Dynamic Resource Provisioning for Energy Efficiency in Wireless Access Networks: a Survey and an Outlook

Łukasz Budzisz, Fatemeh Ganji, Gianluca Rizzo, Marco Ajmone Marsan, Michela Meo, Yi Zhang, George Koutitas, Leandros Tassiulas, Sofie Lambert, Bart Lannoo, Mario Pickavet, Alberto Conte, Ivaylo Haratcherev and Adam Wolisz

Abstract—**Traditionally, energy efficiency aspects have been included in the wireless access network design space only in the context of power control aimed at interference mitigation, and for the increase of the terminal battery lifetime. Energy consumption of network components has also, for a long time, not been considered an issue, neither in equipment design, nor in network planning and management. However, in recent years, with the user demand increasing at nearly exponential pace, and margins rapidly shrinking, concerns about energy efficiency have been raised, with the objective to reduce network operational costs (not to mention the environmental issues). Installing more energy-efficient hardware does not seem to fully solve the problem, since wireless access networks are almost invariably (over)provisioned with respect to the peak user demand. This means that efficient resource management schemes, capable of controlling how much of the network infrastructure is actually needed and which parts can be temporarily powered off to save energy, can be extremely effective and provide quite large cost reductions.**

Considering that most of the energy in wireless access networks is consumed in the radio part, a dynamic provisioning of wireless access network resources is crucial, to achieve energy-efficient operation. The consensus on this approach in the research community has been wide in the last few years, and a large number of solutions was proposed. In this paper, we survey the most important proposals, considering the two most common wireless access technologies, namely cellular and WLAN. Main features of the proposed solutions are analyzed and compared, with an outlook on their applicability in typical network scenarios that also include cooperation between both access technologies. Moreover, we provide an overview of the practical implementation aspects that must be addressed to achieve truly energy-efficient wireless access networks, including current standardization work, and trends in the development of energy-efficient hardware.

Ł. Budzisz and A. Wolisz are with Telecommunication Networks Group, TU Berlin, Germany (e-mail: [budzisz, wolisz]@tkn.tu-berlin.de).
F. Ganji is now with Department of Security in Telecommunication, TU Berlin/Telekom Innovation Laboratories (e-mail: fatemeh.ganji@telekom.de).
G. Rizzo is currently with the University of Applied Sciences of Western Switzerland (HES-SO VS), Sierre, Switzerland (e-mail: gianluca.rizzo@hevs.ch).
M. Ajmone Marsan, M. Meo and Y. Zhang are with the Department of Electronics and Telecommunications, Politecnico di Torino, Italy (e-mail:[marco.ajmone, michela.meo, yi.zhang]@polito.it).
G. Koutitas and L. Tassiulas are with the Computer Engineering and Telecommunications, University of Thessaly, Volos, Greece (e-mail: george.koutitas@gmail.com; leandros@inf.uth.gr)
S. Lambert, B. Lannoo and M. Pickavet are with Department of Information Technology, Ghent University - iMinds, Belgium (e-mail: [sofie.lambert, bart.lannoo, mario.pickavet]@intec.ugent.be).
A. Conte and I. Haratcherev are with Alcatel-Lucent Bell Labs, France (e-mail: [alberto.conte,ivaylo.haratcherev]@alcatel-lucent.com).

Index Terms—**Green networking, on/off switching, sleep modes, cellular networks, WLAN, femtocells**

