because in our opponents (LP-WUR and C-ICIC) switching power is also not considered. Therefore, for fairness, we did not apply the switching power effect over our power consumption model in this work and we set it a future work item.

The QoS provided to the users can be increased if the amount of interference is lowered. Therefore, we aim in the DESAS at enhancing the received SINR to reach a required threshold T_{SINR},

$$\frac{g_{ij}P_{ij}}{\sigma^2 + \sum_{k\neq j} I_{ik}} \geq T_{SINR}.$$

Maximization of energy- and spectral-efficiency can be achieved through minimization of power consumption and the number of required RBs while satisfying UEs' throughput requirements. Therefore, we formulate the DESAS problem given an initial x_{ij} assignment as

$$\text{finding } x_{ij}, P_{ij}, z_j$$

$$\text{to minimize } \sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{M}} x_{ij}\left(\alpha_P P_{ij} + \alpha_S \tau_{ij}\right) \tag{6a}$$

$$\text{subject to } \sum_{j\in\mathbb{M}} x_{ij} = 1, \quad \forall i \in \mathbb{N}, \tag{6b}$$

$$\sum_{i\in\mathbb{N}} x_{ij}\tau_{ij} \leq \tau_j, \quad \forall j \in \mathbb{M}, \tag{6c}$$

$$\tau_{ij} \geq \tau^*_{ij}, \quad \forall i \in N, \ \forall j \in \mathbb{M}, \tag{6d}$$

$$x_{ij} \leq z_j, \tag{6e}$$

where

$$x_{ij} = \begin{cases} 1 & \text{if UE } i \text{ is assigned to BS } j \text{ and } z_j = 1 \\ 0 & \text{otherwise} \end{cases}$$

The objective of the DESAS problem is to minimize energy consumption and to reduce the number of required RBs by finding UEs should be assigned to which BSs (x_{ij}), how much power need to be assigned by BSs to their associated UEs (P_{ij}) and which BSs should be kept active (z_j) while maintaining the throughput and satisfying UEs requirements. In (6a), we employ normalization coefficients α_P and α_S defined as $w_P\rho_P$ and $w_S\rho_S$ for energy- and spectral-efficiency, respectively [14]; w is a weight for adding a preference among energy- and spectral-efficiency, and ρ is a normalization factor. DESAS is a UE-to-BS assignment problem where each UE has to be served by only one BS (6b). To make sure we can assign a UE to a BS, the corresponding BS should have enough available capacity to satisfy the requested RBs by UE (6c) and the minimum required RBs of users have to be satisfied (6d). A UE can be assigned to a BS only when the BS is active (6e).

The DESAS problem can be reduced to the generalized assignment problem with a space complexity of $O(2^{N\times M} + 2^{N\times M} + 2^M)$ which is an NP-hard problem [15]. By using classical knapsack terminology, we can describe the generalized assignment problem (GAP) [15] as assigning item i to knapsack j with a profit value of P_{ij}. When item i is assigned to knapsack j it has a weight and each knapsack has a capacity. In the DESAS problem, $(\alpha_P P_{ij} + \alpha_S \tau_{ij})$ can be considered as profit value of assigning UE i to BS j, τ_{ij} is the weight of UE i in case it is assigned to BS j and the capacity of BS j is represented as τ_j. Since the DESAS problem is an NP-hard problem, a global solution to the DESAS is not feasible and practicable. We consequently resort to dividing the DESAS into two separate problems with a lower complexity that we present in the next two sections. These two subproblems are briefly:

- *BS density adaptation (BDA) problem* defines how density of BSs can be controlled based on network capacity and traffic load to conserve energy. By solving the BDA problem, we identify which BSs can be turned off while the minimum required throughput can still be satisfied.
- *Power adaptation (PA) problem* defines how the transmission power of BSs can be adapted by considering the density of BSs while maintaining coverage.

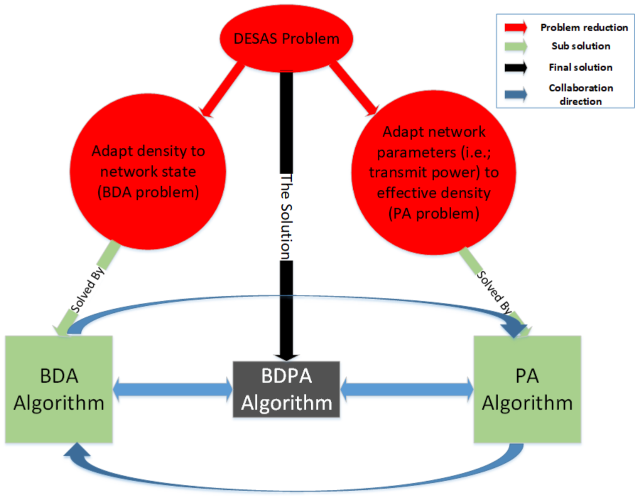

FIGURE 2. **The overall methodology.**

We firstly formulate these problems to evaluate their complexities. After that, we propose heuristic solutions to them. The overall methodology of this paper is presented in Fig. 2.

A. BS DENSITY ADAPTATION (BDA) PROBLEM

Let the minimum number of RBs that have to be allocated to UE i by BS j be τ^*_{ij}. Then, the minimum required bandwidth of BS j to satisfy its associated UEs is denoted as B^*_j (Hz) becomes

$$B^*_j = \sum_{i \in \mathbb{N}} B^*_{ij} = \sum_{i \in \mathbb{N}} \tau^*_{ij} b_\tau. \tag{7}$$

By adapting density of BSs, network resources can be used more efficiently by balancing load among cells and then redundant BSs can be turned off. We have to control BS density for maximizing energy efficiency while preserving throughput. The number of active BSs can be minimized by finding out the whether or not a BS should be active (z_j) subject to QoS constraints. Therefore, we define the BDA problem as

$$\text{finding} \quad x_{ij}, z_j$$

$$\text{to minimize} \quad \sum_{j \in \mathbb{M}} z_j \tag{8a}$$

$$\text{subject to} \quad \sum_{i \in \mathbb{N}} x_{ij} B_{ij} \leq B_j, \quad \forall j \in \mathbb{M}, \tag{8b}$$

$$x_{ij} B_{ij} \geq B^*_{ij}, \quad \forall j \in \mathbb{M}, \tag{8c}$$

$$\sum_{j \in \mathbb{M}} x_{ij} = 1, \quad \forall i \in \mathbb{N}, \tag{8d}$$

$$\text{given} \quad x_{ij} = \begin{cases} 1 & \text{if UE } i \text{ is assigned to BS } j \text{ and } z_j = 1 \\ 0 & \text{otherwise} \end{cases} \tag{8e}$$

In the BDA problem, we try to minimize the number of active BSs (8a) by making sure that the amount of bandwidth allocated to UEs are always smaller than the bandwidth (8b) allocated to BSs. In other words, the minimum required bandwidth by each UE must be satisfied by its corresponding BS (8c). UE i will initially be assigned to BS j if the amount of received SINR value is the highest with respect to its neighbor cells ($x_{ij} = 1$). This is carried out for only initializing the system model and after that UE assignment will be considered as the optimization value in this paper.

The BDA problem is a binary integer linear program (BILP) with a search space complexity of $O(2^{N \times M} + 2^M)$, that is NP-hard [16], [17] with a smaller search space complexity compared to the DESAS problem. Turning off or on a BS impacts other network parameters such as available bandwidth in other cells or received interference. In case the effect of activation and deactivation of BSs is considered constant, this problem can be reduced to the well-known bin packing problem [18]. In the bin packing problem, the goal is to minimize the number of required bins to be used for packing objects with different sizes and values. In the BDA problem, objects are UEs with different bandwidth requirements (values) and bins are active BSs whose count we try to minimize. Therefore, the BDA problem can be reduced to a multi-dimensional bin packing problem that is NP-hard. To solve this problem, we need to define a utilization factor for each cell to predict network behavior in case of activation and deactivation of a BS. The corresponding utilization ratio will be introduced in the next section.

B. POWER ADAPTATION (PA) PROBLEM

In the PA problem, we want to adapt transmit power by considering the effective BS density to maintain network coverage and to prevent coverage holes while satisfying UEs' traffic requirements. This problem can be formulated as the minimization of overall power which is allocated by BS j to its associated UEs.

Assume the overall density of small BSs and density of UEs in the network are represented as λ_b and λ_u, respectively where generally $\lambda_u \geq \lambda_b$. We assume a UE is served by the BS that provides the highest SINR value. The bandwidth per user depends on the density of UEs (λ_u) and the size of cells. For the sake of simplicity, we assume the whole bandwidth (B) is divided among UEs homogeneously; i.e., $B_u = B\frac{\lambda_b}{\lambda_u}$. Therefore, the average rate per UE i is

$$R_{ij} = B_u \mathbb{E}\left[log_2(1 + SINR_{ij})\right]. \tag{9}$$

We define the coverage probability $P_c(T_{SINR})$ as the probability of the received $SINR_{ij}$ by UE i from its closest BS j to be greater than a threshold value T_{SINR},

$$P_c(T_{SINR}) = \texttt{Prob}\left(SINR_{ij} > T_{SINR}\right).$$

In the PA problem, we want to find a solution to maximize the energy efficiency while enhancing the spectral efficiency (ϕ) without degrading the provided QoS level in terms of network coverage and UEs' throughput. To maintain the network coverage and satisfy the UEs' throughput requirement, the maximization of energy efficiency can be converted to the minimization of overall power consumption including the transmit power (P_{ij}) for all UEs and the operational power

(P_0) in the whole network. Given an initial UE-to-BS assignment x_{ij}, we formulate the PA problem as

$$\begin{aligned}
&\text{finding} \quad x_{ij}, P_{ij} \\
&\text{to minimize} \quad \sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{M}} x_{ij}\left((P_{ij}+P_0)\right) && \text{(10a)}\\
&\text{subject to} \quad P_c(P_{ij}) \geqslant (1-\epsilon)P_c^{max}, && \text{(10b)}\\
&\qquad R_{ij}(P_{ij}) \geqslant min\left\{R, \phi^{max}B_u(\lambda_b)\right\}, && \text{(10c)}\\
&\qquad x_{ij} \in \{0,1\}, \quad \forall i\in\mathbb{N},\ \forall j\in\mathbb{M}, && \text{(10d)}\\
&\qquad \sum_{j\in\mathbb{M}} x_{ij} = 1, \quad \forall i\in\mathbb{N}, && \text{(10e)}
\end{aligned}$$

The objective is to minimize the overall power consumption in the network given x_{ij} where $x_{ij} = 1$ if UE i is connected to BS j, and zero otherwise (10d), and each UE can be connect to just one BS (10e). We need to make sure that the probability of coverage for all UEs should be always higher than an acceptable outage value if the power allocated by BS j to UE i is set to P_{ij}. By considering P_c^{max} as the probability of coverage when there is no interference, we want $P_c(P_{ij})$ (coverage probability in practice) to be greater than ϵ less of P_c^{max} to achieve the required coverage and to satisfy the throughput requirement of UEs.

