A Survey on Green Mobile Networking: From The Perspectives of Network Operators and Mobile Users

Muhammad Ismail, *Member, IEEE,* Weihua Zhuang, *Fellow, IEEE,* Erchin Serpedin, *Fellow, IEEE,* and Khalid Qaraqe, *Senior Member, IEEE*

Abstract—Efficient usage of energy in wireless networks represents a major concern in academia and industry, mainly because of environmental, financial, and quality-of-experience considerations. Various solutions have been proposed to enable efficient energy usage in wireless networks, and these approaches are referred to as *green* wireless communications and networking. In this survey, we mainly focus on energy efficient techniques in base stations and mobile terminals as they constitute the major sources of energy consumption in wireless access networks, from the operator and user perspectives, respectively. Unlike the existing articles and surveys, we aim to present a unified treatment of green solutions and analytical models for both network operators and mobile users. Such a unified treatment will help in the future to develop green solutions that enable an improved and balanced efficient usage of energy by operators and end users.

Index Terms—Energy efficiency, green communications, power consumption modeling, resource on-off switching, scheduling techniques, traffic modeling.

I. INTRODUCTION

During the past decade, there has been an increasing demand for wireless communication services, which have extended beyond telephony services to include video streaming and data applications [1]. This trend has been accompanied by a wide deployment of wireless access networks. In general, a wireless access network is defined as a wireless system that uses radio base stations (BSs) or access points (APs) to interface mobile terminals (MTs) with the core network or the Internet [2]. Hence, a wireless access network is mainly composed of BSs/APs, core network, and MTs [3]. The BSs/APs are responsible for radio resource management and user mobility management, and provide access to the Internet. The core network serves as a backbone network with Internet connectivity and provides data services [3]. Currently, MTs are equipped with processing and display capabilities that enable them to provide not just voice services but also video streaming and data applications. In addition, MTs have

M. Ismail and K. Qaraqe are with the Department of Electrical and Computer Engineering, Texas A&M University at Qatar, Doha, Qatar, e-mail:{m.ismail, khalid.qaraqe}@qatar.tamu.edu. Part of this work was done while the first author was with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Canada.

W. Zhuang is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Canada, e-mail: wzhuang@uwaterloo.ca.

E. Serpedin is with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77840, USA, email: serpedin@ece.tamu.edu.

multiple radio interfaces and mobile users can enjoy single-network and/or multi-homing services [4] - [6].

A. The Need for Green Communications

The BS is the main source of energy consumption in the wireless access network, from the operator side [2]. It has been estimated that more than 57% of the operator total energy consumption is in the BS [2], [7], [8]. In total, there are about 3 million BSs worldwide that consume 4.5 GW of power [9]. From the user side, it has been estimated that there are around 3 billion MTs in the world with power consumption of $0.2 - 0.4$ GW [10]. The high energy consumption of wireless access networks has promoted increased environmental and financial concerns for both service operators and users, and quality-of-experience (QoE) considerations for the mobile users.

From an environmental perspective, the telecommunications industry contributes 2% of the total CO_2 emissions worldwide, and this percentage is expected to increase to 4% by 2020 [11]. In addition, the expected lifetime of rechargeable batteries is around $2 - 3$ years and manifests in 25000 tons of disposed batteries per year, which triggers environmental concerns (and financial considerations for mobile users as well) [12]. Moreover, the high energy consumption of BSs and MTs results in high heat dissipation and electronic pollution [13]. From a financial perspective, a significant portion of annual operating expenses of a service provider are energy costs [14], [15]. It has been estimated that the cost of energy bills of service providers range from 18% (in mature markets in Europe) to 32% (in India) of their operational expenditure (OPEX) [16], [17]. For cellular networks outside the power grid, the energy expenses reach up to 50% of the OPEX [18], [19]. Finally, from a user QoE perspective, it has been shown that more than 60% of mobile users complain about their limited battery capacity [20]. The gap between the demand for energy and the MT offered battery capacity is increasing exponentially with time [21]. Hence, the MT operational time between battery charging has become a significant factor in the user perceived quality-of-service (QoS) [22].

Due to the aforementioned concerns, there has been an increasing demand for energy efficient solutions in wireless access networks. The research works carried out in this direction are referred to as *green* solutions. The term *green* emphasizes the environmental dimension of the proposed solutions. Therefore, it is not sufficient to present a cost

effective solution if it is not friendly to the environment. For instance, it is not acceptable within the green paradigm to have a cost effective electricity demand schedule for a network operator that relies on different electricity retailers, in a liberated electricity market, without ensuring that the proposed solution is also environmental friendly in terms of the resulting CO_2 emissions [23]. The green wireless communications and networking paradigm aims to reduce energy consumption of communication devices while taking into account the environmental impacts of the proposed solutions.

B. Technical Contributions

In the current literature, there exist several surveys that present different approaches for energy efficient communications and networking solutions for wired networks [24], optical networks [25], wireless networks [17], [26] - [28], and mobile users [29] - [31]. The existing surveys investigate energy efficiency either from a network operator perspective [24] - [28] or a mobile user perspective [29] - [31]. However, green solutions should balance energy efficiency between network operators and mobile users. There exists a trade-off in energy saving between service providers and mobile users, yet the existing research does not account for it. For instance, consider the BS on-off switching approach that is proposed in literature to save energy for network operators in a low call traffic load condition. This approach can lead to higher energy consumption for MTs in the uplink due to a larger transmission distance. In this case, such a solution only shifts the energy consumption burden from the network operator to the mobile users, leading to battery drain for MTs at a faster rate. Therfore, it is necessary to develop a better understanding of green solutions for both network operators and mobile users, in order to develop energy efficient solutions that account for and balance the trade-off in energy saving between network operators and mobile users. In addition, call traffic dynamics and power consumption modeling play a significant role in developing effective green solutions, but have not been carefully studied in literature. Moreover, other performance metrics conflicting with achieving energy efficiency should also be considered. The contributions of this survey are next briefly summarized[1]:

- A presentation of various power consumption models for BSs and MTs to capture transmission power, circuit power, and reception power consumption.
- A detailed description of different energy efficiency definitions proposed in literature, under different call traffic load conditions for network operators and under different networking settings for mobile users (e.g., single-user or multi-user with or without fairness considerations).
- A review of different models proposed for call traffic load dynamics that take into account the spatial and temporal fluctuations at different scales (long-term, short-term, flow-level, and packet-level).
- An overview of the performance metrics conflicting with achieving energy efficiency in wireless communications

and networking.
- A unified treatment of energy efficient solutions for network operators and mobile users. Specifically, we classify the energy efficient solutions based on the call traffic load condition into low and high call traffic load solutions. Using such a classification, we discuss the similar approaches adopted by network operators and mobile users to save energy for BSs and MTs.
- Identifying future research directions that help to develop effective green solutions, which can balance energy saving among network operators and mobile users.

C. Organization

Throughout this study, we aim to present a complete picture of energy efficient (green) models and solutions for BSs and MTs that enable new approaches to balance the energy saving for both network operators and mobile users. In order to develop/analyze a green communication solution, an appropriate definition of energy efficiency for network operators and mobile users should be introduced. Such a definition relies on power consumption, throughput, and traffic load models for network operators and mobile users. Moreover, the green communication solution should satisfy some conflicting performance metrics. Hence, the first part of the paper is dedicated to energy efficiency definitions and power consumption, throughput, and traffic load models for network operators and mobile users, along with conflicting performance metrics. After introducing the background concepts, the second part of the paper focuses on state-of-the-art green communication solutions and analytical models for network operators and mobile users at different traffic load conditions. Finally, we discuss the impact of green communication solutions from the perspectives of network operators and mobile users, respectively, aiming to balance the existing trade-offs.

The rest of this paper is organized as follows. In Section II, different definitions are presented to describe energy efficiency in wireless networks at different traffic load conditions, along with throughput and power consumption models, from the network operator and mobile user perspectives. Section III discusses different models to capture the temporal and spatial fluctuations in traffic load, along with some performance metrics that conflict with the general objective of achieving energy efficiency in wireless networks. Green solutions and analytical models are reviewed for network operators and mobile users at low and high traffic load conditions in Sections IV and V, respectively. Finally, future research directions and conclusions are given in Sections VI and VII, respectively.

II. MODELING OF ENERGY EFFICIENCY IN WIRELESS NETWORKS

In this section, we present different definitions that have been proposed in the literature to assess the energy efficiency of wireless networks from the network operator and mobile user perspective, respectively. As an important component of energy efciency denitions, we rst present different throughput and power consumption models for BSs and MTs.

---

[1]In this survey, we mainly focus on direct communications (i.e., no multi-hop or relaying techniques) for licensed networks and users.

A. Throughput Models

In what follows, we first introduce the concepts of aggregate network capacity C_n, area spectral efficiency T_n, and the user achieved data rate R_m. These concepts are necessary in the energy efficiency definitions that we will later present.

To measure the network aggregate capacity C_n, Shannon formula is used [23]

$$C_n = B_n \log_2 \det(I + PHH^{\dagger}), \qquad (1)$$

where B_n represents network n total bandwidth, I denotes the identity matrix, P stands for the transmission power matrix of every BS of network n to every MT m in service, H is the channel matrix between each BS of network n and MT m, and $\dagger$ denotes the transposition operation. The channel matrix H might account for fast fading, noise, and interference affecting the radio transmission. C_n in (1) has the unit of bits per second. On the other hand, in a low call traffic load condition, the area spectral efficiency T_n is used in the energy efficiency definition rather than the aggregate capacity given in (1) [15], as will be shown. The area spectral efficiency measures the network throughput while considering the coverage probability. Let $\mathbb{P}(\gamma_{x \to u} > \zeta)$ denote the success probability of the signal-to-noise ratio (SNR) γ of a BS at location x received by an MT at location u, satisfying a certain QoS threshold ζ. The coverage probability $\mathbb{P}_n(\zeta)$ is obtained by averaging the success probability over the propagation range to location u. Hence, for network n with BS density λ_n, the area spectral efficiency T_n is given by

$$T_n = \lambda_n \mathbb{P}_n(\zeta) \log_2(1 + \zeta) \qquad (2)$$

where T_n is measured over a unit area.

In literature, two key definitions have been proposed to assess the data rate R_m achieved by MT m. The first definition is based on the knowledge of the instantaneous channel state information (CSI) [13], [32] - [34], and takes the form

$$R_m = B_m \log_2(1 + \frac{\gamma_m}{\Gamma}), \qquad (3)$$

where B_m denotes the allocated bandwidth on the uplink to MT m, γ_m is the SNR of MT m received at the destination, and Γ is the SNR gap between the channel capacity and a practical coding and modulation scheme. For Shannon formula, $\Gamma = 1$. However, the instantaneous CSI requires a feedback from each MT to the serving BS, which results in a large overhead. Hence, statistical CSI can be used to reduce the amount of overhead. In this case, R_m is described in a statistical average sense [35], i.e.,

$$R_m = E_H[B_m \log_2(1 + \frac{\gamma_m}{\Gamma})], \qquad (4)$$

where E_H denotes the expectation over channel state H, which affects the SNR γ_m. R_m in (3) and (4) is measured in bits per second.

B. Power Consumption Models

Different models have been proposed in literature to represent power consumption for the network, P_n, and MTs, P_m, as shown in Table I.
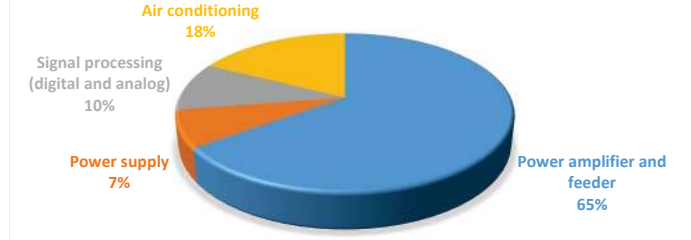

Fig. 1. Percentage of power consumption at different components of a large-cell BS [9].

The network power consumption P_n can be modeled as the aggregate power consumption of the network BSs. Practical measurements of power consumption at BSs are reported in [7], [9], [19], [36]. Figure 1 shows the percentage of power consumption at different components of a large-cell BS. On the other hand, Table II shows the power consumption profile for a femto-cell BS. By comparing the values in Figure 1 and Table II, we observe that a femto-cell BS consumes most of the power in the signal processing part as opposed to a large-cell BS (65.6% and 10% for femto and large-cell BSs, respectively). In addition, the RF transmission/reception energy consumption part in a femto-cell BS is almost half of that of a large-cell BS, with only 19.6% of the power consumed in the femto-cell BS power amplifier as opposed to 65% in a large-cell BS.

Let P_b denote the power consumption for BS b. In literature, different models are used to represent P_b. The first one is an ideal load dependent power model [37]. This model assumes that the BS consists only of energy proportional devices and hence, it assumes no power consumption in an idle state, i.e.,

$$P_b = \rho P_{tb}, \qquad (5)$$

where ρ stands for the system traffic load density and P_{tb} denotes the BS transmitted power. However, such a model is not realistic as the power consumption of some components shown in Figure 1 does not scale with the call traffic load. A more sophisticated model captures the power consumption of different BS components, and assumes the expression [38]

$$P_b = \frac{\frac{P_{tb}}{\xi(1 - \sigma_{\text{feed}})} + P_{\text{RF}} + P_{\text{BB}}}{(1 - \sigma_{\text{DC}})(1 - \sigma_{\text{MS}})(1 - \sigma_{\text{cool}})}, \qquad (6)$$

where P_{RF} is the radio frequency (RF) power consumption, P_{BB} represents the baseband unit power consumption, ξ denotes the power amplifier efficiency, and σ_{feed}, σ_{DC}, σ_{MS}, and σ_{cool} stand for the losses incurred by the antenna feeder, DC - DC power supply, main supply, and active cooling, respectively. For simplicity, the model (6) can be further approximated by a linear model [11], [15], [18], [23], [37], [38]. In such a linear model, two components are introduced to represent P_b. The first component is a fixed power term that captures the power consumption at the power supply, cooling, backhaul, and other circuits, and is denoted by P_f. The second component is proportional to the traffic load. The linear (affine) model is described by

$$P_b = \Delta_p P_{tb} + P_f, \qquad (7)$$

TABLE I
A SUMMARY OF DIFFERENT POWER MODELS IN LITERATURE.