I. INTRODUCTION

In recent years, wireless data communications have become increasingly popular: the advent of smartphones, tablets and laptops has enabled the widespread use of new bandwidth-intensive applications, such as mobile web browsing and mobile video streaming. This has resulted in an immense growth of mobile data use, which is expected to continue in the coming years: a more than ten-fold increase in mobile data traffic between 2013 and 2018 is predicted in recent forecasts from Ericsson and Cisco [1, 2]. Fixed data traffic, including traffic transmitted through WLAN (typically based on IEEE 802.11) access points (APs), will show a slower growth, but remain dominant in absolute volume. According to Cisco, traffic from WLAN and mobile devices combined will more than triple between 2013 and 2018, exceeding the amount of IP traffic from wired devices by 2016 [2]. As more and more wireless data is being transported, the speed at which this data can be transmitted also shows an impressive growth: historically, data transmission rates in both cellular and WLAN networks have been rising by a factor of about ten every five years [3]. To offer this high-speed wireless access, network operators are deploying increasingly complex and dense access networks, with more radios, i.e., cellular base stations (BSs) and WLAN APs per unit area, thus vastly increasing their energy consumption [3, 4]. This escalation of energy consumption in wireless access networks is of great concern for a number of reasons.

First, growing environmental concerns mandate a reduction of greenhouse gas emissions, which have been recognized as a global threat to environment protection and sustainable development. Communication networks are an important contributor to the carbon footprint. Their share of the global electricity consumption has increased by almost 40 % from 2007 to 2012, rising from 1.3 % to 1.8 % [5]. Based on the mentioned trends of rapid growth in mobile data traffic, it is easy to predict that this share will keep growing, unless drastic actions are taken to improve energy efficiency.

Second, coupling growing energy consumption with increasing energy costs, we can expect the energy bill due to wireless access network operation to become a critical component of the network operational expenditures (OPEX). In the case of cellular networks, the energy cost is already an important component of business models, accounting in some countries for as much as 50 % of network OPEX [3]. Energy efficiency is thus becoming a significant competitive advantage for cellular network operators. Moreover, corporate and social responsibility in terms of greenhouse gas emissions has an impact on the corporate brand image in modern society, giving companies an additional incentive to invest in energy efficient wireless access networks [6].

Consequently, the need to improve energy efficiency in wireless access networks has been established, and has led to numerous research works on the topic called *green radio communication* in recent years. This trend has to be distinguished from the research on the energy efficiency of single mobile devices (user-end perspective) and wireless sensors that has been present since long, because of their battery-limited operation mode, e.g., a good survey of energy-efficient cross-layer optimizations for end-user devices can be found in [7]; a survey of efficient routing techniques recommended for wireless multimedia sensor networks (WMSN) is provided in [8]. Recently, due to financial and environmental considerations mentioned above, the research on green radio communication has quickly taken off. Since the main contributor to the carbon footprint and energy consumption of wireless access networks is the operation of the Radio Access Network (RAN), this is where research efforts have been first directed [3, 6, 9, 10]. Energy savings in the RAN can be achieved in many different ways; for a broad overview of energy saving approaches in wireless access networks we refer the reader to the following surveys [9–12]. Both further improvements of hardware, and use of newer access technologies, e.g., LTE (Long-Term Evolution), can help to significantly improve energy efficiency of RANs, e.g., see study [13] on how carrier aggregation, heterogeneous deployment, and the extended support for multiple-input multiple-output (MIMO) antennas can enable the power saving potential in the most recent generation of cellular networks. However, due to the rising number of radios, solutions that reduce the power consumption of a RAN as a whole are required. Therefore, in this article we will look in more detail into a group of particularly promising approaches, namely, energy-efficient management of BSs and APs, where unused radios are partly or completely switched off in periods of low load, trying to adapt the amount of active resources to the fluctuations in the traffic demand [14]. This forms the main scope of this article and, to the best of our knowledge, it is the first article surveying to that extent the BSs/APs management strategies for both cellular and WLAN environments.

The article is further structured as follows. In Section II the significant potential of BS/AP management algorithms to increase the energy efficiency in RANs is demonstrated, together with a discussion of the support provided by the current cellular and WLAN architectures. Next, in Section III a taxonomy to classify these management schemes is proposed, to bring better structure to our survey. The most prominent solutions that have been proposed so far in the literature, some of which were developed within the 7th EU Framework Programme TREND project [15], are then discussed in more detail in Section IV. This section also provides a comprehensive overview of management algorithms for both analyzed wireless access network technologies, and proposes indications on the further research directions. Whenever possible, we also try to assess the energy saving potential of the different approaches, in order to provide the reader with an indication of the benefits achievable with the different management algorithms: of course, energy savings heavily depend on the specific network to which the management algorithms are applied, so our indications should be only considered as rough estimations. When devising BS/AP management strategies, there is a number of technological and regulatory developments which can pave the way to a widespread development of these solutions in energy-efficient wireless access networks of the future. To this end, Section V analyzes the main risks and opportunities inherent in the practical application of the analyzed strategies. In Section VI, an outlook on how technologies can be adapted to make the use of BS and AP management algorithms feasible and effective, is provided. Finally, general conclusions are given in Section VII.