By considering an interference limited environment, we assume the maximum achievable spectral efficiency is ϕ^{max}. When the bandwidth allocated to a UE in a network with density λ_b is $B_u(\lambda_b)$, the maximum achievable data rate becomes $\phi^{max}B_u(\lambda_b)$. This is the upper bound for the UEs' demand rate, R. Then, $R_i(P_j)$ is the achievable rate by UE i from its associated BS j when the transmission power of the BS is equal to P_j. It must always be equal or larger than the UEs' demand rate R (10c).

The PA problem, similar to the DESAS problem, can be reduced to the GAP. The overall power that is going to be assigned to UEs is the profit value, allocated rate to UE i ($R_i(P_j)$) when transmit power is set to P_j is the weight of assigned power and the maximum achievable data rate by BS j ($\phi^{max}B_u(\lambda_b)$) can be considered as the capacity in the PA problem. Therefore, the PA problem can be reduced to generalized assignment problem with the space complexity of $O(2^{N\times M} + 2^M)$ which is known to be an NP-hard problem [19] with a smaller search space compared to the DESAS problem.

III. BS DENSITY AND POWER ADAPTATION (BDPA) ALGORITHM

We first need to evaluate various scenarios that we may face in a dynamic network to justify the BDPA algorithm and show how can we tackle the problems defined in the previous section by adapting density and transmit power of BSs to the network state.

A. SCENARIOS

We define four different scenarios that we may face while running our algorithm among small cells. These four scenarios (S1, S2, S3, S4) and the effect of applying BDPA algorithm is depicted in the left and right side of Fig. 3 respectively. We assumed in the figures that macro BSs are connected to a MEC.

- *Scenario S1*: As it can be seen in Fig. 3a, when the traffic load in some of small cells is low and in case their assigned UEs are transferred to the associated macro cell, the utilization ratio of macro BS (that we define in Section III-C) will still remain under the threshold value. In this case, we can turn off redundant small cells and transfer their load to the macro cells.
- *Scenario S2*: Small cells can cause interference to each other that may degrade UE satisfaction as shown in Fig.3b. To enhance QoS and save energy, we may turn off interferer BSs and transfer their UEs to a neighbor BSs with a higher SINR value if the utilization ratio of the neighboring BS still remains below the expected threshold.
- *Scenario S3*: In the third scenario (Fig.3c), when a small cell's load is below the threshold and if its neighboring small cells can handle its associated UEs without causing any degradation in the coverage and the throughput, and if the macro cell is overloaded, then that small cell can be turned off and its neighbors need to expand their coverage area by amplifying their transmit power until the required received SINR can be obtained for all UEs.
- *Scenario S4*: Maintaining QoS may have a higher priority in comparison to saving energy. If the received SINR in a BS is below the minimum required received SINR value to satisfy UEs, the traffic load needs to be distributed to the neighboring cells to enhance the QoS level. As it is shown in Fig. 3d, other cells can boost their transmit power to expand their coverage area for providing service to the unsatisfied UEs located at the edges. The system may force those UEs to handover to their neighboring cells for balancing the load.

In these scenarios, the main goals are enhancing UEs QoS by lowering received interference and increasing UEs throughput while the transmit power of BSs and the density of active BSs are optimized.

B. DISTRIBUTED, CENTRALIZED OR DECENTRALIZED SOLUTION?

For solving the DESAS problem, there are basically three approaches. One is to design a centralized approach where the input size to the problem is the number of BSs. The other approach is the fully distributed approach where each BS tries to solve the problem individually. Since the DESAS is an NP-hard problem, a centralized solution is not feasible in practice, because of the huge search space. When the number of BS's the increases, the solution space grows exponentially. On the other hand, scheduling BSs for conserving energy may affect the overall network throughput if it is done individually by each BSs in a distributed fashion. Since proposing a fully centralized or a distributed solution is not feasible, we design

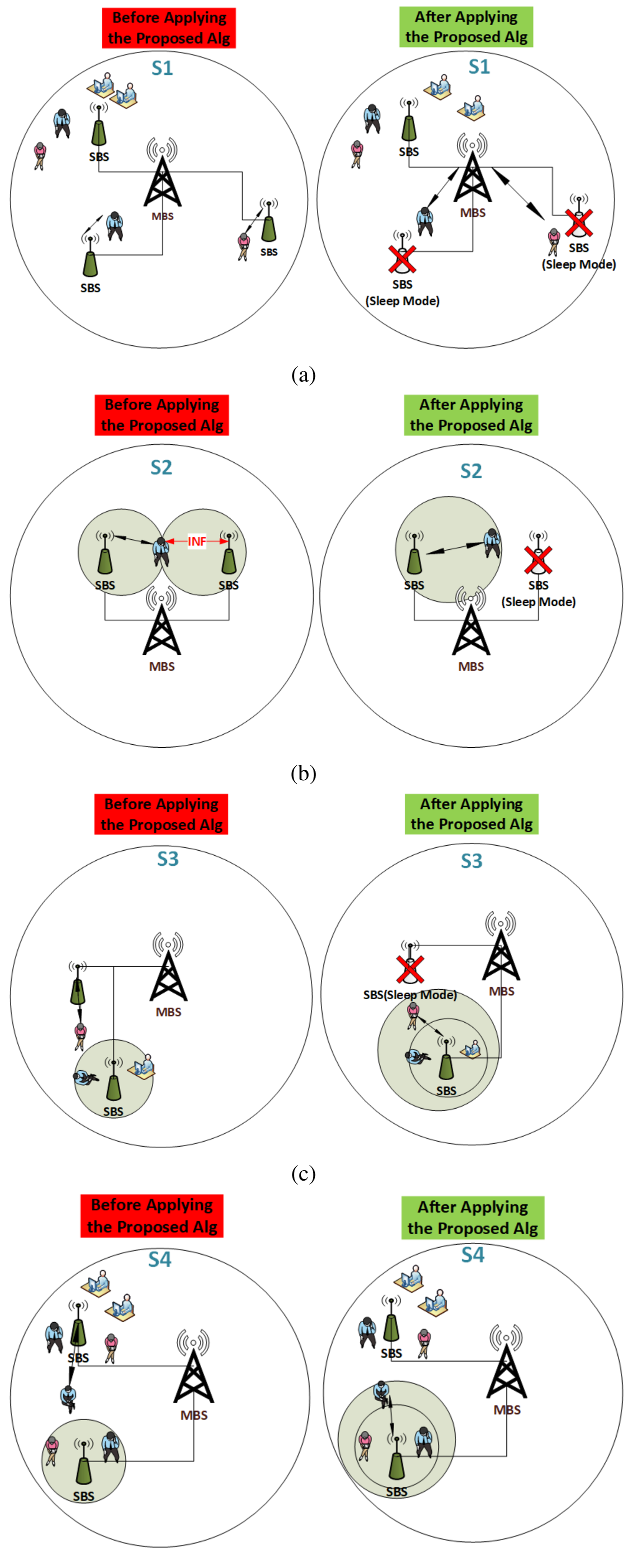

FIGURE 3. Four different scenarios (S1, S2, S3, S4) that we may face in a HetNet (on the left) and the effect of the BDPA algorithm over these scenarios (on the right). (a) Energy conservation by moving load from small to macro BSs. (b) Reducing interference (INF) by adapting BS density. (c) Cell-zooming for conserving energy.(d) Cell-zooming for user satisfaction.

a hybrid solution using multi-access edge clouds (MECs) to augment the capabilities of BSs. We off-load some management and control tasks such as coverage control and sleep scheduling from the core network to the edges of the network.

MECs, by accessing multiple BSs and communicating with each other, have a larger perspective (accessing data which is not limited to just one cell) over the network state in comparison to distributed models. The overall cost to collect network state for a MEC is lower than a centralized model since the divide-and-conquer approach is employed. Although centralized models have the largest perspective over the network state in comparison to MEC and fully distributed models, they are not scalable; the problems become intractable as the network grows. Since MECs are located to the edge of network and close to the BSs, they can collect information from cells as fast as fully distributed models while the delay for collating the information at the core network is huge. With a similar argumentation, we can claim that the amount of processing power and energy needed by centralized models are large in comparison to MECs. All in all, although centralized models can allocate resources optimally, reaching the optimal result may not always be feasible if the size of network is too large. Therefore, by employing MEC we can reach a sub-optimal solution faster with a lower amount of energy and processing power. The overall comparison among decentralized, MEC and fully centralized solutions are summarized in Table 2.

We assume a two-tiered HetNet where under the macro BS's coverage, there are a number of small cells. In the BDPA algorithm, macro BSs will always be up and operational, and the proposed scheduling model will be executed over small cells to maintain the network coverage. The BDPA algorithm consists of BDA and PA algorithms to solve the DESAS problem (8a-10a).

C. BS DENSITY ADAPTATION (BDA) ALGORITHM

Let's assume the total number of available RBs in BSj during the time interval T is represented as $\tau_j(T)$. The resource utilization ratio of BS j can be defined as the percentage of allocated resource blocks to satisfy the minimum required throughput in that BS is defined as

$$A_j = \frac{\sum_{i=1}^{N} \tau_{ij}}{\tau_{j(T)}} = \frac{\sum_{i=1}^{N} \frac{R^*_{ij}}{b_\tau log_2(1+SINR_{ij})}}{\frac{b_j}{b_\tau} T}, \tag{11}$$

where b_j is the total available bandwidth of BS j. To incorporate the resource utilization ratio, we can redefine $SINR$ in (1) as

$$SINR_{ij} = \frac{g_{ij} P_{ij}}{\sigma^2 + \sum_{k=1,k\neq j}^{M} A_j g_{ik} P_{ik}}, \tag{12}$$

We need to prevent coverage holes when BSs are turned off. By employing MEC, we can calculate the probability of network coverage for different SINR values by considering the density of BSs (λ_b). In this paper, we validate our results

TABLE 2. The comparison of implementing a fully distributed, a decentralized or a centralized solution to the DESAS problem.