	Model			Comments	References
BS	Operation only	Large-cell	Ideal	The BS consumes no power when idle, i.e., the BS consists only of energy proportional devices.	[37]
			Realistic	The model captures the BS traffic load independent power consumption.	[11], [15], [18], [23], [37], [38]
		Femto-cell	Load independent	The BS power consumption does not depend on the offered traffic load.	[39]
			Load dependent	The BS power consumption relies on traffic load, packet size, and has an idle part.	[40]
	Operation and Embodied			Besides the operation power, it accounts for the consumed energy in BS manufacturing and maintenance.	[16]
MT	Transmission power only		Without power amplifier efficiency	The models does not account for the transmitter power amplifier efficiency.	[34], [42], [43]
			With power amplifier efficiency	The model accounts for the transmitter power amplifier efficiency.	[13], [32], [44], [45]
	Including circuit power		Constant	The circuit power consumption is given by a constant term independent of the bandwidth and data rate.	[13], [20], [32], [33], [35], [46]
			Bandwidth scale	The circuit power consumption scales with the MT assigned bandwidth.	[47]
			Data rate scale	The circuit power consumption scales with the MT achieved data rate.	[45]
	Including reception power			Besides the transmitter and circuit power consumption, the model also accounts for the receiver power consumption.	[13], [47]

TABLE II
POWER CONSUMPTION PROFILE FOR A FEMTO-CELL BS [36].

Hardware component	Power Consumption (W)	Percentage (%)
Microprocessor	1.7	
Associated memory	0.5	26.4
Backhaul circuitry	0.5	
FPGA	2	
Associated memory	0.5	39.2
Other hardware functions	1.5	
RF transmitter	1	
RF receiver	0.5	34.3
RF power amplifier	2	

where Δ_p is the slope of the load dependent power consumption.

For a femto-cell BS, the power consumption model [39] is expressed as

$$P_b = P_{mp} + P_{FPGA} + P_{tx} + P_{amp}, \qquad (8)$$

where P_{mp}, P_{FPGA}, P_{tx}, and P_{amp} denote the power consumption of the microprocessor, FPGA, transmitter, and power amplifier. While the power consumption model in (8) accounts for most of the components in Table II, it does not exhibit any dependence on the call traffic load. Experimental results in [40] have illustrated the dependence of the femto-cell BS power consumption on the offered load and the data packet size. Hence, the power consumption model for a femto-cell BS is given by [40]

$$P_b = P_b(q, s) + P_f, \qquad (9)$$

where $P_b(q, s)$ and P_f denote the BS power consumption that relies on the traffic load q (expressed in Mbps) and packet size s (expressed in bytes) and the idle power consumption component, respectively.

The models (5) - (9) focus mainly on the BS operation power, which is expressed in watts. In a more general model, the power consumption is described in terms of the BS operating energy, E_o, and embodied energy, E_e, which represents $30-40\%$ of the BS total energy consumption [16]. The embodied energy accounts for the energy consumed by all processes associated with the manufacturing and maintenance of the BS, and is calculated as 75 GJ over the BS lifetime [16]. It consists of two components, the first refers to the initial embodied energy E_{ei} and comprises the energy used to acquire and process raw materials, manufacture components, and assemble and install all BS components, and it is accounted for only once in the initial BS manufacturing. The second component stands for the maintenance embodied energy E_{em} and includes the energy associated with maintaining, repairing, and replacing the materials and components of the BS throughout its lifetime. Hence, the BS total energy consumption, in joules, throughout its lifetime is expressed as

$$E_b = E_e + E_o = (E_{ei} + E_{em}) + E_o, \qquad (10)$$

where $E_{em} = P_{em}T_{lifetime}$ and P_{em} and $T_{lifetime}$ denote the BS maintenance power and lifetime, respectively, and $E_o = P_o T_{lifetime}$, where P_o is defined in terms of the BS operating power given in (5) - (9). The model in (10) is useful for measuring the BS total power consumption in the network design stage, for instance, while designing a multi-tier wireless network, as will be discussed.

Practical measurements of power consumption at MTs are summarized based on different experiments in [29]. Besides the results reported in [29], we present experimental results from [41] and [31], in Tables III and IV, respectively, to discuss the models to be presented later. Different models have been proposed in literature for the MT power consumption P_m. In the simplest model, P_m represents the MT transmission power P_{tm} [34], [42], [43]. When the effect of the power amplifier efficiency is considered, the MT power consumption, in watts, is expressed as [13], [32], [44], [45]

$$P_m = \frac{P_{tm}}{\zeta_m}, \qquad (11)$$

where ζ_m denotes the power amplifier efficiency for MT m, $\zeta_m \in (0, 1]$. With such a power consumption model, for a data call, the minimum energy consumption is attained by using the modulation of the lowest order while satisfying the QoS constraints (e.g., time delay) [33]. However, in practice, the MT circuit power should be captured in the power consumption model P_m. Three different models have been proposed to capture the MT circuit power P_{cm}. In the first model, the circuit power consumption is modeled as a constant, independent of the achieved data rate R_m [13], [20], [32], [33], [35], [46]. With constant circuit power consumption, transmitting with the lowest modulation order is no longer the best strategy as the energy consumption is proportional to the transmission duration [33]. The constant power consumption model, however, does not reflect the effect of transmission bandwidth and data rate on the MT circuit power consumption. From Table III, it is evident that different radio interfaces consume different circuit power, which for one reason is due to the different operating bandwidth. To account for the effect of the allocated bandwidth, the circuit power consumption is modeled via two components [47]. The first component refers to the digital circuit power consumption, which is modeled as a linear function of the transmission bandwidth (as bandwidth increases, more computations and baseband processing are required), i.e.,

$$P_{bm} = P_b^{ref} + \alpha \frac{B_m}{B_{ref}}, \qquad (12)$$

where P_b^{ref} denotes the reference digital circuit power consumption, in watt, for a reference bandwidth B_{ref} and α is a proportionality constant. The second component captures the power consumption of the radio frequency (RF) chain, represented by a constant to account for power consumption in the digital-to-analog converter, RF filter, local oscillator, and mixer. However, the model in (12) does not account for the effect of transmission data rate on the power consumption, as indicated in Table IV. To account for the transmission data rate, the circuit power consumption is modeled as a linear function of the achieved data rate, under the assumption that the clock frequency of the MT digital chips scales with the achieved data rate [45]. Therefore, the circuit power consumption is expressed as

$$P_{cm} = \mu + \beta R_m, \qquad (13)$$

where μ and β are two appropriately chosen constants, measured in watt and watt per bit per second, respectively. Besides the transmission and circuit power modeling in P_m, a constant term is considered to represent the MT receiver circuit power consumption [13], [47].

It is worth mentioning that R_m and P_m can be defined as the sum of corresponding terms over multiple subcarriers assigned to MT m for orthogonal frequency division multiple access (OFDMA) networks [13], [32] - [34], [42], [43] or the sum

TABLE III
MT POWER CONSUMPTION FOR DIFFERENT TECHNOLOGIES [41].

Technology	Action	Power (mw)
WiFi IEEE 802.11 (Infrastructure mode)	In connection	868
	In disconnection	135
	Idle	58
	Idle in power save mode	26
	Downloading at 4.5 Mbps	1450
WiFi IEEE 802.11 (ad hoc mode)	Sending at 700 kbps	1629
	Receiving	1375
	Idle	979
2G	Downloading at 44 kbps	500
	Handover to 3G	1389
3G	Downloading at 1 Mbps	1400
	Handover to 2G	591

TABLE IV
MT POWER CONSUMPTION FOR DIFFERENT DATA RATES OF AUDIO STREAMING AND DOWNLOADING A 200 MB FILE USING WIFI [31].

Bit Rate (kbps)	Nokia E-71 (mW)	Nexus S (mW)	Samsung Galaxy S3 (mW)
128	990	350	419
192	1004	390	440
256	1007	390	452
File download	1092	998	1012

over multiple radio interfaces for MT m with multi-homing capability [20], [46].

C. Energy Efficiency Models

Based on the throughput and power consumption models discussed before, we next present different energy efficiency definitions for networks and MTs in this subsection. A list describing the energy efficiency definitions proposed in literature is presented in Table V.

A generic definition that can be used as a measure of energy efficiency is referred to as energy consumption gain, and it is defined as the ratio of the energy consumed by a base system (BS or MT) to the energy consumed by the system under test, assuming the same conditions [7], [48]. Formally, this is expressed by

$$\eta = \frac{E_{\text{base}} - E_{\text{test}}}{E_{\text{base}}} \quad (14)$$

where E_{base} and E_{test} are measured in joules. The definition in (14) is a relative definition that can be used in any call traffic load condition. Next, we present the absolute energy efficiency definitions.

In a low call traffic load, the mobile user demands do not require that the network operates at its full power. Therefore, one way to measure energy efficiency for network operators at a low call traffic load condition is by means of the ratio between the network output power (energy) and the total input power (energy) [2], [19], i.e., the energy efficiency η_n for

network n is expressed as

$$\eta_n = \frac{P_t}{P_n}, \quad (15)$$

where P_t and P_n denote the network output power (i.e., the power of the RF transmitted signal) and input (consumed) power, respectively. Hence, η_n is unitless. In addition, due to the low service demands, it is not necessary to guarantee that the network achieves a full coverage. It is sufficient to satisfy an acceptable coverage probability. As the definition in (15) does not reflect the achieved network coverage, another definition of energy efficiency measures the power consumed to cover a certain area [15], [19]. Hence, energy efficiency can be defined as [15]

$$\eta_n = \frac{T_n}{P_n}. \quad (16)$$

In (16), η_n has unit of watt^{-1}.

From an operator perspective, in a high call traffic load condition, energy efficiency is defined as the ratio of the aggregate network capacity to the total power consumed by the entire network [2], [9]. Therefore, in a high call traffic load condition, energy efficiency of network n isexpressed as

$$\eta_n = \frac{C_n}{P_n}, \quad (17)$$

in bit per second per watt.

For mobile users, energy efficiency is expressed as a measure of the maximum amount of bits that can be delivered per

TABLE V
A SUMMARY OF DIFFERENT ENERGY EFFICIENCY DEFINITIONS IN LITERATURE.

	Model		Comments	References
BS/MT	Energy consumption gain		A ratio of the energy consumed by a base system to the energy consumed by the system under test. It is a relative measure that can be used at any traffic load.	[7], [48]
BS	Low traffic load	Output – input power	A ratio of network output to input power.	[2], [19]
		Area spectral efficiency – input power	The definition measures the power consumed for certain area coverage. It is used at low traffic load.	[15], [19]
	High traffic load Network capacity – input power		A ratio of the aggregate network capacity to the total power consumed by the network.	[2], [9]
MT	Single-user system	Without error consideration	A ratio of throughput to power consumption.	[13], [20], [32], [42], [46]
		With error consideration	A ratio of goodput to power consumption.	[42], [43]
	Multi-user system	Without fairness consideration	It can be sum rate of all MTs to total power consumption or sum of energy efficiency for individual MTs.	[33] – [35], [43], [44]
		With fairness consideration	It is a geometric mean of energy efficiencies of all MTs.	[33]

joule of energy consumed [13], [20], [32], [42], [46]. This can be described, for mobile user m, as

$$\eta_m = \frac{R_m \Delta T}{\Delta E_m} = \frac{R_m}{\Delta E_m / \Delta T} = \frac{R_m}{P_m}, \qquad (18)$$

where ΔE_m stands for the consumed energy during time interval ΔT by MT m. However, the definition in (18) does not account for the energy consumed for correct reception of data. Another definition measures the net number of information bits that are transmitted without error per joule [42], [43], and it is expressed as

$$\eta_m = \frac{R_m f(\gamma_m)}{P_m}, \qquad (19)$$

where $f(\gamma_m)$ stands for the packet transmission success rate for a given SNR γ_m for MT m. The definition in (19) assumes a ratio of the goodput to power consumption, as compared to the definition in (18) which assumes the ratio of throughput to power consumption. In general, the packet transmission success rate follows an S-shaped (sigmoidal) function, exhibiting an increasing trend with respect to γ_m. In addition, $f(\gamma_m)$ approaches zero as γ_m approaches zero and $f(\gamma_m)$ approaches unity as γ_m approaches infinity [42], [43]. The unit of η_m in (18) and (19) is bit per second per watt.

The definitions in (18) and (19) have been proposed for a single-user scenario [13], [20], [32], [42], [46]. A multi-user system is considered in various scenarios due to the shared bandwidth [33] or interference [34] caused by simultaneous transmissions. In a multi-user scenario, energy efficiency is defined as the ratio between the sum rate of all MTs to the total power consumption [44]

$$\eta_{\text{total}} = \frac{\sum_m R_m}{\sum_m P_m}. \qquad (20)$$

The definition in (20) treats all mobile users as a single unit, and takes into account only the total achieved throughput and power consumption. In order to model the system as a set of distinct mobile users, another definition is used, which represents the total energy efficiency expressed as the sum of energy efficiency for individual MTs [33] - [35], [43]

$$\eta_{\text{total}} = \sum_m \eta_m. \qquad (21)$$

The unit of η_{total} in (20) and (21) is bit per second per watt. However, the definitions in (20) and (21) provide no fairness guarantee for energy efficiency among different MTs. Consequently, some MTs might present high energy efficiencies while others might exhibit low energy efficiencies very close to zero. The geometric mean of energy efficiencies of all MTs promotes fairness among users [33], and is expressed as

$$\eta_{\text{total}} = \sum_m \log(\eta_m). \qquad (22)$$

Unlike (20) and (21), η_{total} in (22) is unitless.

An energy efficiency definition and an associate green solution must be chosen in accordance with an appropriate traffic load model. Hence, in the next section, we discuss the different traffic models proposed in literature.

III. TRAFFIC MODELING AND PERFORMANCE METRICS

Different green solutions can be employed at different call traffic load conditions. In addition, some green solutions exploit the temporal and spatial fluctuations in the call traffic load to save energy. For instance, in order to determine the switch off duration of a BS or MT, as will be discussed later, traffic models are used to probabilistically determine the idle period duration. Furthermore, performance evaluation of the proposed green solutions should be carried out using an appropriate traffic model. Hence, it is necessary to gain a better understanding of different traffic load models proposed in literature. After specifying an appropriate energy efficiency definition and a traffic load model, the next step is to define the performance metrics that may conflict with the objective of enhancing the system energy efficiency. Therefore, before we discuss various green networking solutions, in this section, we present traffic models for BSs and MTs. Then, we discuss performance metrics that are considered for developing green networking solutions.