II. TECHNOLOGIES TO SUPPORT ENERGY-EFFICIENT RAN OPERATION

RAN technologies and algorithms, including those of cellular networks and *dense WLANs*, i.e., featuring thousands of APs per km^2, have traditionally been designed targeting performance maximization at full load. Indeed, the full load working condition is the most critical and challenging one in terms of efficient use of the available resources, and, consequently, of capital expenditure (networks are typically dimensioned so as to efficiently sustain the estimated peak traffic). However, most of the time, networks work at low to medium load. This is especially true for RANs in which the user aggregation level is very limited: cells cover limited areas in which users exhibit similar behaviors, so that the load profile presents large variations between peak and off-peak values, with long periods of low load. As an example, Fig. 1 shows the traffic measured on cells of an Italian mobile network in operation: the solid lines refer to a cell in a business area; the dashed lines refer to a cell in a consumer area; the empty markers identify the profile of a week-day; the solid markers refer to a weekend day. Traffic values are obtained by averaging the measurements (at 15 minute intervals) collected during a week, and are then normalized to the peak average value in the cell. Since the end user behavior is given by the combination of user activity and mobility patterns, we see that the traffic fluctuates significantly during a day, and that periods of low activity are long [16].

While typical operative working regions of networks are at low to medium load, the dimensioning of individual devices, as well as that of the entire network, has somehow neglected the low to medium load operative points. Similarly, energy consumption and operation costs have been rarely taken into account as design variables, energy being traditionally perceived as always available and very cheap. As a consequence,

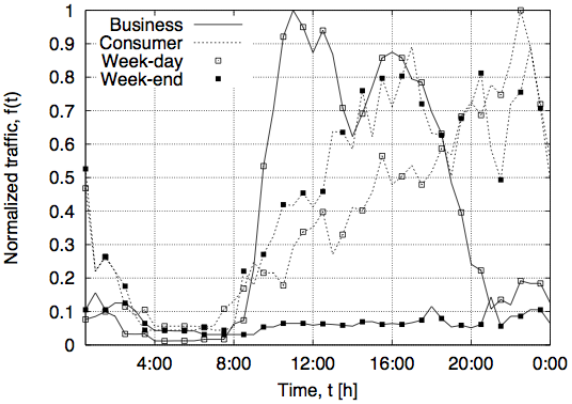


Fig. 1. Daily traffic profiles for a cell in a business area and a cell in a consumer area, week-day and week-end profiles measured in a network in operation.

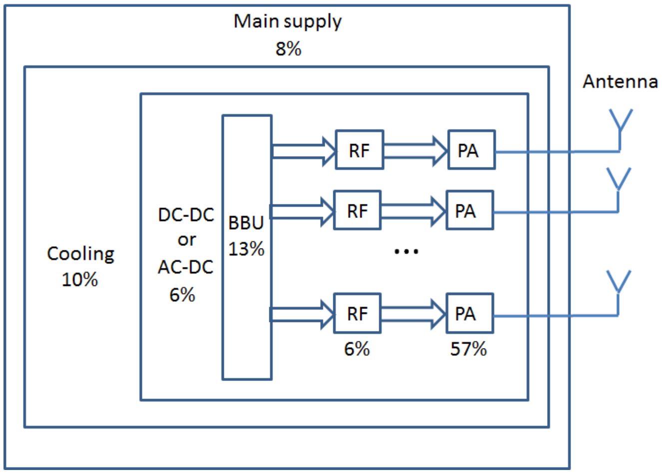


Fig. 2. Components of a BS and their power consumption percentages.