	Fully Distributed	Decentralized (MEC)	Centralized
Awareness of other cells' conditions	Not known	Aware	Aware
Information collection cost	Very low	Low	Very high
Scalability	Scalable	Scalable	Not scalable
Perspective	Per BS individually	Over BSs under its supervision	For the whole network
Time resolution	Very fast	Fast	Very slow
Computational power	Low	High	Very high but not feasible
Resource allocation	Not optimal	Suboptimal	Optimal

by employing the probability of coverage, $P_C(T_{SINR})$ as proposed in [20],

$$P_C(T_{SINR}) = \frac{2}{\gamma}\int_0^\infty t^{\frac{2}{\gamma}-1} e^{-tTN_0\alpha^{-\frac{2}{\gamma}}} e^{-t^{\frac{2}{\gamma}}} \times \left(-\frac{2}{\gamma}\frac{Tt^{\frac{2}{\gamma}}F(1,1-\frac{2}{\gamma};2-\frac{2}{\gamma};-T)}{1-\frac{2}{\gamma}}\right) dt, \quad (13)$$

where

$$\alpha = \frac{\lambda_b \pi E[(P_d S_d)^{\frac{2}{\gamma}}]}{K^2}$$

is the distribution of shadowing S_d that is considered as arbitrary except $E[S^{\frac{2}{\gamma}}] < \infty$, and $F(a, b; c; z)$ is the hyper-geometric function [20]. In this model, T is the received SINR value in the network, $K = 6910\ km^{-1}$ for urban environment, which can be obtained from the COST Walfisch-Ikegami model and γ is a path-loss exponent. Here, $E[(P_d S_d)^{\frac{2}{\gamma}}]$ is the propagation invariance which S_d and P_d are independent random vectors. S_d represent the effect of propagation of a signal from its origin at S_d and P_d is the transmit power of that signal. The proposed model can be easily evaluated numerically if we consider the noise power zero ($N_0 = 0$) as follows:

$$P_C(T_{SINR}) = \left[-\frac{2}{\gamma}\frac{Tt^{\frac{2}{\gamma}}F(1,1-\frac{2}{\gamma};2-\frac{2}{\gamma};-T)}{1-\frac{2}{\gamma}}\right]^{-1}. \quad (14)$$

In this model, BSs are distributed based on Poisson model and by considering this fact that BSs are not completely turned off (they are always aware of the network condition through the MEC in order to be activated and deactivated during each time interval), the Poisson distribution assumption will still be valid throughout the life-cycle of the network. By obtaining P_C from (13), we can find the corresponding required *SINR* value to keep $P_C \geqslant T_{SINR}$. In (12), *SINR* depends on the activity ratio of cells. Therefore, we can determine the maximum activity ratio each BS can tolerate to keep the P_C above the network threshold by knowing the required minimum *SINR* value.

For adapting density of small BSs, we need to make sure user's traffic requirements are satisfied and coverage can still be preserved. For offloading the traffic when a small cell is deactivated, we need to evaluate the capacity of neighboring small cells and the macro BS to find out whether or not they can handle the offloaded traffic. We also need to analyze the effect of applying the scheduling algorithm over the network coverage to prevent holes.

The pseudocode of the proposed BDPA algorithm is presented in Algorithm 1. Due to heterogeneity of the network and variations of BS density, we need to calculate a threshold SINR value to maintain network coverage and minimum cell utilization factor (A_{min}) in each run (lines 2-4). A MEC may access data from a wider range of base stations and it has a larger computation power than base stations. Therefore, the density of BSs can be estimated by using a density estimator at the MEC [21]. By obtaining density of BSs, the minimum T_{SINR} based on the required probability of coverage (P_c) can be computed; then, it can be used to calculate the minimum cell utilization factor for the next step. We can obtain the minimum required throughput in each small cell (R^*_{ij}) by comparing the achievable throughput in the *basic model* (a network where density and power adaptation is not employed) and set it as R^*_{ij} in the BDPA algorithm. We then check the utilization factor of each BS to compare with A_{min} (line 5). If utilization factor of BS j is above the threshold and if its associated macro BS has enough capacity (A_{mj}) to handle BS j load (line 6) we can turn off that base station and transfer its traffic load to its associated macro BS. If macro BS is overloaded, we can transfer BS j's load to one of its neighbors, if possible (lines 12-19). However, if its neighbor cannot handle BS j's traffic load, we need to keep this BS on and use our power adaptation algorithm (*PA()*) that will be explained in the next section. If the utilization factor of BS j is below the threshold and by running our power adaptation algorithm the utilization factor of BS j still remains below the threshold, we can turn on redundant BSs and recalculate A_{min} until we make sure the required QoS (users' traffic requirements) can be satisfied (lines 27-39).

The computational complexity of the joint BDPA algorithm is O($(N_{macro} + N_{small}) \times \tau_{tot}$). Due to the distribution of computational tasks among MECs, each MEC is responsible only for part of BSs and the BDPA algorithm runs in polynomial time.

D. POWER ADAPTATION (PA) ALGORITHM

To maintain coverage and reduce interference in a network in addition of considering the capacity of other cells, we employ the cell zooming technique that we partially presented in [22]. In this work, we expand or reduce the cell size for enhancing the overall network throughput.

Algorithm 1 The BS Density and Power Adaptation Algorithm

for $j = 1$:*All BSs connected to MEC* **do**
  $\lambda \leftarrow$ density of BSs
  $T_{SINR} \leftarrow$ find T_{SINR} such that $P(T_{SINR}, \lambda) \geq T_{cov}$
  $A_{min} \leftarrow \frac{\sum_{i=1}^{N} \frac{R_{ij}^{*}}{b_\tau log_2(1+SINR_{ij})}}{\frac{b_\tau}{b_j} T}$
  if $A_{min} \leq A_j$ **then**
    if $A_{min} \leq A_j + A_{mj} \leq 1$ **then**
      Turn off BS j
      $\lambda \leftarrow \frac{N_{small}-1}{Area}$
      $DeactivatedBSs + 1$
    end
    else
      Find a neighbor small cell k
      with the highest received SINR
      if $A_{min} \leq A_j + A_k \leq 1$ **then**
        if $SINR_k \geq T_{SINR}$ **then**
          Turn off BS j
          $\lambda \leftarrow \frac{N_{small}-1}{Area}$
          $DeactivatedBSs + 1$
        end
        else
          #Our heuristic cell-zooming
          #algorithm needs to be run to
          #maintain QoS in the network
          PA()
        end
      end
    end
  end
  else
    PA()
    if $A_j < A_{min}$ && $DeactivatedBSs > 0$ **then**
      while $A_j \geq A_{min}$ *or* $DeactivatedBSs > 0$ **do**
        Turn on a BS
        $\lambda \leftarrow \frac{N_{small}+1}{Area}$
        $DeactivatedBSs - 1$
        $T_{SINR} \leftarrow$ recalculate T_{SINR}
        $A_{min} \leftarrow \frac{\sum_{i=1}^{N} \frac{R_{ij}^{*}}{b_\tau log_2(1+SINR_{ij})}}{\frac{b_\tau}{b_j} T}$
      end
    end
  end
end

In the first step, based on the determined SINR threshold from (13), we need to adapt transmit power of BSs by finding out an adequate threshold value (T_{SINR}) which can keep P_C above the minimum requirement of network operators (Fig.4, step a) for enhancing the network throughput while the coverage in the whole network is preserved. In the second step, the appropriate modulation and coding scheme (MCS) can be obtained based on the channel quality indicator (CQI) value that can be obtained by BSs from their associated UEs (Fig.4, step b) [23].

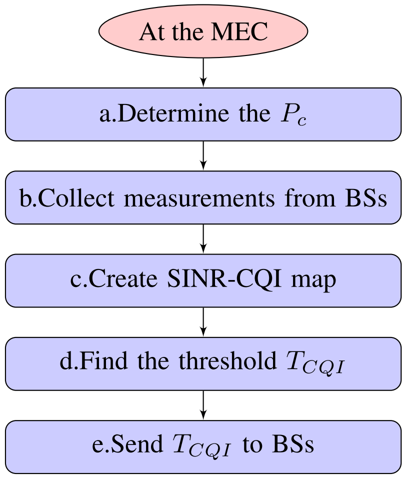

FIGURE 4. The mobile edge tasks flow chart.

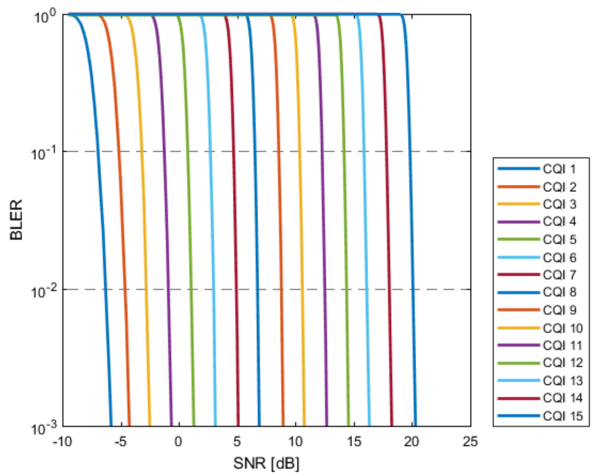

FIGURE 5. The BLER-SINR map for different CQI value.

After obtaining T_{SINR}, we need to find the corresponding CQI value based on the channel condition. In the third step (Fig.4, step c), the SINR values can be mapped to their corresponding CQI levels by calculating the block error rate (BLER) for each CQI level as it is shown in Fig. 5. In Fig. 5, we calculate the BLER for each CQI value to find the corresponding SINR value when the BLER reach to 10%. The corresponding SINR-CQI map is shown in Table 3.

In the last step of the flowchart shown in Fig. 4, at step d, the MEC needs to determine an appropriate CQI value to satisfy network coverage and enhance the throughput simultaneously (T_{CQI}). There is a trade-off between choosing high CQI value and maintaining the network coverage. To maintain the network coverage, we need to use a lower MCS to make sure the received signal is interpretable by the receivers even in

TABLE 3. The SNR-CQI map table.

CQI	SNR	CQI	SNR	CQI	SNR
1	-6.93	6	2.69	11	12.28
2	-5.14	7	4.69	12	14.17
3	-3.18	8	6.52	13	15.88
4	-1.25	9	8.57	14	17.81
5	0.76	10	10.36	15	19.82

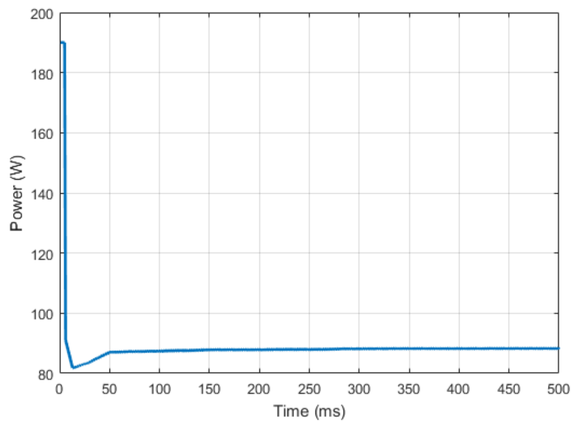

FIGURE 6. The effect of applying fluctuation reduction over the PA algorithm.

low quality channel conditions (low received SINR value). By decreasing the MCS, the size of transport block and the number of bits per symbol will be decreased which can drop the network throughput significantly. Therefore, MEC by considering the probability of coverage and SINR-CQI map which are obtained from previous steps, can choose an appropriate threshold value (T_{CQI}) and inform BSs to adapt their modulation and coding schemes and their power consumption based on it (Fig.4, step e).