Traffic modeling can be classified in two categories, as shown in Table VI. The static model assumes a fixed set of MTs, $\mathcal{M}$, that communicate with a fixed set of BSs, $\mathcal{B}$ [20], [33], [34], [43], [44], [46], [49] - [51]. The static model does not capture the mobility of MTs in terms of their arrivals and departures. Also, the static model does not consider the call-level or packet-level dynamics in terms of call duration, packet arrival, etc. On the other side, the dynamic model captures the spatial and temporal fluctuations of the traffic load, and is discussed in more details next.

A. Traffic Spatial Fluctuation Models

It has been shown that traffic is quite diverse even among closely located BSs [52]. Therefore, different models have been proposed to capture the spatial fluctuations in call traffic load [15], [37], [53].

One approach to model traffic spatial fluctuations is by defining a location-based traffic load density [37]. In this case, a geographic region is served by a set $\mathcal{B}$ of BSs. The geographic region is partitioned into a set of locations. In a location x, the file transfer requests arrive following an inhomogeneous Poisson point process (PPP) with an arrival rate per unit area $\lambda(x)$. The file sizes are independently distributed with mean $1/\mu(x)$ at the location. Hence, the traffic load density is expressed as $\varrho(x) = \lambda(x)/\mu(x) < \infty$, and it is used as a measure to capture the spatial traffic variability.

While the above approach uses a predefined set of BSs, $\mathcal{B}$, with specific locations, an alternative approach defines the locations of available BSs using stochastic geometry theory [15]. In this case, the BSs of network n are located according to a homogeneous PPP, Θ_n, with intensity λ_n in the Euclidean plane. MTs are distributed according to a different independent stationary point process with intensity λ_m. With a stationary PPP Θ_n, the distribution of the distance between an MT and its serving BS, D_m, is the same regardless of the MT exact location. The probability density function (pdf) of D_m is given by [15]

$$f_{D_m}(d) = 2\pi\lambda_n d \exp^{-\lambda_n \pi d^2}, \quad d > 0. \quad (23)$$

The above two models capture the traffic spatial variability among different cells. To model the spatial distribution of MTs within a given cell i, a finite-state Markov chain (FSMC) model is used [53]. This model categorizes the MTs into S classes according to the radius of cell i. Assuming there are M MTs in cell i, an S^M spatial location distribution is considered within the cell. Hence, the FSMC presents $\mathcal{L} = \{L_1, \ldots, L_{S^M}\}$ states. The state transition probability $\Pr\{L_i(t+1) = v_i | L_i(t) = u_i\}$ is the probability of spatial distribution of the MTs within cell i at time slot $t+1$ being v_i given that it was u_i at time slot t, where $u_i = \{u_{i,1}, \ldots, u_{i,M}\}$ and $v_i = \{v_{i,1}, \ldots, v_{i,M}\}$. Using such a model, the dynamic fluctuations in the number of MTs in different regions within the cell can be captured.

B. Traffic Temporal Fluctuation Models

Traffic temporal fluctuations can be observed over two different time scales [11], [38]. The first scale is a long-term traffic fluctuation model that captures the traffic variations over days of the week. This model is very useful in investigating energy efficient solutions for network operators as it captures both high and low call traffic load conditions. The second scale is a short-term traffic fluctuation that models the call (packet) arrivals and departures of the MTs. This model plays an important role in investigating energy efficient resource allocation schemes for MTs and BSs. Next, we will describe the two scales in more details.

1) Long-term Traffic Fluctuations: Real traffic traces show a sinusoidal traffic profile in each cell [14], [54]. Traffic during day time (11 am - 9 pm) is much higher than that during night time (10 pm - 9 am) [52], [14]. Moreover, the traffic profile during weekends and holidays, even during the peak hours, is much lower than that of a normal week day [14]. During a weekday, the traffic profile is 10% less than its peak value 30% of the time, increasing to 43% of the time during weekends [14]. Such a behavior can be modeled using an activity parameter $\psi(t)$ to specify the percentage of active subscribers over time t [38]. Hence, if the population density is p users per km^2, the number of operators is N, each being able to carry $1/N$ of the total traffic volume, and the fraction of subscribers is s_k with an average data rate r_k for terminal type k (e.g., smart phone and tablet), then the traffic demand, in bit per second per km^2, is given by

$$A(t) = \frac{p}{N}\psi(t)\sum_k s_k r_k. \quad (24)$$

It has been shown that the traffic load difference between two consecutive days for 70% of the BSs is less than 20% [52]. Therefore, the long-term fluctuations in call traffic load are estimated from the historical mobile traffic statistics (i.e., the activity parameter $\psi(t)$ and the average data rate r_k can be inferred from historical data).

2) Short-term Traffic Fluctuations: The short-term traffic fluctuation models can be classified into two categories, namely call (flow)-level and packet-level models. The call (flow)-level models are useful in investigating green resource scheduling algorithms at the BSs and MTs in a high call traffic

TABLE VI
A SUMMARY OF DIFFERENT TRAFFIC MODELS.

Model				Comments	References
Static				It does not capture the MT mobility and the traffic dynamics.	[20], [33], [34], [43], [44], [46], [49] − [51]
Dynamic	Spatial	Regional traffic load density		It defines a location-based traffic load density.	[37]
		Stochastic geometry		BSs and MTs are located according to homogeneous Poisson point process.	[15]
		FSMC		It models the spatial distribution of MTs within a cell.	[53]
	Temporal	Long-scale		The model captures traffic fluctuations over the days of the week.	[14], [38], [52], [54]
		Short scale	Flow-level	Poisson-exponential	It models call arrivals as Poisson process and call departures as exponential distribution.	[11], [55] − [57]
				FSMC	The number of calls within a cell is represented by a state in a Markov chain.	[53]
			Packet-level	Infinite buffer	It models the number of backlogged packets in an MT buffer with infinite capacity.	[58]
				Finite buffer	It models the number of backlogged packets in an MT buffer with finite capacity.	[59]

load. One approach that can be used for myopic resource allocation solutions models the call arrivals using a Poisson process with rate λ_a and call durations with an exponential distribution [11], [55] - [57]. For dynamic resource allocation solutions, traffic dynamics in terms of call arrivals and departures are modeled using an FSMC [53]. The number of calls in a given cell i is represented by an M-state Markov chain, with state set $\mathcal{M} = \{0, 1, \ldots, M - 1\}$. The state transition probability $\Pr\{M_i(t + 1) = h_i | M_i(t) = g_i\}$ is the probability of having h_i MTs within cell i at time slot $t + 1$ given that there were g_i MTs at time slot t, where h_i , $g_i \in \mathcal{M}$.

Packet-level traffic models are useful in investigating green resource solutions (on-off switching) at the BSs and MTs, in a low call traffic load condition, through modeling the BS/MT buffer dynamics in terms of packet arrival and transmission [58], [59]. For instance, for an infinite buffer size, the MT buffer dynamics are represented by

$$o_m(t+1) = \max\{o_m(t) + a_m(t+1) - z_m(t), a_m(t+1)\}, \quad (25)$$

where $o_m(t)$, $a_m(t)$, and $z_m(t)$ denote the numbers of backlogged packets in the buffer, arriving packets, and transmitted packets, for MT m in time slot t, respectively. For a buffer with finite size O, the MT buffer dynamics can be expressed as

$$o_m(t + 1) = \min\{o_m(t) + a_m(t + 1) - z_m(t), O\}. \quad (26)$$

The models (25) and (26) are used to design and evaluate optimal on-off switching schemes for radio interfaces of MTs

to achieve energy efficient communications in low call traffic load conditions, a topic which will be addressed later herein survey.

C. Performance Metrics

Improving energy efficiency in wireless networks may conflict with other performance metrics. Therefore, there is a trend on improving energy efficiency while not violating the target performance metric. Such performance metrics can be divided into two main categories. The first category deals with the quality of the ongoing application and hence is referred to as application quality requirements. The second deals with connection establishment and hence is referred to as admission quality requirements. These are summarized as follows.

1) Application Quality Requirements:
- SNR [52], [53]: Improving energy efficiency reduces the transmission power which reduces the overall received SNR. The receiver may not be able to decode the transmitted signal. Hence, a minimum SNR should be satisfied while improving energy efficiency.
- Data rate [20], [46], [50], [51], [53], [56]: With a reduced transmission power level, transmission data rate is reduced. For some applications, a minimum required data rate should be achieved [46], [52], [53] or a constant required data rate should be satisfied [20], [50].
- Delay time [59]: A reduced data rate can lead to a violation of the required delay deadline for delay sensitive

applications. An equivalent representation to ensure a minimum required data rate is not to violate a maximum delay bound for data transmission.

- Video quality [60], [61]: For video streaming applications, using a lower transmission rate for improved energy efficiency can result in video packets missing their delay deadlines. Hence, the resulting video quality is degraded. It is required to maintain the achieved video quality higher than a target value.

2) Admission Quality Requirements:

- Call blocking probability [11], [23]: Improving energy efficiency for wireless networks can be achieved by switching-off some BSs with a low call traffic load, as explained in the next section. This can lead to an increased call blocking probability. Hence, it is required to maintain the call blocking probability below a certain threshold.
- Coverage probability [15]: BSs can improve their energy efficiency by reducing their transmission power. However, this may result in failure in service coverage. It is required to maintain a target performance level in terms of coverage probability $\mathbb{P}_n(\zeta)$, as explained in the previous section.

Given the background provided in the previous sections, in the next section we will present state-of-the-art green communication solutions and analytical models for network operators and mobile users at different traffic load conditions. In green wireless networks, the proposed solutions/models to enhance/analyze energy efficiency can be divided into two categories based on the call traffic load condition. At a low and/or bursty call traffic load, resource on-off switching techniques are adopted, while scheduling techniques are employed at a high and/or continuous call traffic load. These are discussed in detail in the next two sections.

IV. Green Solutions and Analytical Models at Low and/or Bursty Call Traffic Load

In this case, on-off switching of radio resources is adopted to enhance energy efficiency, as shown in Table VII. Network operators can employ on-off switching mechanisms for their BSs at a low call traffic load. Similarly, MTs can switch on-off their radio interfaces in a bursty traffic condition. The related research issues and modeling techniques are discussed in the following.

A. BS On-Off Switching

In network planning, the cell size and capacity are in general designed based on the peak call traffic load. As discussed in Section III, the call traffic load exhibits significant spatial and temporal fluctuations. At a low call traffic load, the network is over-provisioned which results in energy waste. It has been argued that switching off some of the available radio resources (e.g., radio transceivers of BSs) at a low call traffic load can save energy and maintain acceptable performance metrics. However, studies have shown that, when a BS is active, the energy consumption of processing circuits and air conditioner

amounts to 60% of the BS total power consumption (which is represented by the fixed power component in (7)) [51]. Hence, an effective approach for energy saving at a low call traffic load is to switch off some of the network BSs while satisfying the required performance metrics. Switching BSs on and off according to call traffic load conditions is referred to as dynamic planning [11]. Two important issues must be addressed while designing an effective BS switching mechanism, namely, user association and BS operation.

Switching BSs on and off is coupled with the user association problem. In order to switch off some BSs, the call traffic load should be first concentrated in a few BSs, which is achieved through user association. Newly incoming MTs have to be associated with a subset of active BSs, and MTs already in service should perform handover when the serving BSs are switching off. One can identify two research directions related to the MT association problem. The first direction deals with developing new user association mechanisms [37], [51], [62] - [64], while the second one focuses on deriving analytical models to assess the performance of different association mechanisms [55]. In developing a MT association mechanism, two approaches can be adopted to meet the MT required QoS while concentrating traffic load in a few BSs. The first approach adopts an objective function that minimizes the networks' energy consumption while satisfying the user required QoS constraints, while the second approach aims to balance the trade-off between flow level performance for MTs (e.g., data rate or delay) and energy consumption of the network [37]. In the later case, the problem is a multi-objective one with a weighting factor. When the weighting factor equals zero, the MT association is determined based on the flow level performance and, as the weighting factor increases, the MT association decision focuses more on the network power consumption performance. As the weighting factor goes to infinity, the MT connects to the BS that maximizes the network energy efficiency performance in bits per joule. The MT association mechanism can be implemented in a centralized or a decentralized architecture [51]. Both architectures aim to concentrate MTs in a few BSs while satisfying the data rate requirements of MTs and bandwidth limitations of BSs. In the centralized mechanism, a central controller performs MT association based on global network information that is related to channel conditions and user requirements. On the other hand, an MT locally selects the BS with the highest call traffic load that can serve its required data rate in the iterative decentralized mechanism. One challenge with designing such a mechanism is related to computational complexity, due to the binary nature of the decision variables related to the BS on-off switching, and hence the mixed-integer nature of the optimization problem. Therefore, greedy algorithms are mainly adopted to reach a good switching decision [37], [51]. In designing such algorithms, a decision criterion should be defined. For instance, in a user-BS distance decision criterion, the greedy algorithm tends to switch off the BSs with the longest user-BS distance to improve energy efficiency of the network [62]. The rationale behind such a decision criterion is that the longer the user-BS distance, the greater the transmission power required to meet the target service quality of the

TABLE VII
A SUMMARY OF GREEN SOLUTIONS AND ANALYTICAL MODELS AT LOW AND/OR BURSTY CALL TRAFFIC LOAD.