the power versus load profiles of BSs and APs, and of the network, exhibit very limited *load proportionality*, with values of the energy consumption at zero or low load that are large fractions, typically about 60-80%, of the consumption at full load. The limited load proportionality of the wireless access devices can be a serious obstacle to the objective of saving energy. Indeed, the use of little load-proportional devices, subject to a traffic profile like these shown in Fig. 1, causes the devices to be empty or almost empty for long periods of time, with a potentially large energy waste due to the large consumption even at zero load. This is why this paper specifically focuses on the energy savings achieved by turning off network elements when their traffic load is low, as well as on the aspects of how to manage the on/off procedure in a proper way.

One of the main design objectives today is, thus, to reduce the energy consumption of individual devices and of the entire network when low or no traffic is carried, by making the consumed power more load proportional. The research efforts in this field have taken two main directions. On one hand, manufacturers are focusing on the design of devices that consume less, and whose consumption is more load proportional; on the other hand, new network architectures are being proposed to make the entire network consume less, and in a more load proportional fashion. Considering that the electricity cost of the RAN is one of the main contributions to the OPEX of a network, targeting the energy efficiency of the RAN also leads to achieving cost efficiency.

A. Cellular networks

In cellular wireless access networks, end users connect to BSs through wireless channels, and each BS, in turn, is connected to some other network elements through either a wired or a wireless point-to-point link, which is part of the backhaul network. Several generations of cellular wireless access technologies exist, with the most recent being LTE (the Long-Term Evolution of UMTS), also commonly called 4G. Using LTE, the downlink data transfer can reach data rates up to 100 Mbps. Assuming the power necessary to serve each user [W/user] as a relevant power consumption metric for (cellular) wireless access networks, it is claimed in [17] that, in urban areas with a typical user density of 300 users/km^2, LTE requires only 18 W/user, versus 27 W/user for WiMAX, and 68 W/user for HSPA, due to the fact that an LTE BS can serve much more users. Therefore, LTE is a more energy-efficient technology, in W/user, compared to earlier-generation technologies.

In an urban area, an LTE macro BS can cover an area of about 0.22 km^2 with a range of about 500 m. In suburban/rural environments, the covered area can grow to 2.6 km^2 with the same transmission power. Considering that the power consumption of the macro BS is around 1 kW, the power consumption per unit area is approximately 4500 W/km^2 for the urban area, and 400 W/km^2 for the suburban/rural area, respectively [17].

As shown in Fig. 2, a BS is comprised of a baseband unit (BBU) and one or more transceivers, each one of which contains a radio frequency part (RF), a power amplifier (PA), and an antenna, connected through a feeder. In addition, a BS, which normally uses an input voltage of 48 V, also contains an AC-DC converter, a DC-DC converter, and a cooling system (sometimes just a fan) [18].

In traditional deployments, the BS equipment is located far from the antenna, so that long feeder cables are necessary, and power losses occur. In the case of LTE BSs, quite often the RF and PA components can be located close to the antenna, so as to eliminate the feeder cable losses. This layout is called Remote Radio Unit (RRU) or Remote Radio Head (RRH). The RF and PA components can even be integrated into the antenna. Besides allowing fewer losses, an additional advantage of the RRU layout is that in some cases cooling becomes unnecessary. LTE adopts OFDM (Orthogonal Frequency-Division Multiplexing) modulation schemes with high Peak-to-Average-Power Ratio (PAPR), forcing the amplifier to operate in a linear region between 6 and 12 dB lower than the saturation point. This reduces the Adjacent Channel Interference (ACI), but increases power consumption. The numbers in Fig. 2 show the percentage contributions to

TABLE I
PARAMETERS OF THE CONSUMPTION MODEL FOR LTE BSS

Macro BS type	N_{TX}	P_{max} [W]	P_0[W]	Δ_p
No RRU	6	20	130	4.7
With RRU	6	20	84	2.8

power consumption of various components of an LTE macro BS [18]. The most power hungry element is the PA.