This procedure has to be repeated in case of any changes in the network topology such as increasing the amount of maximum transmit power, type of base stations, density of BSs or environmental conditions. In Fig. 4 the summary of mobile edge tasks is presented.

When all base stations are informed about the threshold value which is obtained in the previous step by the MEC; each base station employs the power allocation algorithm presented in Algorithm 2 that is based on the received CQI feedback from it's UEs. BSs compare CQI measurements of user-specific reference signals [24] with the threshold value (line 5) to increase or decrease the power allocation to each RB individually per transmit time interval (1 ms).

In the basic power allocation model where power is allocated to each UE homogeneously, UEs located at the edges experience a high amount of interference from neighboring base stations. The amount of power allocated to UEs located at the cell-center has to be reduced to overcome the interference effect over UEs located at the edges in the PA algorithm.

Each base station independently monitors the feedback from active UEs. When the amount of received CQI value is higher than the threshold, the BS will reduce the power (P) continuously until it makes sure the allocated power is higher than the minimum possible transmission power (P_{min}) and the new received CQI value is still equal or higher than the threshold (lines 15-18). To enhance the network throughput, BSs need to allocate more power to the UEs located at the edges in comparison with other UEs. Therefore, if the amount of received feedback from a UE is less than the threshold, BSs will consider it as the edge UE. In this case, BSs need to amplify the allocated power for 1% (Δ) until they make sure the new CQI feedback is higher than the threshold and the amount of allocated power is still below the maximum transmission power (P_{max}) (lines 6-9). By choosing a larger (Δ) value, the precision of the PA algorithm can reduce, but the algorithm will converge faster. BSs periodically communicate with MECs to adapt their transmission powers by any changes happen in the network condition. To decrease the convergence time in our algorithm because mainly large portion of UEs are located close to the BSs; at first, we reduce the power to the half then we run the algorithm to decrease the power adaptively based on UEs conditions (line 13). To reduce the fluctuation during power allocation, BSs maintain the history of last 10 power allocation to their associated UEs. In case, the amount of power allocated to a UE after power reduction for about (Δ) is not lower than the average of power which have been allocated to the UE in the last 10 time slots (P_{avg}) (line 12), BSs will allocate the same power as the previous step. In Fig. 6, we present the effect of applying fluctuation reduction (FR) over the proposed algorithm. As we can see, the overall power consumption in the network will converge after 50 *ms* and by applying FR over our algorithm; the power fluctuation is reduced.

IV. RESULTS AND DISCUSSION

For the evaluation, we use the Vienna-LTE simulator which is a system level simulator to implement the downlink channel model and the network environment of a multi-user OFDMA system such as LTE [25]. In this simulator, collisions on the random-access channel and other parameters such as noise, interference, shadowing, fading, antenna size, BSs height, number of transceivers, angle of antennas, handover, channel model, traffic model, walking model are considered based on real-life LTE networks and applied in the proposed algorithm. We modified the power allocation model and developed our scheduling module in this simulator.

A. SIMULATION MODELS AND PARAMETERS

The scheduling algorithm is implemented for sparse and dense small cell scenarios to show its capabilities in dynamic networks. We simulate two different network scenarios:

- Sparse network: where the density of BSs varies between 5 to 20 BS/km^2, and 100 UEs are uniform randomly distributed in the region of interest.
- Dense network: the density of BSs varies between 20 to 100 BS/km^2, and 1000 UEs are uniform randomly distributed in the region of interest.

Algorithm 2 PA Algorithm

Receive the threshold value from MEC
for *All available RBs* **do**
  $P_{max} \leftarrow P_{max} = \frac{\text{Sector Maximum Downlink Power}}{\text{Number of Available RBs}}$
  $P_{min} \leftarrow$ minimum applicable power per RB
  if $CQI \leq T_{CQI}$ **then**
    while $P \leq P_{max}$ *and* $CQI < Threshold$ **do**
      $P \leftarrow \text{P} + \Delta$
      $CQI \leftarrow$ Request for a new CQI feedback
    end
  end
  if $CQI \geq T_{CQI}$ **then**
    if $P - \Delta \leq P_{avg}$ **then**
      $P \leftarrow \frac{P}{2}$
      $CQI \leftarrow$ Request for a new CQI feedback
      while $P \leq P_{min}$ *and* $CQI < Threshold$ **do**
        $P \leftarrow P - \Delta$
        $CQI \leftarrow$
        Request for a new CQI feedback
      end
    end
  end
end

To provide a fair comparison among the BDPA algorithm, the LP-WUR [7] and the C-ICIC [8] are described in sections IV.D.1 and IV.D.2, respectively, we employed different types of bursty traffic loads in our simulation which are modeled based on real-life LTE networks [26]. Simulation parameters and traffic models are summarized in Table 4 and Table 5, respectively.

B. VIENNA SIMULATOR, IMPLEMENTATION DETAILS

Vienna LTE system level simulator consists of different modules including antennas, channel models, network generation, schedulers, traffic models, walking models and etc. Each BS includes three sectors and in each sector, small cells are implemented. UEs and BSs are distributed uniform randomly in the region of interest. A wide variety of schedulers exist in this simulator such as best CQI, proportional fair, round robin, and max-min. All of these schedulers will be invoked through *lteScheduler.m* file. To make the BDPA algorithm applicable over different scheduling models, we implement our algorithm in the scheduler coordinator (*lteScheduler.m*) under the power allocation module. We also modify the traffic and the mobility models for different scenarios as explained in the sequel.

C. WHAT'S HAPPENING AT MEC?

Our main aim is reducing the amount of energy consumption in the network by gathering information at the MEC to apply the BDPA algorithm. In MEC, by considering the density of BSs, the appropriate T_{SINR} and T_{CQI} values which can be used as the threshold value to maintain minimum throughput and coverage requirements in the network, needs to be obtained through (13).

TABLE 4. Simulation parameters and their values.

Parameters	Value	Ref
Frequency	2.14 GHz	[27]
K	6910 KM^{-1}	[20]
γ	4	[20]
N_0	$10^{-15.82}$	[20]
Subcarrier Frequency	15 kHz	[27]
Macro BSs Max Power	10 W	[27]
Small BSs Max Power	100 mW	[7]
Small BSs Sleep Power	15 mW	[7]
LP-WUR Power	10 μW	[7]
TTI	1 ms	
Simulation Area	2000 m $\times$ 2000 m	
RRH Antenna Gain	Omni-directional	[28]
Path Loss Model	$128.1 + 37.6log_{10}^{R}$, R in km	[27]
Noise Power Spectral Density	-174 dBm/Hz	[27]
Receiver Noise Figure	9 dB	[27]
Density of BSs (BS/km^2)	5,10,15,20,35,50,75,100	
Active UEs	20,50,100,1000	
UE Speed	5,20,40 km/h	
Feedback	CQI	[28]
Feedback Delay	3 TTI	
Scheduler	Proportional Fair	
Number of Monte-Carlo Runs	30	
Simulation Length	100 TTI	

TABLE 5. Traffic types in OFDMA-based networks [26].

Traffic Type	Transmission Category	Traffic Load
FTP	Best effort	10%
HTTP	Interactive	20%
Video Stream	Streaming	20%
VoIP	Real-time	30%
Gaming	Interactive & real-0time	20%

In Fig.7, the probability of coverage for various SINR thresholds is presented. We also examine three different BS densities (10, 50, 100 BS/km^2) to show how T_{SINR} can be changed under various conditions. As we can see, when the BS density is low (10 BS/km^2), if we set T_{SINR} at -10 *dB* the probability of coverage will be around 57%. However, by increasing the density of BSs to 50 BS/km^2 and 100 BS/km^2, the probability of coverage will be boosted up to 98% and 100%, respectively. By evaluating Fig.7 we can find out if BS density is very low, we may face with too much coverage holes which can degrade the QoS in the network. Moreover, by increasing the density of BSs the amount of received interference will also increase, but this improvement will be negligible when the BS density reaches to its optimum level. Therefore, we need to adapt density of BSs to the network condition to enhance QoS and prevent resource wastage in the network.

The SNIR threshold (T_{SINR}), which is found at the edge cloud will be used by the BDA algorithm and by employing the snapshot of the network condition, it has to be mapped to a CQI value since BSs characterize the channels' quality by using 4-bit CQI values. Therefore, by evaluating BLER in the network, the SINR value which is obtained from the

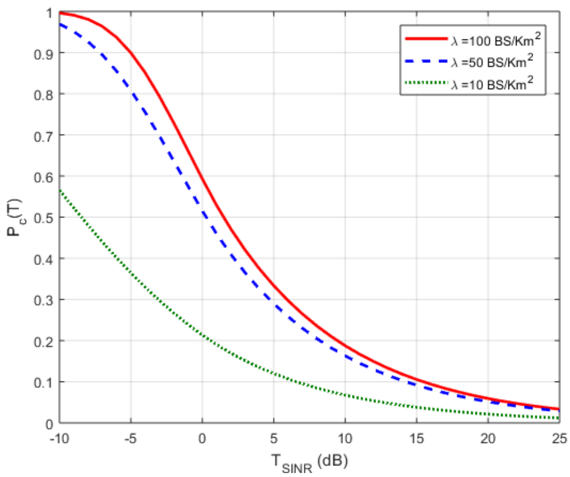

FIGURE 7. **The probability of coverage for different SINR thresholds.**

previous step can be mapped to its corresponding CQI level (Fig.5) to achieve (T_{CQI}) which will be forwarded to BSs to be used in the PA algorithm. Therefore, the BDPA algorithm will be invoked to optimize the network resources when the corresponding information is received through MECs.