Solution/Analytical Model			Comments	References
BS on-off switching	User association		This phase concentrates the MTs in a few BSs to enable switching off other BSs.	[37], [51], [55], [62], [63], [64],
	BS operation		This phase specifies which BSs should be turned off and how.	[11], [51], [65] - [69]
MT radio interface on-off switching	With downlink traffic	Without traffic shaping	An MT switches its radio interface if no data packets are available for the MT at the BS.	[48], [70], [71]
		With traffic shaping	Traffic shaping at the MT or BS is introduced to enable longer idle duration for the MT.	[72], [73], [74]
	With uplink traffic		Besides radio interface on-off switching, an MT controls the transmission power and modulation and coding.	[58], [75]
	With bi-directional traffic		This case deal with presence of both uplink and downlink traffic while switching on and off the MT radio interface.	[76]

users. The network-impact notion is introduced in [63] as a key decision criterion, which quantifies the effect of switching off a given BS on the network performance. Specifically, switching off a given BS results in additional load increments into the neighboring BSs. Besides, switching off a BS can result in a positive impact on the neighboring BSs due to a reduced inter-cell interference. By quantifying the two aforementioned effects, the network-impact criterion modifies the switching off decision as a BS selection problem, aiming at finding the BSs that when switched off leads to the highest network-impact [63]. In addition, an important problem associated with BS on-off switching is related to coverage holes. Hence, another decision metric is related to avoiding coverage holes. In [64], it is shown that finding the optimal set of BSs that minimizes the network power consumption while avoiding coverage holes is closely related to the minimum-weight disk cover problem, which is known to be an NP-hard problem and hence a greedy algorithm is proposed to switch off BSs while maintaining network coverage in polynomial time complexity. To assess the performance of different MT association mechanisms, queueing models are used [55]. Specifically, the MT association process in the overlapped coverage of different BSs is modeled as a customer joining a queue with $V = |\mathcal{B}||\mathcal{M}|$ servers, where $|\mathcal{B}|$ and $|\mathcal{M}|$ denote the number of BSs with overlapped coverage and the maximum number of MTs accommodated in each BS, respectively. Consider a two-BS scenario with three service areas. In service areas 1 and 2, an MT is served by the BS covering that area. In service area 3, an MT can be served by either of BSs with overlapped coverage. A BS is switched off and hence its corresponding $|\mathcal{M}|$ servers are shut down, if no MT is assigned to it. Using the queueing model, analytical expressions are derived for call blocking probability, average number of MTs assigned to each BS, and average power and energy consumed by the network operator to serve one MT [55]. The model can be approximated to account for the case with multiple-BS overlapped coverage.

Based on the MT association phase, the BS operation decision is specified. Hence, BSs with a concentrated call traffic load become active, while light loaded BSs are switched off. The BS operation problem deals with three concerns, namely accommodating future traffic demands, determining BS wake-up instants for switched off BSs, and finally how to implement the BS on-off switching decisions. For the first concern, it should be noted that the BS operation decision lasts for a long duration (i.e., several hours), as frequent BS on and off switching is not desirable due to the increased energy consumption in the BS start-up phase [11] and the unavailable service for the off cells during the decision computation phase [51]. As a result, the BS operation decision should address the future call traffic load either by reserving some resources to account for the future demands [51] or by exploiting the historical call traffic load pattern [11]. In [65], an online stochastic game theoretic algorithm is proposed, where neighboring BSs communicate with each other to predict their traffic profiles, which eventually will lead to optimal switching decisions and result in minimum network energy consumption. As for the second concern dealt with in the BS operation problem, it should be noted that switching off some cells is executed given that active BSs extend their coverage areas to provide service for the cells with inactive BSs. As the call traffic load of the inactive cells increases beyond the capacity limitation of the active BSs, some of the inactive BSs are switched on. Hence,
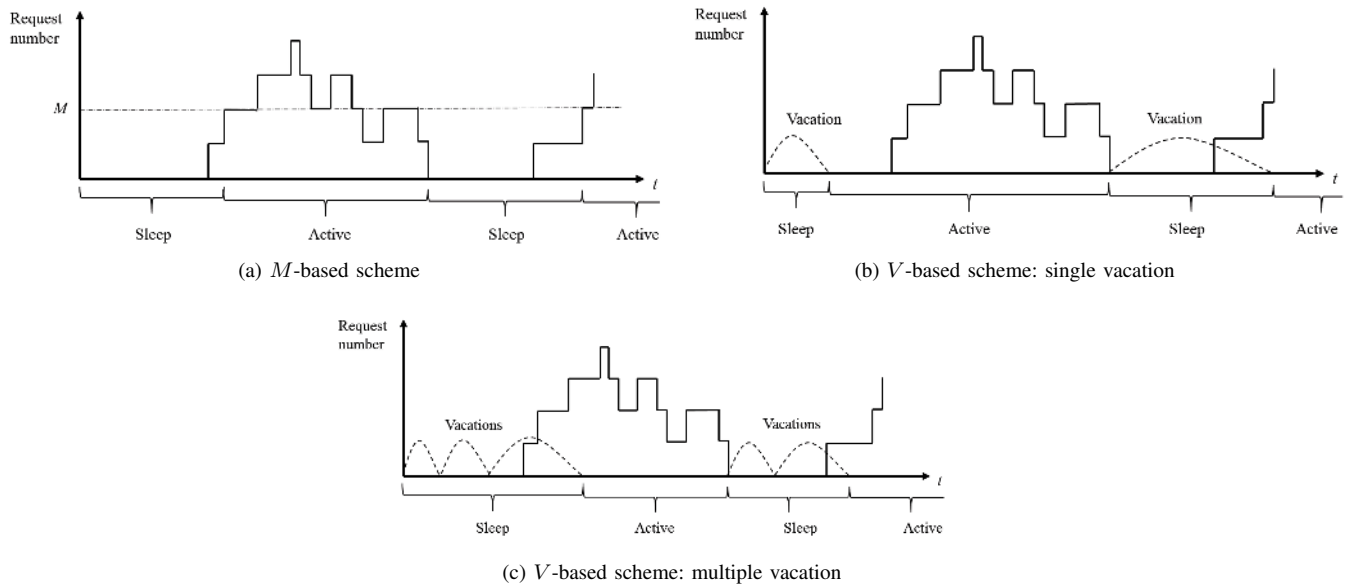
(a) M-based scheme

(b) V-based scheme: single vacation

(c) V-based scheme: multiple vacation

Fig. 2. BS wake-up schemes [66]. In the M-based scheme, the BS is switched off in an idle condition and it wakes up when M users arrive at the system. On the other hand, for the V-based scheme, the BS remains sleep for a period of vacation time before waking up. In a single vacation case, the BS remains awake after the vacation period even if there is no call request to serve, while in a multiple vacation case, the BS goes back to sleep if it wakes up and finds no call request to serve.

in addition to specifying which BSs to be switched off, another equally important research direction aims to determine the wake up instants for switched off BSs. For instance, two wake up schemes are presented in [66], namely number M-based and vacation time V-based schemes, as shown in Figure 2. One limitation with the M-based scheme is the requirement that the BS needs to continuously monitor the user request arrivals, which translates into an advantage for the V-based scheme. For femto-cell BSs with overlapped coverage with macro-cell BSs, three wake up modes are presented in literature, namely BS controlled, MT controlled, and network controlled modes [67]. In the BS controlled mode, the femto-cell BS performs continuous sensing for user activity for wake up, while in the MT controlled mode, the MT sends wake up messages for a sleeping femto BS. Finally, in the network controlled mode, the core network controls the femto BS operation through wake up messages over the backhaul link. The three different modes of operation yield different performance in terms of BS and MT energy consumption and signaling overhead. Specifically, the BS controlled mode results in less energy saving for the BS, the MT controlled mode increases energy consumption for the MT, and the network controlled mode results in additional signalling overhead [67]. MDP-based optimal wake up schemes are proposed in [68] for network operated femto BSs overlapping with a macro BS. To wake up the right femto BSs, which serve the extra traffic demand and still result in efficient energy usage, information regarding call traffic load and user localization within the macro cell is required. In absence of the traffic localization information, the femto BS wake up problem can be formulated as a partially observable MDP [68]. The last issue dealt with in the BS operation problem deals with switching off mode entrance and exit, which are two important design stages in implementing the BS operation decision [69]. The switching off mode entrance stage should specify how the transition from the on state to the off state is implemented. If a BS is switched off very fast, the corresponding MTs may not be able to execute successfully their handover procedures and their calls will be dropped. One reason is a strong received signal from the BS that the MT is associated with, which prevents the MT from hearing signals from nearby BSs. Hence, if the BS that an MT is connected to is suddenly switched off, the MT will not be able to synchronize and associate with another active BS. Another reason is the maximum number of handovers that can occur simultaneously towards a new BS, due to the limited signaling channel capacity. As a result, a progressive switching off operation can be used, an operation that is referred to as BS wilting [69], as shown in Figure 3a. During this process, the MTs associated with the wilting BS initiate a handover process to the neighboring BSs and the BS switching off operation is suspended if the handover process of MTs is unsuccessful. On the other hand, the switching off mode exit specifies how the transition from the off state to the on state is implemented. A BS that is switched on too fast can generate a strong interference to MTs in service. Hence, a progressive switch on process can be used, and such an approach is referred to as BS blossoming [69], as shown in Figure 3b.

B. MT Radio Interface On-Off Switching

Similar to BS on-off switching, an MT with a low or bursty traffic load can switch off from time to time its radio interface to save energy. Designing an appropriate on-off switching schedule for the MT radio interface varies according to whether the MT establishes communications on the downlink [48], [70] - [74], uplink [58], [75], or both links [76].
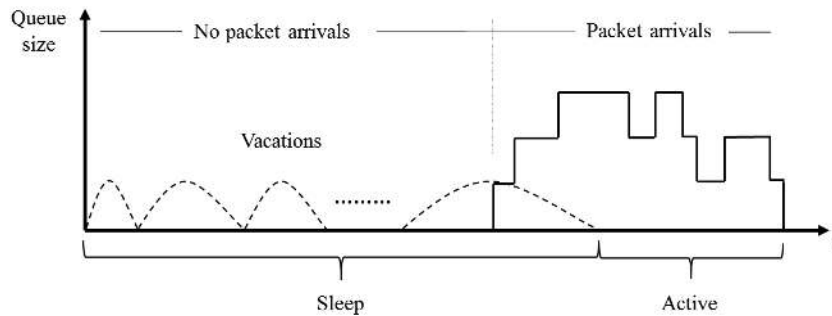
Fig. 4. Modeling of MT on-off switching as a server with repeated vacations [48]. The model is similar to the BS V-based scheme with multiple vacations.
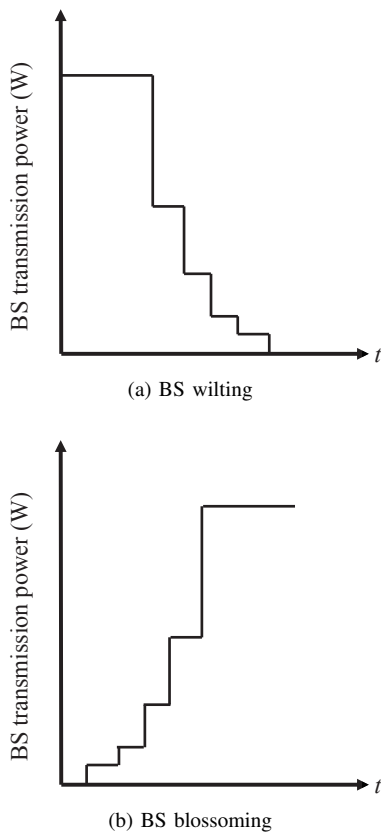

(a) BS wilting

(b) BS blossoming

Fig. 3. BS switching off mode entrance and exit [69]. In BS wilting, the BS transmit power is progressively halved until the BS is off. In BS blossoming, the BS transmit power is progressively doubled until the BS is on.

For downlink communications, the MT radio interface on-off switching mechanism can be classified into two categories based on whether it implements a traffic shaping technique or not [48], [70] - [74]. In the absence of traffic shaping techniques, an MT (with a low or bursty traffic load) switches off its radio interface if no data packets are available for the MT at the serving BS. Hence, the MT establishes a switching on-off schedule that specifies the switching off intervals and switching on instants. At a switching on instant, the MT listens to its serving BS to check if there are any packets available for it. If no packets are available, the MT assumes a switching off interval; otherwise, the MT keeps its radio interface active to receive the available packets. During the MT switch off interval, the incoming packets are buffered at the BS until the MT next switch on instant. While a long switch off interval can enhance the energy savings for the MT, it increases the buffering delay of the packet at the BS until it is received by the MT. In addition, incoming data packets for the MT may be discarded in case of a buffer overflow at the BS. Furthermore, unnecessarily switching on the MT to check packet availability at the BS buffer results in MT energy losses. Hence, the main research challenge in this case is how to design a switching schedule for the MT radio interface that maximizes its energy saving while reducing the buffering delay of packets available at the BS. One approach is to model the MT radio interface as a server that assumes repeated vacations [48], [70], as shown in Figure 4. Hence, analytical expressions can be derived for the expected number of switching off intervals until a packet is available for the MT at the BS. Using the analytical expressions, myopic optimization problems can be formulated to minimize the MT energy consumption rate while achieving acceptable performance in terms of the message response time, which is defined as the time interval from arrival time of an arbitrary message at the BS to the time it leaves the system (BS) after service completion [48]. Besides myopic optimization techniques, dynamic programming can be used to design a switching off schedule that minimizes a cost function consisting of a weighted sum of the energy consumed for radio interface on-off switching and a target performance metric (e.g., the buffering delay at the BS for the MT when switched off) [70]. In addition to queueing models coupled with myopic and dynamic optimization techniques, a Llyod-max algorithm can be used to design a schedule that specifies the switching on instants for the MT radio interface [71]. One limitation with the aforementioned works is that, if the packet inter-arrival time of the application is too small, the MT cannot switch off its radio interface to provide acceptable QoS performance. In addition, the MT consumes a significant amount of energy to switch on its radio interface. Further, every time the MT finds a single packet available at the BS buffer, an interruption signal is triggered by the MT radio interface to activate the MT data bus and central processing unit (CPU). If the MT experiences frequent interrupts, it will not be able to enter a deep sleep state and only a small amount of energy will be saved. Therefore, traffic shaping techniques are introduced to enable a longer

idle duration for an MT[2]. Such a traffic shaping technique can be implemented by the MT itself, where the MT buffers the incoming data packets for a short period at its radio interface, without activating its data bus and CPU, and then releases the data packets as a burst so as to reduce the interruption trigger events and to save more energy [72]. For transmission control protocol (TCP) applications, an alternative approach can be triggered by the MT, where the MT forces the BS to send data packets in bursts and can enjoy a longer idle duration by announcing a zero congestion window size. Hence, the data packets are buffered at the BS for a longer period, until the MT announces an appropriate window size to allow the BS to release data packets in bursts [73]. While the above traffic shaping research deals with a single-user scenario, the main objective in a multi-user environment is to schedule the on-off switching of radio interfaces for different MTs so as to satisfy their target QoS and save energy by switching off the MT radio interfaces for a long enough time [74]. An MT stores sufficient data at its buffer to satisfy its QoS and switches off its radio interface to save energy while the BS serves another MT. The MT switches on its radio interface only when no sufficient data is available at its buffer to satisfy the QoS requirement.

For uplink communications, besides adapting the physical layer parameters such as controlling the transmission power and modulation/coding schemes, an MT can switch on and off its radio interface to further save energy. In [75], it is shown that different parameters such as the packet arrival rate and packet delay constraint affect the practicality of adopting such a switching approach. Specifically, it is practical to employ an on-off switching mechanism for energy saving at MTs for small packet arrival rate and/or large packet delay constraint. In such scenarios, the research challenge is how to jointly adapt the power control, modulation and coding schemes, and switching on and off the MT radio interface to save energy in presence of stochastic traffic and channel conditions (i.e., no a-priori knowledge of traffic arrivals and channel conditions). In this case, an MT can choose to switch off its radio interface and hold data packets in its buffer to transmit them as a burst in better channel conditions. Besides saving energy, the transmission mechanism should satisfy the target QoS in terms of data packet delay and should avoid an overflow event at the MT buffer. A Markov decision process (MDP) problem can be formulated to control the data packet transmission throughput (and hence the amount of buffered data packets), the achieved bit error probability, and the MT radio interface state (switch on or off) so as to balance energy saving with QoS guarantee (i.e., minimizing data packet delay and avoiding buffer overflow) [58].

A general model for MT radio interface on-off switching is captured in the context of bi-directional communications [76]. In such a scenario, no BS buffering delay is experienced by incoming downlink traffic during uplink transmission, as the MT radio interface is already switched on. Hence, a finite general Markovian background process can be used to model the uplink activity and downlink traffic so as to

derive analytical expressions for the buffer occupancy and downlink packet delay statistics [76]. Such expressions can be useful in developing an efficient on-off switching mechanism for the MT radio interface for both uplink and downlink communications.

C. Discussion

Based on the above review, BS on-off switching aims to exploit spatial and temporal fluctuations in call traffic load to achieve energy saving. As a result, using static call traffic models for switching schedule design (i.e., to determine decisions on switch-off and wake-up instants) and/or performance evaluation, as in [64], is not realistic. Instead, the call traffic models should reflect a joint spatial and long-term temporal fluctuation behavior, as in [11] and [62]. Traffic models that capture joint spatial and short-term temporal call-level fluctuations, such as [37] and [51], are not capable of assessing the daily switching schedule performance due to a time varying traffic demand. Furthermore, traffic models that capture only long-term (as in [62] and [65]) or short-term (as in [55]) temporal call-level fluctuations fail to exploit the spatial dimension of the problem and stand unrealistic for performance evaluation in large-scale networks with multiple BS sites. For BS power consumption models, both static and dynamic components, as in (7) and (9), should be accounted for, which is the case for the algorithms developed in [11], [37], and [64]. Power consumption models under the assumption of constant transmission power, as in [51], [55], [63], and [65], neglect the scaling of transmission power with the call traffic load, which is unrealistic. Overall, the reported solutions in Section IV.A. aim to minimize the network energy consumption, which is somehow similar in concept to maximizing the energy consumption gain given in (14). However, such an expression does not assess the network gain (in terms of transmitted power as in (15) or network coverage as in (16)) versus the incurred cost (in terms of the network consumed power). The reported solutions minimize the network energy consumption while satisfying a target performance metric. For BS on-off switching solutions, the target performance metrics are based on admission quality requirements, as in [11], [51], [55], [64]. Few works account for application quality requirements, as in [37]. In practice, a solution should satisfy both admission and application quality requirements, as in [62] and [63], to better serve the users required QoS.

On the other hand, MTs can save energy by switching off their radio interfaces during idle periods of bursty traffic. Hence, static traffic models for a fixed number of backlogged data packets ready for transmission, as in [75], are not realistic to determine the MT idle period and hence will not help in developing practical sleep schedules for the MTs. Instead, practical traffic models should capture the packet-level short-term temporal fluctuations, as in [48], [58], [70], [71], [73], [74], and [76]. While some solutions account for both active and idle power consumption values, as in [48], [70], [71], and [73], and reception power consumption, as in [74] and [76], these solutions do not include the circuit power consumption component of the MTs. Both transmission

---

[2]The idle duration in this context represents the interval during which an MT is not receiving any data packets.

and circuit power consumptions should be accounted for as in [58] and [75]. However, such models assume fixed circuit power consumption and neglect the dynamic circuit power component as given in (12) and (13). The reported solutions in Section IV.B. minimize the MT energy consumption while satisfying application quality performance metrics. However, such a modeling approach overlooks the network capacity limitations, e.g., in terms of available bandwidth, which may lead to call blocking. Hence, the proposed solutions should aim to satisfy both application and admission quality requirements.

V. Green Solutions and Analytical Models at High and/or Continuous Call Traffic load

Energy efficient scheduling techniques are adopted to satisfy the required QoS at reduced energy consumption when switching on-off techniques are infeasible due to a high and/or continuous call traffic load. Various scheduling techniques are proposed in literature for network operators and mobile users, and can be divided into four categories, as shown in Table VIII. The categories include scheduling for single-network access, multi-homing access, small size cells, and scheduling with different energy supplies, which are discussed next in more details.

A. Scheduling for Single-network Access

In this case, a mobile user connects to a single wireless access network at a time. Two system models are adopted in literature for single-network access. The first model assumes that a single network covers a given geographical region, which can be referred to as a homogenous wireless medium. The second model deals with the availability of multiple networks with overlapped coverage in the geographical region, which is referred to as a heterogeneous wireless medium. For the homogeneous wireless medium, the network operator aims to assign its resources to MTs so as to reduce the total power consumption of its BSs. Such an objective can be achieved by minimizing the transmission power while providing acceptable QoS performance, a technique that is referred to as margin adaptive strategy [9]. An approach to implement the margin adaptive strategy is via a score-based scheduler. For instance, in an OFDMA system, the BS calculates a score for every resource block q to be assigned to MT m [77]. The calculated score ensures that the BS would consume the least transmission power by assigning resource block q to MT m. Moreover, the score promotes fair resource allocation among MTs, as a penalty function can be included based on the number of already allocated resource blocks for MT m. A low score indicates a more desirable resource block. Fairness consideration is also investigated in [78] following a proportional rate constraint, which ensures that each user eventually obtains a specific proportion of the system throughput. Admission control policies can also be employed to implement a margin adaptive strategy, where a new session is admitted into the system as long as the sub-frame energy in an OFDMA-based BS is kept below a certain threshold [79]. Moreover, a margin adaptive strategy can be implemented through a discrete rate adaptation policy

that controls the transmission rate and power according to channel conditions, so as to maximize the achieved energy efficiency while satisfying a bit-error-rate constraint [80]. Similarly, a channel driven rate and power adaptation strategy can be achieved by jointly adapting modulation and coding schemes (MCS) and transmission power to optimize the trade-off between goodput and energy efficiency [81]. In addition, a margin adaptive strategy can be implemented by scheduling resources among MTs based on their traffic delay tolerance [54]. Specifically, delay tolerant traffic (e.g., video and data) can be served in an opportunistic way during periods of good channel conditions (i.e., soft real time service). One limitation with the margin adaptive strategy is the requirement of CSI to allocate the transmitted power, which requires using pilot symbols. These pilot symbols will incur some energy consumption. Two approaches can be used for pilot energy assignment [9], namely constant single pilot energy and constant total pilot energy. In the former approach, each pilot keeps the same energy level independent of the number of pilot symbols. Hence, the larger the number of pilot symbols is, the more accurate CSI is available, yet the higher energy consumption is. The later approach assigns a constant energy to all pilots, resulting in reduced energy per pilot for a larger number of pilots, which can lead to inaccurate CSI. On the other hand, in a heterogeneous wireless medium, energy can be saved by assigning MTs to the BSs that reduce energy consumption for a set of operators with BSs of overlapped coverage [49]. In addition, each BS in such a heterogeneous environment may choose between two modes of operation, i.e., point-to-point or point-to-multi-point. Hence, the problem can be decomposed into two sub-problems, one for BS selection and the other for BS mode choice. While the work in [49] controls transmission power only through BS operation mode selection, a joint BS selection and power control mechanism is proposed in [82], which aims to associate MTs to BSs with overlapped coverage while minimizing the BS transmission powers to reduce the interference among different links. Furthermore, offloading techniques can be adopted to enhance energy efficiency in a heterogeneous wireless medium. Specifically, through mobility prediction and using the pre-fetching feature, data traffic can be offloaded from cellular networks to WiFi hotspots or femto-cells [83]. Hence, delay tolerant traffic can be downloaded when mobile users are close to the WiFi access point or femto-cell instead of using the macro-cell [84]. Overall, offloading can be either network or user driven [85]. Various factors affect the energy efficiency performance in terms of user mobility, backhaul throughput, data size, and WiFi and/or femto-cell densities [83].

Similarly, MTs can save energy by appropriate resource scheduling on the uplink, based on the network multiple access scheme. Various energy efficient mechanisms are proposed for OFDMA-based networks [33], [34], [43], [57]. The mechanisms mainly enhance energy efficiency through subcarrier allocation, power control, and joint subcarrier allocation and power control [43]. Both centralized and decentralized architectures can be adopted to implement the mechanisms [33], [34]. In a centralized architecture, the BS in each cell jointly performs subcarrier allocation, modulation

TABLE VIII
A SUMMARY OF GREEN SOLUTIONS AND ANALYTICAL MODELS AT HIGH AND/OR CONTINUOUS CALL TRAFFIC LOAD.

Solution/Analytical Model			Comments	References
Single-network	BS	Margin adaptive strategy	It minimizes the transmission power while providing an acceptable QoS.	[9], [54], [77] - [81]
		User association in Heterogeneous wireless medium	It assigns the MTs to the BSs, with coverage overlap, that reduce energy consumption for a set of operators.	[49], [82] – [85]
	MT	OFDMA network — Sub-carrier allocation	It enhances energy efficiency through joint sub-carrier allocation and power control.	[33], [34], [43]
		OFDMA network — Carrier aggregation	It employs both PCC and SCC carrier components for energy saving.	[57]
		TDMA network	The MT energy efficiency is maximized in TDMA through opportunistic transmission.	[59]
Multi-homing	BS	Network Cooperation	The MT receives required data rate from multiple BSs simultaneously. The BSs coordinate their transmitted power for energy saving.	[50], [53]
	MT	BS selection and power allocation	The MT specifies a set of BSs for uplink transmission and determines the allocated transmission power for each radio interface.	[20], [46]
Small Cells			It divides the cell into several tiers of smaller cells to reduce transmission range for BSs and MTs.	[36], [56], [86], [87]
Multiple Energy Sources	BS	Multiple retailers	The network operator decides how much electricity to procure from each retailer.	[23], [88]
		On-grid and green energy sources	The objective is to maximize the utilization of green energy and saves the on-grid energy.	[52]
		Complementary renewable sources	The BSs are powered using only renewable sources.	[89], [90], [92] – [97]
	MT	Multiple batteries	It employs the recovery effect of batteries.	[98]

order adaptation, and power control for MTs. In a distributed mechanism, given a subcarrier assignment, an MT adjusts its modulation order and transmission power to optimize its own energy efficiency. In a multi-cell environment, multi-cell interference should be taken into account via energy efficient uplink resource allocation scheduling [34], [43]. In addition to subcarrier allocation and power control, energy efficiency is maximized for OFDMA-based networks through dynamic carrier aggregation [57]. In general, while an MT served by all carrier components will enjoy an enhanced throughput, its energy consumption also increases. Following a dynamic carrier aggregation technique, an MT is assigned to the queue of a given carrier component, which is referred to as primary carrier component (PCC). Whenever the queue of a carrier component is empty, it helps other carrier components through aggregation and therefore, it is referred to as supplementary carrier component (SCC). Two mechanisms can be adopted for SCC assignment [57]. The first mechanism aggregates all SCCs to support the PCC with the longest queue. The second mechanism orders PCCs according to queue length and SCCs are circularly allocated to the ordered PCCs in a round-robin fashion. For time division multiple access (TDMA)-based networks, energy efficiency is maximized for a set of MTs by opportunistic transmission [59]. Specifically, a scheduler is designed at the BS to select an MT for transmission and to determine its transmission rate. The problem complexity is reduced by decomposing it into two sub-problems. The first is a user scheduling sub-problem which selects an MT opportunistically for transmission, based on channel conditions and backlog information. The second sub-problem determines the transmission rate for the selected MT to minimize the transmission power by transmitting packets in queue such that the average delay constraint is satisfied with equality.

B. Scheduling for Multi-homing Access

Recently, the wireless communication medium has become a heterogeneous environment with overlapped coverage due to different networks. In such a networking environment, MTs are equipped with multiple radio interfaces. Through multi-homing capability, an MT can maintain multiple simultaneous associations with different networks. Besides enhancing the achieved data rate through bandwidth aggregation, multi-homing service can enhance energy efficiency for network operators and mobile users. This is because an MT experiences different channel conditions and bandwidth capabilities over its different radio interfaces.

Different network operators can reduce the transmission power of their BSs by supporting multi-homing services. The motivation behind employing multi-homing to enhance energy efficiency can be explained using the power-rate curve, which can be divided into two regions [50]. In the first region, power consumption increases slowly with the growth of data rate, while in the second region power consumption increases dramatically with data rate. Hence, a multi-homing threshold, R_b, of data rate can be determined to start multi-homing transmission if the required data rate is larger than R_b [50]. The multi-homing threshold is based on the ratio of channel gain

between the MT and BSs of different networks. In addition, the optimal transmission data rate from each BS can be specified to maximize the energy efficiency of the networks. Moreover, cooperating BSs can control their transmission power using a semi-Markov decision process (SMDP) to minimize the total BS power consumption under a target QoS constraint at the MTs [53].

Similarly, MTs can enhance their energy efficiency through multi-homing service. In this case, an MT determines which and how many BSs will be selected for multi-homing, based on the required data rate and the channel parameters of available BSs [20]. To reduce the complexity, the problem can be decomposed into two sub-problems. The first sub-problem specifies which BSs will be selected for multi-homing and the second sub-problem determines the optimal transmission rate from each selected BS. For a constant data rate service, energy efficiency maximization is equivalent to MT total power consumption minimization. Different from [20], the work in [46] deals with energy efficiency maximization for a variable data rate using multi-homing service through power allocation.