D. VALIDATION OF THE BDPA ALGORITHM

We validate the BDPA algorithm in this section and compare it with the a sleep scheduling technique that we call as the LP-WUR technique which employs low-power wake-up receivers [7]. We also compare our results with a cooperative interference inter-cell control (C-ICIC) model which jointly allocates power and resource among UEs to enhance energy consumption and network throughput [8]. In the LP-WUR, density of BSs are adapted based on UEs request, while in the C-ICIC energy consumption will be optimized by collaboration of neighbor BSs with each other. Therefore, we can compare the effect optimizing energy consumption based on UEs request or collaboration of neighbor cells with a MEC based architecture (BDPA) in this paper. We present its details in the sequel. We evaluate throughput, spectral efficiency, energy efficiency, and the impact of mobility and user density on the performance of the BDPA algorithm in this section.

1) LP-WUR SLEEP SCHEDULING TECHNIQUE

By equipping senders and receivers with wake-up receivers, an on-demand energy-efficient UE-controlled sleep scheduler is presented in [7] that we call as LP-WUR technique. For comparing the BDPA performance we implemented the LP-WUR technique in the Vienna simulator as well. Due to a low amount of power consumption in wake-up radio modules (10 μW approximately), they can be always in the active mode to listen to channels. Therefore, the wake-up radio can trigger a BSs for transmission whenever a wake-up signal is successfully received and small cells can go to the sleep mode when the transmission is terminated. To implement LP-WUR, we trace communications among UEs and BSs and we examine energy consumption, throughput and spectral efficiency for this algorithm, whenever a packet is exchanged between UEs and BSs.

2) COOPERATIVE ICIC (C-ICIC) TECHNIQUE

In this model, authors introduced C-ICIC technique which can enhance the received SINR and exploits communication channel by reducing the inter-cell interference (ICI) effect between adjacent BSs [8]. In the C-ICIC, UEs are classified into a bad radio (BR) group if they are negatively affected due to the ICI (by considering the amount of received wideband CQI feedback), while the rest of UEs will be classified as a good radio (GR). In the C-ICIC, satisfaction functions for UES, cells, sectors are defined. Then, resource blocks and power allocation are based on the satisfaction functions. C- ICIC is a distributed technique facilitating cooperation among adjacent cells.

E. MONTE-CARLO SIMULATION RESULTS

In this section, we present the quality of service, spectral efficiency, energy consumption results. We further evaluate the impact of user density and mobility.

1) QUALITY OF SERVICE (QOS)

As it is explained previously, QoS can be maintained when the minimum required throughput in each small cell (R^*_{ij}) can be satisfied. In this work, for increasing the system efficiency with respect to our opponents (LP-WUR and C-ICIC), we defined R^*_{ij} as the mean throughput which can be achieved in in LP-WUR and C-ICIC. In Fig.8 and Fig.9, we compare the mean throughput of the BDPA algorithm with the LP-WUR and the C-ICIC techniques in sparse and dense network scenarios, respectively. Throughput can be affected by increasing density of UEs and BSs in the network, due to higher transmission rates in the network. As we can see, when BS density is low (Fig.8) the maximum achievable throughput by the BDPA, the LP-WUR and the C-ICIC algorithms are around 10 *Mbps*, 8.5 *Mbps* and 9 *Mbps*, respectively. These values are boosted up to 15 *Mbps* and 11 *Mbps* when BS density is increased (Fig.9). In both scenarios, we can achieve a higher network throughput in different conditions by employing the BDPA algorithm. Because, in the BPDA algorithm by considering the capacity of each cell and reducing the number of active BSs, the total amount of interference received by UEs will be reduced which increases channel quality and improves the overall network throughput. Moreover, unlike the C-ICIC, the BDPA by employing MEC in its architecture can have a higher perspective over the network which can maintain network coverage and adapt BSs density while transmission power is reduced. With the power adaptation feature of the BDPA algorithm, the received SINR for UEs located at the cell borders will be enhanced that improves the network throughput compared to the LP-WUR and the C-ICIC technique. The proposed algorithm can provide a larger network throughput in dense networks and can enhance QoS up to 26% in comparison to the LP-WUR and the C-ICIC techniques.

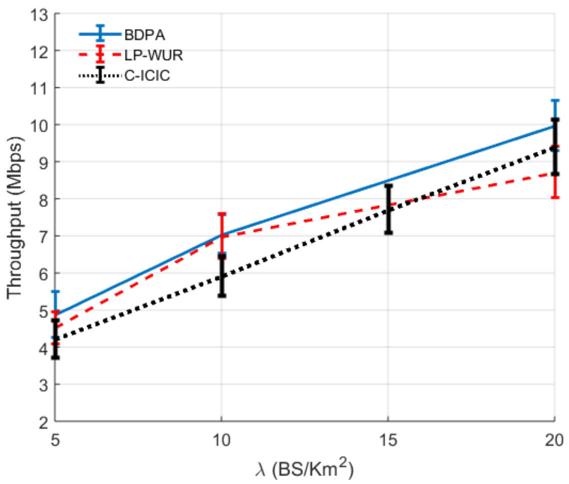

FIGURE 8. The effect of applying the BDPA algorithm on network throughput in a sparse network with 100 UEs for different BS densities.

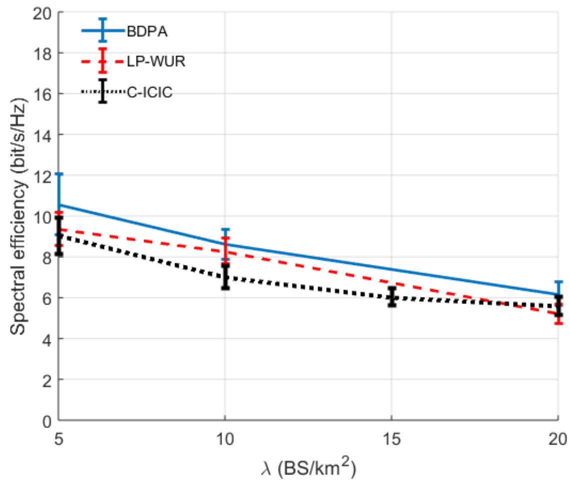

FIGURE 10. The effect of applying the BDPA algorithm in a sparse network with 100 UEs on spectral efficiency for different BSs' densities.

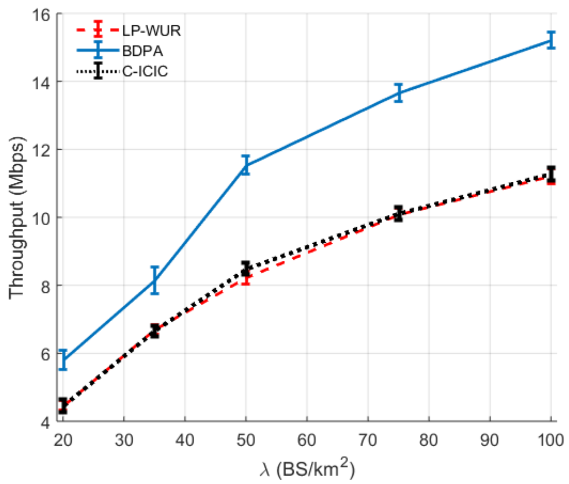

FIGURE 9. The effect of applying the BDPA algorithm on network throughput in a dense network with 1000 UEs for different BS densities.

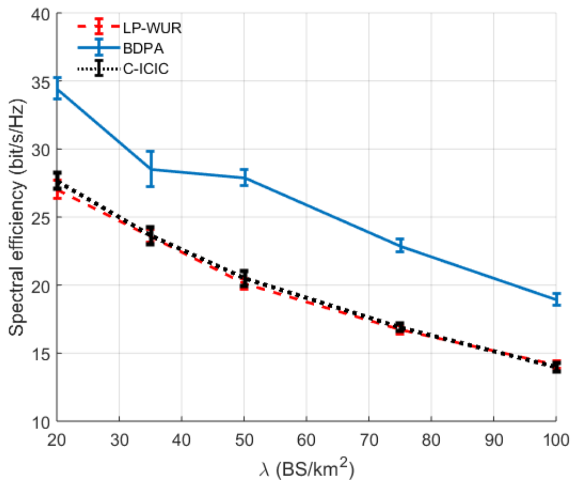

FIGURE 11. The effect of applying the BDPA algorithm in a dense network with 1000 UEs on spectral efficiency for different BSs' densities.

2) SPECTRAL EFFICIENCY

We present the effect of applying the BDPA algorithm on the spectral efficiency in Fig.10 and Fig.11. Spectral efficiency can be enhanced by using a high order MCS which is possible when the interference in the network is low and the network has enough capacity to handle the traffic load. In both cases (sparse and dense) the spectral efficiency will reduce by increasing BS density due to the high amount of interference. The highest spectral efficiency in a sparse network for the BDPA, the LP-WUR and the C-ICIC are when BS density is 5 BS/km^2 (Fig.10) and it is about 11 *bps/Hz*, 9 *bps/Hz* and 8.9 *bps/Hz*, respectively. Additionally, in the dense network (1000 UEs) the highest spectral efficiency for the BDPA, the LP-WUR and the C-ICIC can be achieved is when BS density is 20 BS/km^2 and it is about to 35 *bps/Hz*, 27 *bps/Hz* and 28 *bps/Hz*, respectively. Due to the larger number of deactivated BSs by the BDPA algorithm (which can reduce the interference in the network), the average spectral efficiency per cell can be enhanced by 30%. The BDPA algorithm, by adapting density of BSs and distributing their load to other cells, can use resources more efficiently in comparison to the LP-WUR and the C-ICIC techniques. Although BSs are only activated whenever UEs request in the LP-WUR technique, they can be activated with even a single UE request. Therefore, the LP-WUR technique cannot use resources efficiently and spectral efficiency will always be lower than that of the BDPA algorithm. The C-ICIC adjusts transmit power in two levels based on UEs class (BR and GR) while in the BDPA, power can be adjusted at different levels based UEs condition which can distribute power among UEs more efficiently. Therefore, the BDPA can provide higher throughput and spectral efficiency with respect to the C-ICIC.

3) ENERGY CONSUMPTION

To evaluate the energy efficiency of the BDPA algorithm, we analyze the amount of energy consumption for different

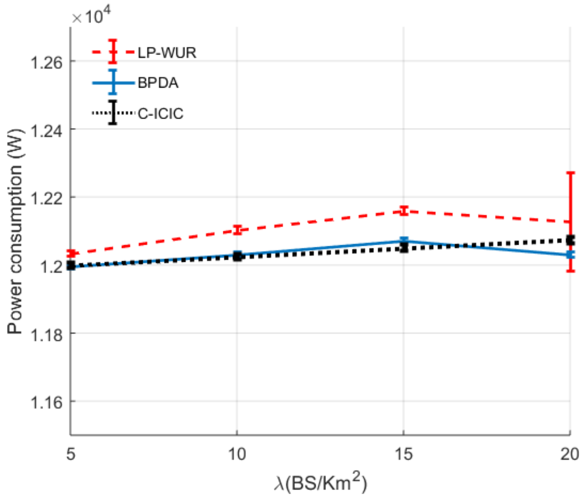

FIGURE 12. The effect of applying the BDPA algorithm on energy consumption in a sparse network with 100 UEs and bursty traffic for different BSs' densities.