C. Scheduling With Small Size Cells

A small cell has a radio coverage of tens to a few hundreds of meters (e.g., pico and femto cells) [36]. As a result, the division of a macro-cell into several tiers of smaller cells replaces a long range transmission with a short range transmission due to the close proximity between small cell BSs and MTs [86]. It is expected that the power consumption of a small cell will be approximately 5 watts by 2020 [87]. Hence, an improved energy efficiency can be achieved. In [86], an expression of the possible power gain $G(J)$ resulting from cell splitting into J smaller cells is provided. It is shown that, for an ideal free space propagation channel model, the achieved gain satisfies $G(J) < 1$, and hence cell splitting should not be implemented. On the other hand, in a non-ideal propagation environment, $G(J) > 1$ and it increases with the number J of small cells[3]. Different configurations are presented in literature for small cell deployment, as shown in Figure 5. It is shown in [87] that the cell-on-edge deployment results in a significant reduction in network energy consumption, as compared to the uniformly distributed configuration, due to lower transmit power for cell edge users.

The main challenge of cell splitting is the associated inter-tier interference. This is mainly due to the limited radio resources. Hence, the radio resources of the macro-BS are shared among the small cells. Multi-cell processing can be employed to mitigate interference [86]. Hence, multiple BSs within a cluster exchange CSI and users' data to support MTs and eliminate interference. Based on the gathered information, beam-forming techniques are used to minimize the total transmit power while satisfying a certain signal-to-interference plus

---

[3]The BS power consumption model in [86] does not capture the BS embodied energy as in [16]. When BS embodied energy is considered, there is a limit on the number of small cells that can be included to enhance energy efficiency.

(a) Cell-on-edge deployment
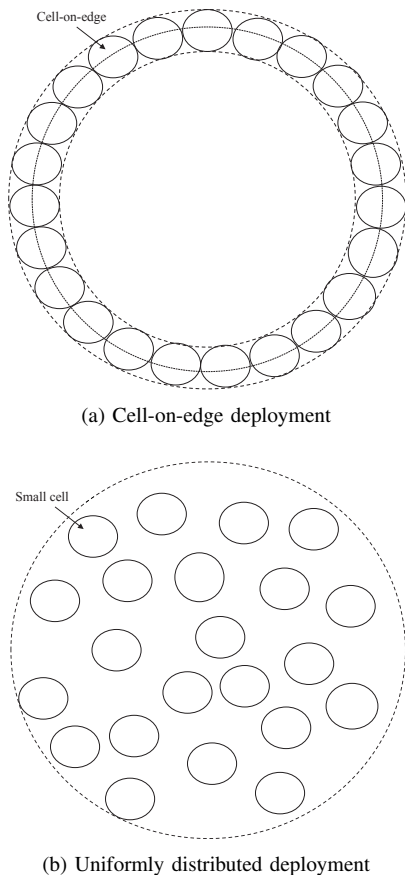

(b) Uniformly distributed deployment

Fig. 5. Configurations for small cell deployment [87]. Cell-on-edge deployment distributes the small cells around the edge of a macro cell. On the other hand, in uniform distribution, small cells are uniformly distributed across the macro cell.

noise ratio (SINR) for different MTs. In addition to multi-cell processing (and in presence of both co-tier and cross-tier interference), admission control with QoS guarantee can play a vital role in mitigating interference, where a joint resource allocation mechanism can be employed among multi-tier networks [56] .

D. Scheduling With Multiple Energy Sources

Various scheduling techniques have been proposed to deal with the presence of multiple energy sources [23], [52], [88] - [98]. The main objective in these works is to jointly control transmission power and select the energy source that minimizes the total energy consumption. For network operators, multiple energy sources deal with the availability of different electricity retailers [23], [88], on-grid and green (renewable) energy [52], and different (complementary) renewable sources [89] - [97]. For MTs, multiple energy sources deal with the availability of multiple batteries [98].

In an electricity market liberalization model, electricity retailers compete with each other and aim to achieve the highest individual profits by adjusting the electricity price offered to users in different regions [23]. Electricity prices offered by different retailers change frequently to reflect variations in the cost of energy supply, which is referred to as real time pricing. Given a set of electricity retailers, a Stackelberg game can be

formulated, where each retailer provides its real time price, to maximize its profit, to the network operator which decides how much electricity to procure from each retailer to power on its BSs and achieve the lowest call blocking with the least cost [23]. In [88], the optimal amount of energy to be procured from each retailer is determined using evolutionary algorithms (Genetic Algorithm and Particle Swarm Optimization), which due to the random nature of the evolution process is shown to outperform the deterministic algorithm developed in [23]. In addition to the presence of multiple electricity retailers, it is argued that the BSs of future cellular networks will be powered by both on-grid and green (renewable) energy (e.g., solar energy) [52]. With such a hybrid energy system, the objective is to optimize energy utilization in such networks by maximizing the utilization of green energy and saving on-grid energy. Network designers are faced with two central issues [52]: 1) how to optimize the green energy usage at different time slots to accommodate the temporal dynamics of the green (solar) energy generation and the mobile traffic, and 2) how to accommodate the spatial dynamics of the mobile traffic with the objective of maximizing the utilization of green energy by balancing the green energy consumption among BSs through cell size adjustment. While the aforementioned works deal with the presence of on-grid energy, the long term objective is to power BSs in appropriate locations using only a combination of complementary renewable sources (e.g., wind in winter and solar in summer) [89]. Furthermore, power cooperation enables different BSs (networks) to share (trade) their green power with each other whenever possible for a sustainable and energy efficient network operation [90]. In utilizing renewable energy sources, renewable energy generation and storage should be investigated. Since renewable energy sources are intermittent, energy storage is used to address this limitation. Hence, the harvested energy is stored in a battery with finite capacity before it is used for transmission [91], [92]. In this context, the energy replenishment process and the storage constraints of the rechargeable batteries need to be taken into account while designing efficient transmission strategies [93]. Two constraints should be accounted for at the energy harvesting battery [94]. The first ensures that the energy drawn from the battery is at most equal to the energy stored at the battery, which is referred to as the causality constraint. The second constraint ensures that the energy level at the battery does not exceed a maximum level to avoid battery energy overflow. Hence, storage sizing is very important to guarantee a sustainable energy at a reduced cost. In addition, BSs have to adapt their data transmission to the availability of energy at a particular instant [95], [96]. Therefore, more studies are needed to minimize the overall power consumption of BSs, through on-off switching at a low call traffic load or scheduling and node cooperation [97] at a high call traffic load, to reduce the required generation potential and storage capacity. A very important aspect of green communications is to consider the environmental dimension of the proposed solution. Hence, while selecting an appropriate energy source (i.e., electricity retailer and/or renewable energy source), it is necessary to guarantee that the CO_2 emission cost is below a target level. The CO_2 emission cost, in kg/hr, related to the
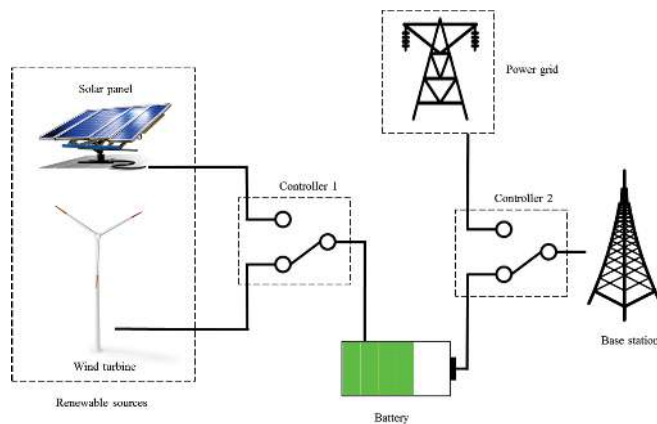
Fig. 6. A green hybrid solution. The hybrid solution uses a combination of renewable and grid energy sources. Complementary renewable energy sources can be used. If the power grid is absent, i.e., the BS is not connected to the power grid and hence controller 2 does not exist, the BS is powered only through renewable sources. An energy harvesting battery is used to overcome the intermittent nature of renewable energy sources.

BS power consumption P_b can be expressed as [23]

$$I(P_b) = \alpha P_b^2 + \beta P_b, \tag{27}$$

where α and β are constants that depend on the pollutant level of the electricity retailer.

For MTs, under a pulsed discharge profile, the battery is able to recover some charges during the interruptions of the drained current (i.e., no transmission periods). Hence, an improved battery performance can be achieved. This phenomena is referred to as the recovery effect. To promote the recovery effect and enhance the battery performance, a package of multiple batteries can be used and a scheduling policy can be developed to efficiently distribute the discharge demand among the multiple batteries connected in parallel [98].

E. Discussion

The majority of research works that investigate green communication solutions at a high traffic load employ static traffic models for resource scheduling and performance evaluation, as in [9], [23], [33], [34], [43], [49], [50], [77], [78], [82], [86], and [88]. Very few works use traffic models that reflect long-term (as in [79] and [87]) or short-term (as in [54], [56], and [57] for call-level and [59] for packet-level) temporal fluctuations. Also, few works use traffic models that capture spatial fluctuations in traffic load, as in [52] and [53]. Spatial and temporal traffic models should be employed for performance evaluation of green resource scheduling solutions. Spatial traffic models are useful in evaluating the algorithm performance in large-scale networks, while temporal models are important to investigate the associated signaling overhead, which may jeopardize the energy saving benefits, if high overhead is expected. In addition, many works account only for transmission power consumption as in [9], [43], [49], [50], [54], [56], [57], [79], [82], [86], and [87]. Both transmission and circuit power consumption should be accounted for, as in [23], [33], [34], [52], [53], [59], [77], [78], and [88]. However, the aforementioned models do not account for dynamic circuit power consumption, as in (12) and (13). Also, BS transmission power consumption should scale with the traffic load as

expressed in (7) and (9). Furthermore, for small-cell and multi-tier deployment, both operation and embodied energy should be accounted for as in (10). Accounting only for operation power consumption in such scenarios can be misleading. While some works aim to minimize the energy consumption, the work in [77] is to maximize an energy consumption gain expression similar to (14). Moreover, the works in [33], [34], [43], [50], [57], and [78] aim to maximize an energy efficiency expression similar to (17), (18), or (19). Such an expression provides a better indication of the performance in terms of the achieved gain (in terms of resulting data rate) versus the incurred cost (in terms of the energy consumed). Almost all reported solutions aim to minimize energy consumption or maximize energy efficiency, while maintaining a satisfactory performance in terms of application quality requirements. The works in [23] and [88] target admission quality requirements. In practice, an effective solution should satisfy both admission and application quality requirements, as in [56].

VI. Future Research

The existing research works mainly focus on enhancing energy efficiency either of network operators or mobile users. However, a green solution implemented at the network operator side can lead to high energy consumption at the mobile user side, and vice-versa. Hence, green solutions should capture the trade-off in energy efficiency among network operators and mobile users and should be jointly designed to balance such a trade-off.

For instance, the BS on-off switching mechanism involves two phases, namely user association and BS operation. Focusing only on energy efficiency of the network operator, a BS on-off mechanism can lead to an energy inefficient user association from the mobile user perspective. Specifically, it can lead to MTs being associated in the uplink with a far away BS in order to switch off a nearby BS. This will result in energy depletion for the MTs and hence dropped services. Thus, a BS on-off switching mechanism should capture the trade-off in the achieved energy efficiency for the network operator and mobile users, and should aim at

balancing them. MTs should be associated with BSs that can balance energy saving for both network operators and mobile users. The existing research, however, focuses on balancing energy consumption performance of a BS with the flow-level performance at the MT, e.g., [37]. Instead, the multi-objective function in [37] should aim to balance energy saving for BSs and MTs while satisfying the MT required QoS. As a result, BS switching off decision criteria in literature, such as user-BS distance [62], call traffic load [51], network impact [63], and network coverage holes [64], should be revised. Specifically, the switching off criterion should include, besides the aforementioned metrics, an MT energy consumption metric. Similarly, the existing mechanisms present only the call traffic load as a wake up criterion [68]. The switching mechanisms should capture the degradation in energy consumption for MTs and include it as a BS wake up decision metric. Furthermore, MTs suffer from inter-cell interference. An uplink scheduling scheme at MTs performs power allocation while dealing with the inter-cell interference negative effect. However, inter-cell interference can be affected by the BS on-off switching decision. Such a dependence can be modeled in the user received SINR using a BS activity parameter, which equals to one if the BS is on, and zero otherwise. In addition, the BS on-off switching decision should promote energy saving at MTs by switching off cells that result in the highest interference during a low call traffic load condition. Moreover, the analytical models used in literature, e.g., the queuing model in [55], mainly assess the network energy saving performance for a given mechanism. Such models should be extended to assess the energy saving performance for both network operators and mobile users.

Similarly, the existing MT radio interface on-off switching mechanisms focus mainly on the energy saving performance at the MT without capturing the impact of the energy saving mechanisms implemented at the BSs. Specifically, the downlink mechanisms allow an MT to switch off its radio interface for a given interval while dealing with only the buffer delay and/or overflow at the BS, e.g., [48], [70] - [74]. However, the impact of BS on-off switching is not considered. If the serving BS is switched off during the MT sleep interval, the MT connection will be dropped and the buffered data will be lost. Hence, the MT radio interface switching schedule design needs to be revised. For instance, in [48], the MT switching on is triggered upon a packet arrival at the BS. Such a model should be extended to account for the BS switching off decision as an additional switching on trigger for the MT radio interface. Moreover, the existing switching off design metrics focus on balancing energy consumption at the MT with the buffer delay at the BS [70]. An extension is required to account for the BS energy consumption due to a delayed switching off decision for the BS while waiting for the MT to wake up. Furthermore, network operators can save energy at BSs by scheduling delay tolerant applications (e.g., data and video) opportunistically in the presence of good channel conditions. MT radio interface on-off scheduling should take account of the delay at the BS due to both MT inactivity and BS opportunistic scheduling of traffic. The radio interface on-off scheduling at an MT and the opportunistic

traffic scheduling at the BS should balance energy efficiency for both network operators and mobile users, while satisfying the target performance metrics. Opportunistic scheduling can also be used for energy saving at MTs. However, such an approach does not always work in practical scenarios (e.g., a stationary user suffering from a slow fading channel), which is not the case for BSs due to spatial user diversity. For the MT energy saving mechanisms at the uplink, power control and radio interface on-off switching mechanisms account in their design only for the channel and traffic dynamics [58]. In addition, the BS on-off switching dynamics should be captured while designing an energy saving mechanism.

Furthermore, the energy efficient resource scheduling mechanisms at a heterogeneous wireless medium assign MTs to the BSs which reduce energy consumption for network operators [49], [82]. Such mechanisms mainly deal with downlink resource scheduling. However, no investigation is performed for MTs with bidirectional traffic, e.g., for video call applications. In this case, two approaches can be implemented for energy saving at both network operators and mobile users. The first relies on single-network access, where the MT is associated with the BS that balances energy saving for the network operators and the mobile users. On the other hand, the second approach employs multi-homing where the MT connects on the uplink to the BS that promotes energy saving for the mobile user while the MT connects on the downlink to the BS that promotes energy saving for the network operators. Moreover, the potentials of the heterogeneous wireless medium should be better exploited to enhance energy saving. In addition, for multi-homing service, as MTs connect to multiple networks simultaneously, radio resources at different radio interfaces can be properly scheduled to enhance energy efficiency. However, existing research works focus only on power allocation schemes at the MT different radio interfaces to save energy in different channel conditions. Given the bandwidth capabilities of different networks, cross-layer designs that incorporate joint bandwidth and power allocation can lead to an improved energy efficiency.

In addition, the existing opportunistic scheduling mechanisms focus on energy saving for network operators [54] or MTs [59]. However, for MTs with bidirectional traffic, opportunistic scheduling should be implemented such that the time slot for uplink and downlink transmission can balance energy saving for both network operators and mobile users. Finally, for radio resource scheduling in BSs powered by renewable energy sources, the existing research focuses mainly on downlink delay tolerant applications [52], [88]. Hence, BSs aim to schedule data transmissions at time slots when energy is available. However, when MT radio interface on-off scheduling is implemented, the BSs need to account for the MT sleep interval, which may conflict with the BS energy limitation due to the finite size of the energy harvesting buffer at the BS and might lead to buffer overflow. Hence, the resource scheduling mechanism should balance energy availability at the BS with energy saving at the MT.

VII. Conclusion

Due to environmental, financial, and QoE considerations, there has been a great emphasis on the need to develop energy efficient solutions for wireless communications and networking. Such solutions are referred to as green solutions. In this survey, we have reviewed the existing research activities dedicated to green wireless communications and networking solutions, from the network operator and mobile user perspectives. The first step towards developing an effective green solution is to identify the call traffic load condition, based on which an appropriate definition of energy efficiency can be proposed, as discussed in Section II. The second step is to use proper models for power consumption and for call traffic loads that account for both spatial and temporal fluctuations, as discussed in Sections II and III, respectively. Besides improving energy efficiency, certain performance metrics should be satisfied, as discussed in Section III, which are determined based on the target application, e.g., voice, data, or video services. Given the call traffic load condition, different green solutions and analytical models can be adopted, as presented in Sections IV and V. At a low call traffic load condition, on-off switching of radio devices (e.g., BSs for network operators and MT radio interfaces for mobile users) can improve the performance of energy consumption. Radio resource scheduling techniques have been proposed for a high call traffic load condition. Despite the various efforts proposed to analyze and design effective green solutions, many open issues remain to be further investigated. As future research, green solutions should capture the tradeoff in energy efficiency among network operators and mobile users and should be designed to balance such a trade-off.

VIII. Acknowledgment

References

[1] M. Ismail and W. Zhuang, "Cooperative networking in a heterogeneous wireless medium," *Springer Briefs in Computer Science*, New York, April 2013.

[2] T. Chen, Y. Yang, H. Zhang, H. Kim, and K. Horneman, "Network energy saving technologies for green wireless access networks," *IEEE Wireless Commun.*, vol. 8, no. 5, pp. 30-38, Oct. 2011.

[3] D. Cavalcanti, D. P, Agrawal, C. Cordeiro, B. Xie, and A. Kumar "Issues in integrating cellular networks, WLANs, and MANETs: a futuristic heterogeneous wireless network," *IEEE Wireless Commun. Magazine*, vol. 12, no. 3, pp. 30-41, 2005.

[4] M. Ismail and W. Zhuang, "A distributed multi-service resource allocation algorithm in heterogeneous wireless access medium," *IEEE J. Select. Areas Commun.*, vol. 30, no. 2, pp. 425-432, Feb. 2012.

[5] M. Ismail, A. Abdrabou, and W. Zhuang, "Cooperative decentralized resource allocation in heterogeneous wireless access medium," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 714-724, Feb. 2013.

[6] M. Ismail and W. Zhuang, "Decentralized radio resource allocation for single-network and multi-homing services in cooperative heterogeneous wireless access medium," *IEEE Trans. Wireless Commun.*, vol. 11, no. 11, pp. 4085-4095, Nov. 2012.

[7] C. Han, T. Harrold, S. Armour, and I. Krikidis, "Green radio: radio techniques to enable energy-efficient wireless networks," *IEEE Commun. Magazine*, vol. 49, no. 6, pp. 46-54, June 2011.

[8] Y. Chen, S. Zhang, S. Xu, and G. Y. Li, "Fundamental trade-offs on green wireless networks," *IEEE Commun. Magazine*, vol. 49, no. 6, pp. 30-37, June 2011.

[9] H. Bogucka and A. Conti, "Degrees of freedom for energy savings in practical adaptive wireless systems," *IEEE Commun. Magazine*, vol. 49, no. 6, pp. 38-45, June 2011.

[10] S. Mclaughlin, P. M. Grant, J. S. Thompson, and H. Haas, "Techniques for improving cellular radio base station energy efficiency," *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 10-17, Oct. 2011.

[11] M. Ismail and W. Zhuang, "Network cooperation for energy saving in green radio communications," *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 76-81, Oct. 2011.

[12] K. Pentikousis, "In search of energy-efficient mobile networking," *IEEE Commun. Magazine*, vol. 48, no. 1, pp. 95-103, Jan. 2010.

[13] G. Miao, "Energy-efficient uplink multi-user MIMO," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2302-2313, May 2013.

[14] E. Oh, B. Krishnamachari, X. Liu, and Z. Niu, "Toward dynamic energy-efficient operation of cellular network infrastructure," *IEEE Commun. Magazine*, vol. 49, no. 6, pp. 56-61, June 2011.

[15] Y. S. Soh, T. Q. S. Quek, M. Kountouris, and H. Shin, "Energy efficient heterogeneous cellular networks," *IEEE Journal Selected Areas Commun.*, vol. 31, no. 5, pp. 840-850, May 2013.

[16] I. Humar, X. Ge, L. Xiang, M. Jo, M. Chen, and J. Zhang, "Rethinking energy efficiency models of cellular networks with embodied energy," *IEEE Network*, vol. 25, no. 2, pp. 40-49, April 2011.

[17] D. Feng, C. Jiang, G. Lim, L. J. Cimini, G. Feng, and G. Y. Li, "A survey of energy-efficient wireless communications," *IEEE Commun. Surveys Tutorials*, vol. 15, no. 1, pp. 167-178, 2013.

[18] S. Tombaz, A. Vastberg, and J. Zander, "Energy- and cost-efficient ultra-high-capacity wireless access," *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 18-24, Oct. 2011.

[19] L. M. Correia, D. Zeller, O. Blume, and D. Ferling, "Challenges and enabling technologies for energy aware mobile radio networks," *IEEE Commun. Magazine*, vol. 48, no. 11, pp. 66-72, Nov. 2010.

[20] X. Ma, M. Sheng, Y. Zhang, and X. Wang, "Concurrent transmission for energy efficiency in heterogeneous wireless networks," *IEEE Trans. Wireless Commun.*, under review.

[21] G. Miao, N. Himayat, Y. Li, and A. Swami, "Cross-Layer optimization for energy-efficient wireless communications: a survey," *Wiley J. Wireless Commun. and Mobile Computing*, vol. 9, pp. 529-542, March 2009.

[22] M. Ismail, W. Zhuang, and S. Elhedhli, "Energy and content aware multi-homing video transmission in heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3600-3610, July 2013.

[23] S. Bu, F. R. Yu, Y. Cai, and X. P. Liu, "When the smart grid meets energy-efficient communications: green wireless cellular networks powered by the smart grid," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 3014-3024, Aug. 2012.

[24] A. P. Bianzino, C. Chaudet, D. Rossi, and J. L. Rougier, "A survey on green networking research," *IEEE Commun. Surveys Tutorials*, vol. 14, no. 1, pp. 3-20, 2012.

[25] Y. Zhang, P. Chowdhury, M. Tornatore, and B. Mukherjee, "Energy efficiency in telecom optical networks," *IEEE Commun. Surveys Tutorials*, vol. 12, no. 4, pp. 441-458, 2010.

[26] R. Bolla, R. Bruschi, F. Davoli, and F. Cucchietti, "Energy efficiency in the future Internet: a survey of existing approaches and trends in energy-aware fixed network infrastructures," *IEEE Commun. Surveys Tutorials*, vol. 13, no. 2, pp. 223-244, 2011.

[27] J. B. Rao and A. O. Fapojuwo, "A survey of energy efficient resource management techniques for multicell cellular networks," *IEEE Commun. Surveys Tutorials*, to appear.

[28] L. Suarez, L. Nuaymi, and J. M. Bonnin, "An overview and classification of research approaches in green wireless networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 142, pp. 1-18, April 2012.

[29] N. Vallina and J. Crowcroft, "Energy management techniques in modern mobile handsets," *IEEE Commun. Surveys Tutorials*, vol. 15, no. 1, pp. 179-198, 2013.

[30] M. Kennedy, A. Ksentini, Y. H. Aoul, and G. M. Muntean, "Adaptive energy optimization in multimedia-centric wireless devices: a survey," *IEEE Commun. Surveys Tutorials*, vol. 15, no. 2, pp. 768-786, 2013.

[31] M. A. Hoque, M. Siekkinen, and J. K. Nurminen, "Energy efficient multimedia streaming to mobile devices - a survey," *IEEE Commun. Surveys Tutorials*, to appear.

[32] G. Miao, N. Himayat, and G. Y. Li, "Energy-efficient link adaptation in frequency-selective channels," *IEEE Journal Selected Areas Commun.*, vol. 58, no. 2, pp. 545-554, Feb. 2010.

[33] G. Miao, N. Himayat, G. Y. Li, an S. Talwar, "Low-complexity energy-efficient scheduling for uplink OFDMA," *IEEE Trans. Commun.*, vol. 60, no. 1, pp. 112-120, Jan. 2012.

[34] C. Y. Ho and C. Y. Huang, "Non-cooperative multi-cell resource allocation and modulation adaptation for maximizing energy efficiency in uplink OFDMA cellular networks," *IEEE Wireless Commun. Letters*, vol. 1, no. 5, pp. 420-423, Oct. 2012.

[35] Y. Rui, Q. T. Zhang, L. Deng, P. Cheng, and M. Li, "Mode selection and power optimization for energy efficiency in uplink virtual MIMO systems," *IEEE Journal Selected Areas Commun.*, vol. 31, no. 5, pp. 926-936, May 2013.

[36] I. Ashraf, F. Boccardi, and L. Ho, "Sleep mode techniques for small cell deployments," *IEEE Commun. Magazine*, vol. 49, no. 8, pp. 72-79, Aug. 2011.

[37] K. Son, H. Kim, Y. Yi, and B. Krishnamachari, "Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks," *IEEE Journal Selected Areas Commun.*, vol. 29, no. 8, pp. 1525-1536, Sep. 2011.

[38] G. Auer, V. Giannini, C. Desset, and I. Godor, "How much energy is needed to run a wireless network?" *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 40-49, Oct. 2011.

[39] M. Deruyck, D. D. Vulder, W. Joseph, and L. Martens, "Modelling the power consumption in femtocell networks" *WCNC'12 Workshop on Future Green Communications*, pp. 30-35, April 2012.

[40] R. Riggio and D. J. Leith, "A measurement-based model of energy consumption in femtocells," *Wireless Days 2012 IFIP*, pp. 1-5, Nov. 2012.

[41] G. P. Perrucci, F. H. P. Fitzek, and J. Widmer, "Survey on energy consumption entities on the smartphone platform" *IEEE VTC'11-Spring*, pp. 1-6, May 2012.

[42] F. Meshkati, H. V. Poor, S. C. Schwartz, "Energy-efficient resource allocation in wireless networks," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 58-86, May 2007.

[43] S. , G. Colavolpe, D. Saturnino, and A. Zapone, "Potential games for energy-efficient power control and subcarrier allocation in uplink multicell OFDMA systems," *IEEE Journal Selected Areas in Signal Processing*, vol. 6, no. 2, pp. 89-103, April 2012.

[44] O. Onireti, F. Heliot, and M. A. Imran, "On the energy efficiency-spectral efficiency trade-off in the uplink of CoMP system," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 556-561, Feb. 2012.

[45] C. Isheden and G. P. Fettweis, "Energy-efficient multi-carrier link adaptation with sum rate-dependent circuit power," *Proc. Globecom'10*, pp. 1-6, Dec. 2010.

[46] O. Galinina, S. Andreev, A. Turlikov, and Y. Koucheryavy, "Optimizing energy efficiency of a multi-radio mobile device in heterogeneous beyond-4G networks," *Performance Evaluation*, under review.

[47] G. Lim and L. J. Cimini, "Energy-efficient cooperative relaying in heterogeneous radio access networks," *IEEE Wireless Commun. Letters*, vol. 1, no. 5, pp. 476-479, May 2012.

[48] A. P. Azad, "Analysis and optimization of sleeping mode in WiMAX via stochastic decomposition techniques," *IEEE Journal Selected Areas Commun.*, vol. 29, no. 8, pp. 1630-1640, Sep. 2011.

[49] K. Ying, H. Yu, and H. Luo, "Inter-RAT energy saving for multicast services," *IEEE Commun. Letters*, vol. 17, no. 5, pp. 900-903, May 2013.

[50] X. Ma, M. Sheng, and Y. Zhang, "Green communications with network cooperation: a concurrent transmission approach," *IEEE Commun. Letters*, vol. 16, no. 12, pp. 1952-1955, Dec. 2012.