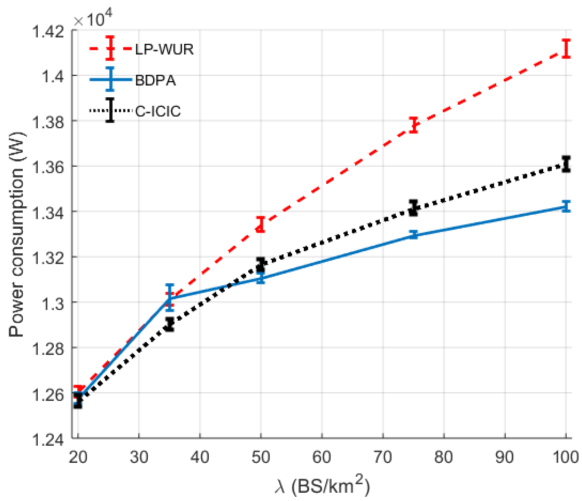

FIGURE 13. The effect of applying the BDPA algorithm on energy consumption in a dense network with 1000 UEs and bursty traffic for different BSs' densities.

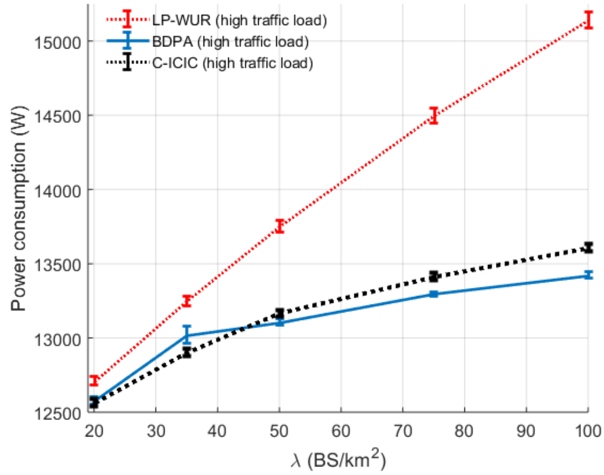

FIGURE 14. The effect of applying the BDPA algorithm on energy consumption in a dense network when traffic load is high for different BS density.

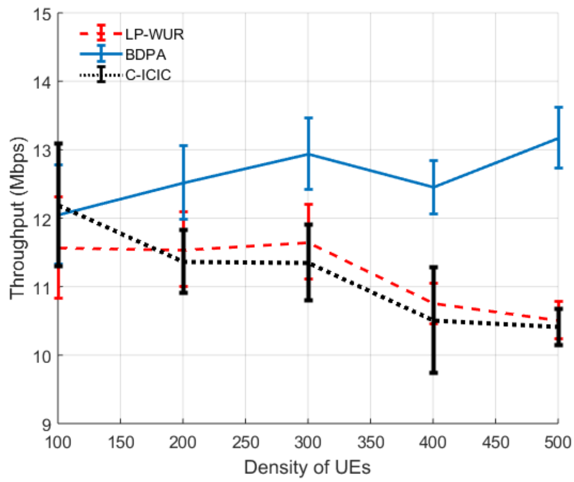

FIGURE 15. The effect of applying the BDPA algorithm on network throughput when density of BSs is fixed at 50 BS/km^2 and UEs' density is varied between 100 to 500.

numbers of UEs and BSs in Fig.12 and Fig.13. As it is shown, by increasing BS density the amount of energy consumption in both scenarios will increase. When the network is sparse (Fig.12) and BS density is 20 BS/km^2, the BDPA, LP-WUR and C-ICIC algorithms consume almost same amount of power for about 12000 W; while, the overall power consumption in the dense scenario (Fig.13) will increase up to 13400 W, 14100 W and 13600 W, respectively when BS density is 100 BS/km^2. In all cases, the highest amount of energy is conserved in the BDPA algorithm due to the low number of active BSs during each transmission in comparison with [7] and [8]. As we can see in these figures, by adapting MCS in the network based on our power adaptation model and offloading BSs, we can always achieve less energy consumption in comparison to the LP-WUR and the C-ICIC techniques. The BDPA algorithm can conserve a higher amount of energy in dense networks and can save up to 5% more energy with respect to the LP-WUR technique which means 3 kW per second or 98.55 *Tera-Watt* per year. The amount of energy conservation will be larger in ultra-dense networks.

Moreover, in the LP-WUR technique, when we have a continuous traffic load all BSs need to be in the active mode to provide service for their associated UEs. However, in the BDPA algorithm, network resources will be used in an efficient manner by offloading the traffic of low-load BSs to neighboring cells and deactivating them to conserve more energy. As we can see in Fig.14, the proposed algorithm can conserve up to 12% more energy when the traffic load is high.

4) USER DENSITY

In dynamic networks, in addition of variation in density of BSs, UE density will also vary in a day. By keeping the density of BSs constant (50 BS/km^2), we evaluate the effect of UE density on throughput, spectral efficiency and power consumption in Fig.15, Fig.16 and Fig.17,

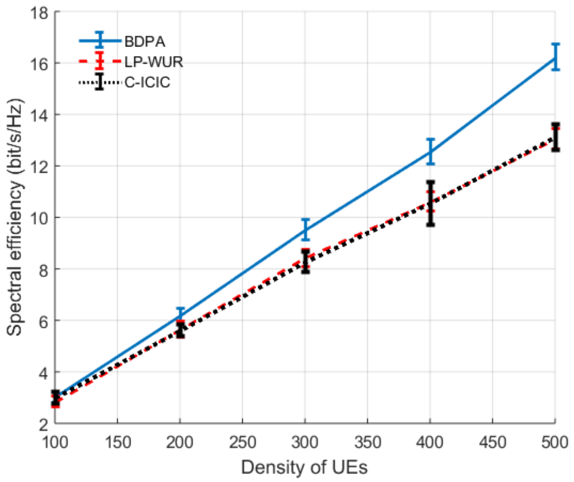

FIGURE 16. The effect of applying the BDPA algorithm on spectral efficiency when density of BSs is fixed at 50 BS/km^2 and UEs' density is varied between 100 to 500.

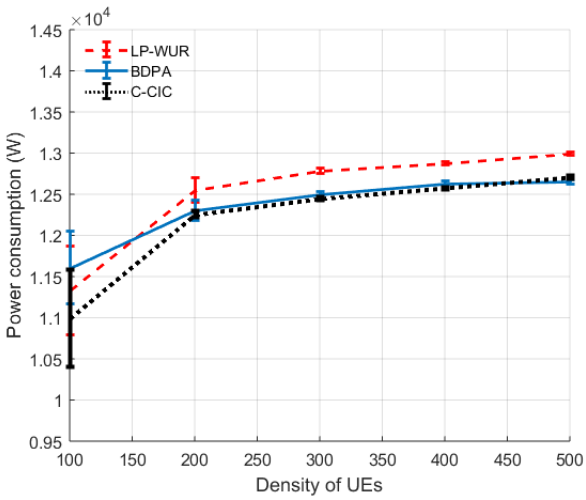

FIGURE 17. The effect of applying the BDPA algorithm on power consumption when density of BSs is fixed at 50 BS/km^2 and UEs' density is varied between 100 to 500.

respectively. As we can see, in all of these cases, the BDPA algorithm can achieve higher performance in comparison to the LP-WUR and the C-ICIC techniques. When the density of UEs are low, they have almost the same performance. As we can see, by increasing UE population, the BDPA algorithm produces better results. In the BDPA algorithm, to satisfy UE constraints such as throughput, when UE density is low and based on the UEs locations we may need to keep more BSs in the active mode in comparison to the LP-WUR and the C-ICIC technique. However, by increasing UEs' density the amount of data transmission will also increase, and LP-WUR needs to turn on and off BSs more frequently. At this point, the BDPA can be more efficient by distributing load among cells and deactivating redundant BSs. Moreover, in BDPA by enhancing *SINR* (12), we can enhance the received interference by UEs more accurately with respect

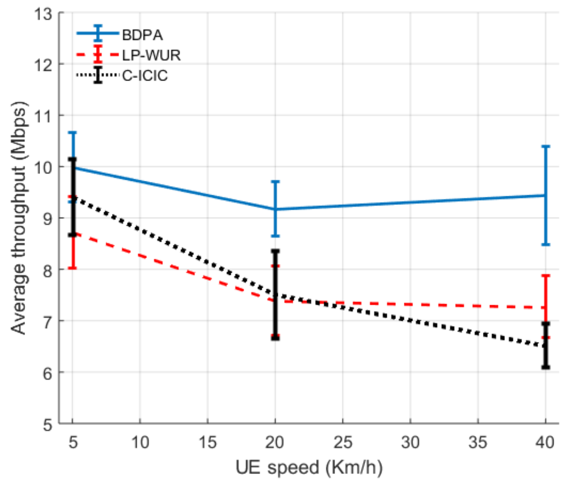

FIGURE 18. The effect of applying proposed algorithm on network throughput in a sparse network with 100 UEs for varied velocities.

to the C-ICIC. Therefore, the overall throughput and spectral efficiency will be enhanced while the energy consumption can be maintained in comparison with the C-ICIC.

5) MOBILITY

To provide a real-life condition in our simulation, the mobility of UEs for different speeds are also considered. The density of UEs and density of BSs are 100 and 20 BS/km^2, respectively and we evaluate 3 different speeds for UEs (5 *km/h*, 20 *km/h* and 40 *km/h*) to simulate pedestrians, cyclists and drivers in our analysis. As it is shown in Fig.18, the proposed algorithm can achieve a higher network throughput in all cases compared to the LP-WUR and the C-ICIC techniques. The main reason that BDPA can perform better is equipping BDPA with MEC and providing higher perspective over network state in comparison with the LP-WUR and the C-ICIC techniques. Therefore, the BDPA algorithm can predict faster and more accurate when a BS needs to be turned off or vice versa.

V. RELATED WORK

In a mobile network, a portion of energy is consumed dynamically by the BSs depending on the amount of the traffic load while another portion of energy is statically consumed in some components such as cooling systems, power supplies or for signal processing. In this paper, the main focus is on dynamic power consumption in BSs. Power consumption in BSs can be managed by utilizing their resources and adapting density of BSs and transmit power of BSs to the network condition.