[51] Z. Niu, Y. Wu, J. Gong, and Z. Yang, "Cell zooming for cost-efficient green cellular networks," *IEEE Commun. Magazine*, vol. 48, no. 11, pp. 74-79, Nov. 2010.

[52] T. Han and N. Ansari, "On optimizing green energy utilization for cellular networks with hybrid energy supplies," *IEEE Trans. Wireless Commun.*, to appear.

[53] S. Zhang and P. Wang, "Optimal trade-off between power saving and QoS provisioning for multicell cooperation networks," *IEEE Wireless Commun.*, vol. 20, no. 1, pp. 90-96, Feb. 2013.

[54] Z. Niu, "TANGO: traffic-aware network planning and green operation," *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 25-29, Oct. 2011.

[55] A. da Silva, M. Meo, and M. Marsan, "Energy-performance trade-off in dense WLANs: a queuing study," *Computer Networks*, vol. 56, pp. 2522-2537, March 2012.

[56] L. B. Le, D. Niyato, E. Hossain, D. I. Kim, and D. T. Hoang, "QoS-aware and energy-efficient resource management in OFDMA femtocells," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 180-194, Jan. 2013.

[57] F. Liu, K. Zheng, W. Xiang, and H. Zhao, "Design and performance analysis of an energy-efficient uplink carrier aggregation scheme," *IEEE Journal Selected Areas Commun.*, to appear.

[58] N. Mastronarde and M. van der Schaar, "Fast reinforcement learning for energy-efficient wireless communication," *IEEE Trans. Signal Processing*, vol. 59, no. 12, pp. 6262-6267, Dec. 2011.

[59] H. Kwon and B. G. Lee, "Energy-efficient scheduling with delay constraints in time-varying uplink channels," *Journal of Commun. and Networks*, vol. 10, no. 1, pp. 28-37, March 2008.

[60] X. Lu, E. Erkip, Y. Wang, and D. Goodman, "Power efficient multimedia communication over wireless channels," *IEEE Journal Selected Areas Commun.*, vol. 21, no. 10, pp. 1738-1751, Dec. 2003.

[61] T. H. Lan and A. H. Tewfik, "A resource management strategy in wireless multimedia communications-total power saving in mobile terminals with a guaranteed QoS," *IEEE Trans. Multimedia*, vol. 5, no. 2, pp. 267-281, June 2003.

[62] T. Han and N. Ansari, "On greening cellular networks via multicell cooperation," *IEEE Wireless Commun.*, vol. 20, no. 1, pp. 82-89, Feb. 2013.

[63] 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.

[64] C. Y. Chang, W. Liao, H. Y. Hsieh, and D. S. Shiu, "On optimal cell activation for coverage preservation in green cellular networks," *IEEE Trans. Mobile Computing*, to appear.

[65] N. Saxena, B. J. R. Sahu, and Y.S. Han, "Traffic-aware energy optimization in green LTE cellular systems," *IEEE Commun. Letters*, vol. 18, no. 1, pp. 38-41, Jan. 2014.

[66] J. Wu, S. Zhou, and Z. Niu, "Traffc-aware base station sleeping control and power matching for energy-delay tradeoffs in green cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 4196-4209, Aug. 2013.

[67] S. Navaratnarajah, A. Saeed, M. Dianati, and M. A. Imran, "Energy efficiency in heterogeneous wireless access networks," *IEEE Wireless Commun.*, vol. 20, no. 5, pp. 37-43, Oct. 2013.

[68] L. Saker, S. E. Alayoubi, R. Combes, and T. Chahed, "Optimal control of wake up mechanisms of femtocells in heterogeneous networks," *IEEE Journal Selected Areas Commun.*, vol. 30, no. 3, pp. 664-672, April 2012.

[69] A. Conte, A. Feki, L. Chiaraviglio, D. Ciullo, M. Meo, and M. A. Marsan, "Cell wilting and blossoming for energy efficiency," *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 50-57, Oct. 2011.

[70] A. P. Azad, S. Alouf, E. Altman, V. Borkar, and G. S. Paschos, "Optimal control of sleep periods for wireless terminals," *IEEE Journal Selected Areas Commun.*, vol. 29, no. 8, pp. 1605-1617, Sep. 2011.

[71] A. Agarwal and A. K. Jagannatham, "Optimal wake-up scheduling for PSM delay minimization in mobile wireless networks," *IEEE Wireless Commun. Letters*, to appear.

[72] R. Wang, J. Tsai, C. Maciocco, T. Y. C. Tai, and J. Wu, "Reducing power consumption for mobile platforms via adaptive traffic coalescing," *IEEE Journal Selected Areas Commun.*, vol. 29, no. 8, pp. 1618-1629, Sep. 2011.

[73] H. Yan, S. A. Watterson, D. K. Lowenthal, K. Li, R. Krishnan, and L. L. Peterson, "Client-centered, energy-efficient wireless communication on IEEE 802.11b networks," *IEEE Trans. Mobile Computing*, vol. 5, no. 11, pp. 1575-1590, Nov. 2006.

[74] H. Zhu and G. Cao, "On supporting power-efficient streaming applications in wireless environments," *IEEE Trans. Mobile Computing*, vol. 4, no. 4, pp. 391-403, Aug. 2005.

[75] Y. Jin, J. Xu, and L. Qiu, "Energy-efficient scheduling with individual packet delay constraints and non-ideal circuit power," *Journal Commun. and Networks*, vol. 16, no. 1, pp. 36-44, Feb. 2014.

[76] K. D. Turck, S. D. Vuyst, D. Fiems, S. Wittevrongel, and H. Bruneel, "Performance analysis of sleep mode mechanisms in the presence of bidirectional traffic," *Computer Networks*, vol. 56, pp. 2494-2505, March 2012.

[77] S. Videv, J. S. Thompson, H. Haas, and P. M. Grant, "Resource allocation for energy efficient cellular systems," *EURASIP Journal on Wireless Communications and Networking*, vol. 181, pp. 1-15, May 2012.

[78] Z. Ren, S. Chen, B. Hu, and W. Ma, "Energy-efficient resource allocation in downlink OFDM wireless systems with proportional rate constraints," *IEEE Trans. Vehicular Technology*, to appear.

[79] Y. L. Chung, "Rate-and-power control based energy-saving transmissions in OFDMA-based multicarrier base stations," *IEEE Systems Journal*, to appear.

[80] L. Chen, Y. Yang, X. Chen, and G. Wei, "Energy-efficient link adaptation on rayleigh fading channel for OSTBC MIMO system with imperfect

CSIT," *IEEE Trans. Vehicular Technology*, vol. 62, no. 4, pp. 1577-1585, May 2013.

[81] L. C. Wang, W. C. Liu, A. Chen, and K. N. Yen, "Joint rate and power adaptation for wireless local area networks in generalized nakagami fading channels," *IEEE Trans. Vehicular Technology*, vol. 58, no. 3, pp. 1375-1386, March 2009.

[82] L. P. Qian, Y. J. Zhang, Y. Wu, and J. Chen, "Joint base station association and power control via benders decomposition," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1651-1665, April 2013.

[83] V. A. Siris, M. Anagnostopoulou, "Performance and energy efficiency of mobile data offloading with mobility prediction and prefetching," *Proc. IEEE WoWMoM'13*, pp. 1-6, June 2013.

[84] N. Ristanovic, J. Y. Le Boudec, A. Chaintreau, and V. Erramilli, "Energy efficient offloading of 3G networks," *Proc. IEEE MASS'11*, pp. 202-211, Oct. 2011.

[85] S. Tarkoma, M. Siekkinen, E. Lagerspetz, and Y. Xiao, "Smartphone energy consumption: modeling and optimization," *Cambridge University Press*, 2014.

[86] T. A. Le, S. Nasseri, A. Z. Esfahani, M. R. Nakhai, and A. Mills, "Power-efficient downlink transmission in multicell networks with limited wireless backhaul," *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 82-88, Oct. 2011.

[87] M. Z. Shaker, K. A. Qaraqe, M. S. Alouini, E. Serpedin, and M. A. Imran, "Green heterogeneous small-cell networks: toward reducing the CO_2 emissions of mobile communications industry using uplink power adaptation," *IEEE Wireless Commun.*, vol. 51, no. 6, pp. 52-61, June 2013.

[88] H. Ghazzai, E. Yaacoub, M. S. Alouini, and A. A. Dayya, "Optimized smart grid energy procurement for LTE networks using evolutionary algorithms," *IEEE Trans. Vehicular Technology*, to appear.

[89] C. McGuire, M. R. Brew, F. Darbari, G. Bolton, A. McMahon, D. H. Crawford, S. Weiss, and R. W. Stewart, "HopScotch-a low-power renewable energy base station network for rural broadband access," *EURASIP Journal on Wireless Communications and Networking*, vol. 112, pp. 1-12, March 2012.

[90] T. Han and N. Ansari, "Powering mobile networks with green energy," *IEEE Wireless Commun.*, vol. 21, no. 1, pp. 90-96, Feb. 2014.

[91] B. Devillers and D. Gunduz, "A general framework for the optimization of energy harvesting communication systems with battery imperfections," *Journal of Commun. and Networks*, vol. 14, no. 2, pp. 130-139, April 2012.

[92] O. Orhan, D. Gunduz, and E. Erkip, "Throughput maximization for an energy harvesting communication system with processing cost," *Proc. IEEE Information Theory Workshop*, pp. 84-88, Sept. 2012.

[93] K. Tutuncuoglu and A. Yener, "Optimum transmission policies for battery limited energy harvesting nodes," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1180-1189, Feb. 2012.

[94] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in OFDMA systems with hybrid energy harvesting base station," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3412-3427, July 2013.

[95] K. Tutuncuoglu and A. Yener, "Sum-rate optimal power policies for energy harvesting transmitters in an interference channel," *Journal of Commun. and Networks*, vol. 14, no. 2, pp. 151-161, April 2012.

[96] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, "Transmission with energy harvesting nodes in fading wireless channels: optimal policies," *IEEE Journal Selected Areas Commun.*, vol. 29, no. 8, pp. 1732-1743, Sept. 2011.

[97] B. Gurakan, O. Ozel, J. Yang, and S. Ulukus, "Energy cooperation in energy harvesting communications," *IEEE Trans. Commun.*, submitted.

[98] C. F. Chiasserini and R. R. Rao, "Energy efficient battery management," *IEEE Journal Selected Areas Commun.*, vol. 19, no. 7, pp. 1235-1245, July 2001.

Muhammad Ismail (S'10, M'13) is an assistant research scientist in the Electrical and Computer Engineering department in Texas A&M University at Qatar. He received his PhD degree in Electrical and Computer Engineering from University of Waterloo, Canada in 2013, and MSc. and BSc. degrees in Electrical Engineering (Electronics and Communications) from Ain Shams University, Cairo, Egypt in 2007 and 2009, respectively. Dr. Ismail research interests include distributed resource allocation, green wireless networks, cooperative networking, smart grid, and biomedical signal processing. He is a co-recipient of the best paper award in IEEE ICC 2014. He served as a TPC member in the ICWMC in 2010 - 2014 and IEEE ICC 2014 and 2015. He served in the IEEE INFOCOM 2014 organizing committee as a web chair. He joined the International Journal On Advances in Networks and Services editorial board since January 2012. He was an editorial assistant for the IEEE Transactions on Vehicular Technology in the period January 2011 - July 2013. He has been a technical reviewer for several IEEE conferences and journals.

Weihua Zhuang (M'93,SM'01,F'08) has been with the Department of Electrical and Computer Engineering, University of Waterloo, Canada, since 1993, where she is a Professor and a Tier I Canada Research Chair in Wireless Communication Networks. Her current research focuses on resource allocation and QoS provisioning in wireless networks, and on smart grid. She is a co-recipient of several best paper awards from IEEE conferences. She received the Outstanding Performance Award 4 times since 2005 from the University of Waterloo and the Premier's Research Excellence Award in 2001 from the Ontario Government. Dr. Zhuang was the Editor-in-Chief of IEEE Transactions on Vehicular Technology (2007-2013), and the Technical Program Symposia Chair of the IEEE Globecom 2011. She is a Fellow of the IEEE, a Fellow of the Canadian Academy of Engineering (CAE), a Fellow of the Engineering Institute of Canada (EIC), and an elected member in the Board of Governors and VP Mobile Radio of the IEEE Vehicular Technology Society. She was an IEEE Communications Society Distinguished Lecturer (2008-2011).

Erchin Serpedin (F'13) received the specialization degree in signal processing and transmission of information from Ecole Superieure DElectricite (SU-PELEC), Paris, France, in 1992, the M.Sc. degree from the Georgia Institute of Technology, Atlanta, in 1992, and the Ph.D. degree in electrical engineering from the University of Virginia, Charlottesville, in January 1999. He is currently a professor in the Department of Electrical and Computer Engineering at Texas AM University, College Station. He is the author of 2 research monographs, 1 textbook, 90 journal papers and 150 conference papers, and serves currently as associate editor for IEEE TRANSACTIONS ON COMMUNICATIONS, Signal Processing (Elsevier), EURASIP Journal on Advances in Signal Processing, Physical Communication (Elsevier), and EURASIP Journal on Bioinformatics and Systems Biology. His research interests include statistical signal processing, information theory, bioinformatics, and genomics. He is an IEEE Fellow.

Khalid Qaraqe (M97,SM00) was born in Bethlehem. He received the B.Sc. degree in EE from the University of Technology, Baghdad in 1986, with honors. He received the M.Sc. degree in EE from the University of Jordan, Jordan, in 1989, and he earned his Ph.D. degree in EE from Texas AM University, College Station, TX, in 1997. From 1989 to 2004 Dr. Qaraqe has held a variety positions in many companies and he has over 12 years of experience in the telecommunication industry. He has worked for Qualcomm, Enad Design Systems, Cadence Design Systems/Tality Corporation, STC, SBC and Ericsson. He has also worked on numerous GSM, CDMA, WCDMA projects and has experience in product development, design, deployments, testing and integration. Dr. Qaraqe joined the department of Electrical Engineering of Texas AM University at Qatar, in July 2004, where he is now professor. Dr. Qaraqe research interests include communication theory and its application to design and performance, analysis of cellular systems and indoor communication systems. Particular interests are in the development of 3G UMTS, cognitive radio systems, broadband wireless communications and diversity techniques.