Green networking and energy-efficiency have motivated researchers, and their works can be categorized into three main classes: (1) Efficient resource allocation models, (2) Load balancing models, and (3) Bandwidth enhancement models. We will briefly survey these categories in relation to our work in the sequel.

To reduce the energy consumption in BSs by turning off BSs' transceivers in idle times, discontinuous transmission (DTX) models can be employed in the network architecture [29], [30]. In DTX based models Multicast Broadcasting Single Frequency Network (MBSFN) sub-frames are allocated by considering traffic load in the network [31]. Although applying DTX schemes can reduce the energy consumption significantly, DTX does not useful during peak load due to the lack of empty sub-frames and it causes a delay for packet delivery [5]. Moreover, by equipping UEs and BSs with LP-WUR, delay for activation and deactivation of BSs can be reduced and BSs can be switched off during their idle periods [7]. Another method for enhancing energy efficiency is the aggregation of RBs [32]. By employing carrier aggregation schemes, the amount of overheads during communication will be reduced, which can increase energy conservation in the network. Another scheme for reducing power consumption in the network is adapting BSs' transmit power to satisfy QoS constraints by minimizing block error rate (BLER) in the network [33]. Also, energy efficiency can be increased by optimizing resource allocation in OFDMA-based networks [34], [35]. In the BDPA algorithm, in addition of adapting BSs transmit power, we adapt density of BSs to network parameters such as throughput and coverage to optimize the network resources more efficiently. Moreover, resource allocation is optimized by increasing cells' utilization factor and deactivating redundant BSs at the same time. Therefore, by adapting density and transmit power of BSs and utilizing the network resources, we simultaneously enhanced the resource management, throughput and coverage in the network from different perspectives. In [36], a joint user scheduling and power control mechanism is proposed which can enhance energy efficiency in ultra-dense networks. Although, in [36], transmit power of base stations and UE allocation are optimized, the density of BSs considered as a pre-configured value which may not applicable in dynamic networks where the density of small cells dynamically changes in time and space. Moreover, by implementing the proposed algorithm in a system level LTE network, we considered the real-life network parameters including traffic models, mobility models and etc. which are not analyzed in the mentioned work. In [37], an energy efficient on-off switching model is proposed where BSs are turned off or on by considering the amount of traffic load in the network. Unlike [37], the main focus is on the effect of increasing density of small cells in future heterogeneous networks in this paper. Moreover, the proposed algorithm in addition of considering the traffic load and the density of UE dynamicity, it also considers dynamic BSs where the density of BSs may also fluctuate. In the BDPA, network parameters (transmit power) will be adapted to the density of active BSs (PA algorithm) and the density of BSs will be also adapted to the network condition while the QoS (network throughput) can be maintained (BA algorithm).

Moreover, energy can be also conserved by applying load balancing schemes such as distance-aware models in the network architecture [38], [39]. By employing distance-aware models, when two BSs are competing with each other for registering UEs in their cells, a BS can be switched off when it has a larger distance to UEs and lower traffic load with respect to other BSs. Energy efficiency can be also improved by applying traffic-aware models to adapt energy consumption by considering traffic variations with respect to time in the network [40]. Energy efficiency can be also achieved by applying UE migration techniques to reduce the number of active BSs with the low traffic load which can be implemented in a distributed or centralized manner [41]. In centralized models, by analyzing traffic load among BSs, a BS with the highest load will be determined. After that, in case the selected BS has enough bandwidth for satisfying its neighbor UEs, UEs from a BS with light traffic load will be migrated and the selected BS will be turned off. In contrast, in distributed models, a BS needs to find its pair in a way that the selected BS has a lower traffic load and its traffic load can be handled by its peer. Therefore, by applying UE migration techniques the amount of energy conservation will be increased significantly in the network [42], [43]. In this work, we employed a combination of these methods with a heuristic and fast solution to enhance the network performance. In our scheduling model, we equipped our algorithm with a UE mitigation technique to provide load balancing in the network; but, unlike [38] and [39], UEs are not transferred to other cells just by considering the distance to their neighbor BSs. We consider cell utilization factor and received SINR value to find out the best candidate among BSs to mitigate UEs. We also employ a cell-zooming technique to maintain throughput and prevent any coverage holes in the network. Moreover, the BDPA algorithm, unlike [40], can enhance the network performance and energy conservation even when the traffic load is high as it is explained in Fig.14 due to the resource utilization technique which is employed in this algorithm. Last but not least, our scheduling model is faster than [41] due to its lower complexity and employing MEC instead of a centralized solution.

Bandwidth expansion models can be also employed for enhancing energy efficiency in the network [44], [45]. For instance, in time compression mode (Tcom) algorithm, energy is conserved by reducing overhead caused by control signals during transmission. Moreover, in Tcom by applying high order modulation and coding schemes, RBs are compressed in either of frequency and time domain which can expand the bandwidth in networks. Therefore, BSs can handle their load faster and signal overhead will be reduced, which can cause energy conservation in the network [46]. In [44] and [45] expanding bandwidth for UEs located at the edge of cells may not possible due to the high amount of interference received by them from their neighbor cells. In this work, by employing a cell-zooming technique we enhance communication for UEs located at the edges by minimizing interference from other cells and providing high order MCS which can expand bandwidth in each cell individually by considering the network condition.

Moreover, networks can be also optimized by employing MEC platforms in their architectures. Employing MEC platforms provide high flexibility and supporting multi-tenancy in the cellular networks, which can deliver a wide variety of services including cloud computing, network slicing, network function virtualization (NFV) and software-defined networks (SDN) [47]. Due to the heterogeneity of future networks, coordination of dynamic network is difficult which in recent works, authors tackle this issue by employing MEC platforms in networks [48], [49]. MEC can also enable an agile and simple solution to maintain connectivity and enhance service management in the future networks [50]. In recent researches, energy efficiency through MEC platforms is achieved by offloading processing loads from UEs to MECs which can alleviate UEs resource constraints [51]. MEC by providing higher computational power can also be used for enhancing routing data among UEs which can reduce the energy consumption in the network [52]–[54]. Moreover, MECs by having higher perspective over the network condition can enhance QoS and interference which can enhance power conservation in the networks [55]. The main advantages of the BDPA algorithm are employing MEC in a distributed manner, implementing it in a system level simulator by considering real-life network parameters (antenna type, height and angle, channel model, shadowing, fading, scheduling, traffic load and etc.) and enhancing the energy efficiency and the throughput by adapting network parameters including transmit power and density of BSs to each other simultaneously through the MEC platform.

VI. CONCLUSION

In this paper, we proposed a density-aware, energy-efficient and spectrum-efficient sleep scheduling technique by applying two heuristic algorithms to conserve energy and enhance the quality of service including users' throughput and satisfying the required network coverage simultaneously in a heterogeneous network. Base station density is adapted by considering cell capacities and by balancing traffic load among cells. We prevent coverage holes and enhance network throughput by applying a cell-zooming technique and improving signal-to-interference-plus-noise ratio for users located at cells borders. We provide higher processing power and higher perspective over the network state in comparison to individual base stations by equipping the network with multi-access edge cloud. The proposed model is examined with a system level simulator to provide reliable results. All in all, we find out future networks can become smarter and more efficient if we consider BS density in our models and adapt it to the network condition. This work can be extended by considering the mobility of base stations in future work.

ACKNOWLEDGMENT

This paper was presented in part at the 26th IEEE Signal Processing and Communications Applications Conference (SIU) Conference.

REFERENCES

[1] E. Björnson, L. Sanguinetti, and M. Kountouris, "Deploying dense networks for maximal energy efficiency: Small cells meet massive MIMO," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 832–847, Apr. 2016.

[2] V. Sciancalepore, V. Mancuso, A. Banchs, S. Zaks, and A. Capone, "Enhanced content update dissemination through D2D in 5G cellular networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7517–7530, Nov. 2016.

[3] A. Abrol and R. K. Jha, "Power optimization in 5G networks: A step towards GrEEn communication," *IEEE Access*, vol. 4, pp. 1355–1374, Apr. 2016.

[4] M. Ding, D. Lopez-Perez, H. Claussen, and M. A. Kaafar, "On the fundamental characteristics of ultra-dense small cell networks," *IEEE Netw.*, vol. 32, no. 3, pp. 92–100, May/Jun. 2018.

[5] K. Kanwal, G. A. Safdar, M. Ur-Rehman, and X. Yang, "Energy management in LTE networks," *IEEE Access*, vol. 5, pp. 4264–4284, Mar. 2017.

[6] C. Liu, B. Natarajan, and H. Xia, "Small cell base station sleep strategies for energy efficiency," *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1652–1661, Mar. 2016.

[7] H. Xu, W. Yu, A. Hematian, D. Griffith, and N. Golmie, "Performance evaluation of energy efficiency with sleep mode in ultra dense networks," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Mar. 2018, pp. 747–751.

[8] M. Yassin, S. Lahoud, M. Ibrahim, K. Khawam, D. Mezher, and B. Cousin, "Cooperative resource management and power allocation for multiuser OFDMA networks," *IET Commun.*, vol. 11, pp. 2552–2559, Nov. 2017. [Online]. Available: https://digital-library.theiet.org/content/journals/10.1049/iet-com.201%7.0201

[9] H. U. Sokun, E. Bedeer, R. H. Gohary, and H. Yanikomeroglu, "Optimization of discrete power and resource block allocation for achieving maximum energy efficiency in OFDMA networks," *IEEE Access*, vol. 5, pp. 8648–8658, May 2017.

[10] P. Yu, L. Feng, W. Li, and X. Qiu, "Energy-saving management mechanism based on hybrid energy supplies for LTE heterogeneous networks," *Mobile Inf. Syst.*, vol. 2016, Jun. 2016, Art. no. 3121538.

[11] M. H. Alsharif, J. Kim, and J. H. Kim, "Green and sustainable cellular base stations: An overview and future research directions," *Energies*, vol. 10, no. 5, p. 587, Apr. 2017. [Online]. Available: http://www.mdpi.com/1996-1073/10/5/587

[12] N. Yu, Y. Miao, L. Mu, H. Du, H. Huang, and X. Jia, "Minimizing energy cost by dynamic switching ON/OFF base stations in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7457–7469, Nov. 2016.

[13] C. Luo and J. Liu, "Load based dynamic small cell on/off strategy in ultra-dense networks," in *Proc. 10th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2018, pp. 1–6.

[14] O. Grodzevich and O. Romanko, "Normalization and other topics in multi-objective optimization," in *Proc. Fields-MITACS Ind. Problems Workshop*, Aug. 2006, pp. 1–13.

[15] A. Mesodiakaki, F. Adelantado, L. Alonso, M. Di Renzo, and C. Verikoukis, "Energy- and spectrum-efficient user association in millimeter-wave backhaul small-cell networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1810–1821, Feb. 2017.

[16] E. Balas and M. W. Padberg, "Set partitioning: A survey," *SIAM Rev.*, vol. 18, no. 4, pp. 710–760, Oct. 1976.

[17] F. Z. Kaddour, M. Pischella, P. Martins, E. Vivier, and L. Mroueh, "Opportunistic and efficient resource block allocation algorithms for LTE uplink networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2013, pp. 487–492.

[18] S. Martello and P. Toth, *Knapsack Problems: Algorithms and Computer Implementations*. New York, NY, USA: Wiley, 1990.

[19] S. Martello and P. Toth, "Generalized assignment problems," in *Algorithms and Computation*. Berlin, Germany: Springer, 1992, pp. 351–369.

[20] H. P. Keeler, B. Blaszczyszyn, and M. K. Karray, "SINR-based k-coverage probability in cellular networks with arbitrary shadowing," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2013, pp. 1167–1171.

[21] O. Yaman, A. Eroglu, and E. Onur, "Density-aware cell zooming," in *Proc. 21st Conf. Innov. Clouds, Internet Netw. Workshops (ICIN)*, Feb. 2018, pp. 1–8.

[22] S. Mollahasani and E. Onur, "Density-aware power allocation in mobile networks using edge computing," in *Proc. 26th Signal Process. Commun. Appl. Conf. (SIU)*, May 2018, pp. 1–4.

[23] A. Rajesh and R. Achar, "A review of heterogeneous resource allocation in LTE-A based Femto cell networks with CQI feedback," *Indian J. Sci. Technol.*, vol. 9, no. 36, pp. 1–16, Sep. 2016. [Online]. Available: http://www.indjst.org/index.php/indjst/article/view/102109

[24] M. Yassin, S. Lahoud, M. Ibrahim, and K. Khawam, "A downlink power control heuristic algorithm for LTE networks," in *Proc. 21st Int. Conf. Telecommun. (ICT)*, May 2014, pp. 323–327.

[25] M. Rupp, S. Schwarz, and M. Taranetz, *The Vienna LTE-Advanced Simulators: Up and Downlink, Link and System Level Simulation* (Signals and Communication Technology), 1st ed. Singapore: Springer, 2016.

[26] G. Karagiannis, G. T. Pham, A. D. Nguyen, G. J. Heijenk, B. R. Haverkort, and F. Campfens, "Performance of LTE for smart grid communications," in *Proc. 17th Int. GI/ITG Conf. Meas., Modelling, Eval. Comput. Syst. Dependability Fault Tolerance*, vol. 8376, Mar. 2014, pp. 225–239.

[27] *Evolved Universal Terrestrial Radio Access (EUTRA)*, document TR 36.942, 3GPP, Radio Frequency (RF) System Scenarios, Oct. 2014.

[28] B. Clerckx, G. Kim, and S. Kim, "MU-MIMO with channel statistics-based codebooks in spatially correlated channels," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov./Dec. 2008, pp. 1–5.

[29] G. Andersson, A. Vastberg, A. Devlic, and C. Cavdar, "Energy efficient heterogeneous network deployment with cell DTX," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.

[30] M. Feng, S. Mao, and T. Jiang, "Base station ON-OFF switching in 5G wireless networks: Approaches and challenges," *IEEE Wireless Commun.*, vol. 24, no. 4, pp. 46–54, Aug. 2017.

[31] D. Migliorini, G. Stea, M. Caretti, and D. Sabella, "Power-aware allocation of MBSFN subframes using discontinuous cell transmission in LTE systems," in *Proc. IEEE 78th Veh. Technol. Conf. (VTC Fall)*, Sep. 2013, pp. 1–5.

[32] C.-T. Tung, Y.-L. Chung, and Z. Tsai, "An efficient power-saving downlink transmission scheme in OFDM-based multiple component carrier systems," in *Proc. 14th Int. Conf. Adv. Commun. Technol. (ICACT)*, Feb. 2012, pp. 116–120.

[33] V. Mor and H. Kumar, "Energy-efficient approaches in wireless network: A review," in *Next-Generation Networks*. Singapore: Springer, Nov. 2018, pp. 135–146.

[34] Z. Zhou, M. Dong, K. Ota, G. Wang, and L. T. Yang, "Energy-efficient resource allocation for D2D communications underlaying cloud-RAN-based LTE-A networks," *IEEE Internet Things J.*, vol. 3, no. 3, pp. 428–438, Jun. 2016.

[35] F. Z. Kaddour, E. Vivier, L. Mroueh, M. Pischella, and P. Martins, "Green opportunistic and efficient resource block allocation algorithm for LTE uplink networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 10, pp. 4537–4550, Oct. 2015.

[36] S. Samarakoon, M. Bennis, W. Saad, M. Debbah, and M. Latva-Aho, "Ultra dense small cell networks: Turning density into energy efficiency," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1267–1280, May 2016.

[37] E. Oh, K. Son, and B. Krishnamachari, "Dynamic base station switching-on/off strategies for green cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2126–2136, May 2013.

[38] A. Bousia, A. Antonopoulos, L. Alonso, and C. Verikoukis, "'Green' distance-aware base station sleeping algorithm in LTE-Advanced," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2012, pp. 1347–1351.

[39] D. Huang, W. Wei, Y. Gao, M. Hou, Y. Li, and H. Song, "Energy efficient dynamic optimal control of LTE base stations: Solution and trade-off," *Telecommun. Syst.*, vol. 66, no. 4, pp. 701–712, Dec. 2017.

[40] M. F. Hossain, K. S. Munasinghe, and A. Jamalipour, "Toward self-organizing sectorization of LTE eNBs for energy efficient network operation under QoS constraints," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2013, pp. 1279–1284.

[41] D. B. Rawat and S. R. Reddy, "Software defined networking architecture, security and energy efficiency: A survey," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 325–346, 1st Quart., 2017.

[42] S. A. Ahmad and D. Datla, "Distributed power allocations in heterogeneous networks with dual connectivity using backhaul state information," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4574–4581, Aug. 2015.

[43] H.-H. Nguyen and W.-J. Hwang, "Distributed scheduling and discrete power control for energy efficiency in multi-cell networks," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2198–2201, Dec. 2015.

[44] Y. Li, W. Liu, B. Cao, and M. Li, "Green resource allocation in LTE system for unbalanced low load networks," in *Proc. IEEE 23rd Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2012, pp. 1009–1014.

[45] S. Almowuena, M. M. Rahman, C.-H. Hsu, A. A. Hassan, and M. Hefeeda, "Energy-aware and bandwidth-efficient hybrid video streaming over mobile networks," *IEEE Trans. Multimedia*, vol. 18, no. 1, pp. 102–115, Jan. 2016.

[46] A. M. Abdelhady, O. Amin, and M.-S. Alouini, "Energy-efficient resource allocation for phantom cellular networks with imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3799–3813, Jun. 2017.

[47] T. Taleb, K. Samdanis, B. Mada, H. Flinck, S. Dutta, and D. Sabella, "On multi-access edge computing: A survey of the emerging 5G network edge cloud architecture and orchestration," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1657–1681, 3rd Quart., 2017.

[48] S. Jošilo and G. Dán, "Selfish decentralized computation offloading for mobile cloud computing in dense wireless networks," *IEEE Trans. Mobile Comput.*, vol. 18, no. 1, pp. 207–220, Jan. 2019.

[49] G. Qiao, S. Leng, and Y. Zhang, "Online learning and optimization for computation offloading in D2D edge computing and networks," *Mobile Netw. Appl.*, vol. 24, pp. 1–12, Jan. 2019. doi: 10.1007/s11036-018-1176-y.

[50] P. Shantharama, A. S. Thyagaturu, N. Karakoc, L. Ferrari, M. Reisslein, and A. Scaglione, "LayBack: SDN management of multi-access edge computing (MEC) for network access services and radio resource sharing," *IEEE Access*, vol. 6, pp. 57545–57561, Sep. 2018.

[51] J. Wen, C. Ren, and A. K. Sangaiah, "Energy-efficient device-to-device edge computing network: An approach offloading both traffic and computation," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 96–102, Sep. 2018.

[52] A. J. Kadhim and S. A. H. Seno, "Energy-efficient multicast routing protocol based on SDN and fog computing for vehicular networks," *Ad Hoc Netw.*, vol. 84, pp. 68–81, Mar. 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1570870518301616

[53] H. Sun, F. Zhou, and R. Q. Hu, "Joint offloading and computation energy efficiency maximization in a mobile edge computing system," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3052–3056, Mar. 2019.

[54] P. Liu, G. Xu, K. Yang, K. Wang, and X. Meng, "Jointly optimized energy-minimal resource allocation in cache-enhanced mobile edge computing systems," *IEEE Access*, vol. 7, pp. 3336–3347, Jan. 2019.

[55] A. K. Rangisetti and B. R. Tamma, "Interference and QoS aware cell switch-off strategy for software defined LTE HetNets," *J. Netw. Comput. Appl.*, vol. 125, pp. 115–129, Jan. 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1084804518303163

SHAHRAM MOLLAHASANI received the B.Sc. degree in computer hardware engineering from IAU, Kashan, Iran, in 2010, and the M.Sc. degree in computer engineering from EMU, Famagusta, Cyprus, in 2013. He is currently pursuing the Ph.D. degree in computer engineering with Middle East Technical University (METU), Ankara, Turkey. He has been a Research Assistant with METU. His research interests include wireless communications, green networks, and 5G mobile networks. He is a member of the Wireless Systems, Networks and Cybersecurity Laboratory (WINSLab), METU. He is a member of the IEEE.

ERTAN ONUR received the B.Sc. degree from Ege University, Izmir, Turkey, in 1997, and the M.Sc. and Ph.D. degrees from Bogazici University, Istanbul, Turkey, in 2001 and 2007, respectively, all in computer engineering. During the M.Sc. and Ph.D. degrees, he was a Project Leader with Global Bilgi, Istanbul, and as an R&D Project Manager with Argela Technologies, Istanbul. He developed and managed many commercial telecommunications applications. After obtaining his Ph.D. degree, he was an Assistant Professor with the Delft University of Technology, The Netherlands. Since 2014, he has been with Middle East Technical University, Turkey. He was an Editor/Convenor of the Personal Networks Group of Ecma International Standardization Body. His research interest includes computer and wireless networks and their security. He is the Founder of the Wireless Systems, Networks and Cybersecurity Laboratory (WINSLab). He is a member of the IEEE.