The peak power consumption of an LTE macro BS is about 1350 W, in the case of 3 sectors, 2 antennas/PA per sector, one carrier with 10 MHz bandwidth, and 2 x 2 MIMO. With a RRU configuration, the peak value decreases to about 800 W, thanks to a reduction of the energy consumed by the PA, and by cooling. The actual instantaneous power consumption of a BS depends on the PA load, which in turn depends on the traffic carried by the BS. The relation between the emitted power and the load ρ can be expressed by a linear function:

$$P(\rho) = a + b\rho \quad 0 \leq \rho \leq 1, \tag{1}$$

where a is the power consumption when the BS is active, but carries no traffic, and b is the load proportionality parameter. The main contribution to a is from the components that have fixed power consumption, independent of the traffic, e.g., cooling units, main supply, fixed energy cost of processing power. On the other hand, the components that have a traffic-dependent power consumption are BBU, RF and PA.

In case of LTE macro BS, the parameters a and b depend on the number of antennas according to the following expression:

$$P(\rho) = N_{TX} \left(P_0 + \Delta_p P_{max} \rho \right) \quad 0 \leq \rho \leq 1, \tag{2}$$

where N_{TX} is the number of antennas, P_{max} is the maximum power out of the PA, P_0 is the power consumed when the RF output is null, Δ_p is the slope of the emission-dependent consumption. By properly optimizing the BS, the top devices can achieve a load proportionality of about 40 %, meaning that at zero load the consumption is about 60 % of the one at full load, and, between zero and full load, power grows linearly with the load.

Table I reports the typical values of the parameters for an LTE macro BS with and without RRU [18].

The most recent BS designs are focusing on the possibility to also integrate the baseband unit of several BSs into a unique unit. This approach, which is usually referred to as Cloud RAN, has two main advantages over traditional BS architectures: first, it allows the signal processing to be optimized at a multi-cell level; second, the number of active BBUs can effectively be adapted to the actual load, leading to high energy efficiency.

Further gains are expected by further reducing the BBU consumption, and the PA consumption. Both these objectives require, in different ways, the use of sleep modes. For the BBU, the idea is to properly reduce baseband processing according to the BS load and by powering off some baseband processing capabilities when load is low. Similarly, the possibility to use stand-by-modes for the PA is under investigation.

Usually, most of the components and subsystems of a BS admit various sleep modes that differ for the degree of functionality that is inhibited when the sleep mode is entered (except for the case that parts of the transceiver chain cannot be switched off separately). Each sleep mode has a specific energy consumption value that depends on how deep the sleep mode is, where deeper sleep modes are associated with higher degrees of reduction of functionality, as well as to lower consumption. Each sleep mode requires a wake-up time, defined as the time needed to restore the normal full functionality of a component or subsystem, starting from the given sleep mode. The deeper the sleep mode is, the longer the wake-up time is. Typical values range from tens of seconds to a couple of minutes for small cells, and up to 10-15 minutes for macro BSs.

Wake-up times have a strong impact on the feasibility of BS management algorithms. One of the major requirements of the operators for the adoption of sleep modes is the fact that they should not affect the quality of service provided to users, and that the network is guaranteed to operate smoothly and continuously. Thus, when using sleep modes, at the occurrence of unexpected traffic increases or bursts, i.e., when capacity is needed, sleeping BSs should quickly be powered on and return to full operation mode.

The definition of sleep modes, that includes energy consumption values and wake-up times, is necessary to properly design the BS management algorithm. The BS management algorithm is the core mechanism to exploit sleep modes, and consists of a strategy with which BSs in a given area enter or exit sleep modes; in other terms, the algorithm defines, at any instant of time, the network configuration that consists of the state of each BS.

Although several recent research works on energy-efficient BSs propose the adoption of dynamic and frequent sleep modes [19–21], still most of the BSs that are deployed today were designed foreseeing only occasional switch-on and switch-off, so that the use of sleep modes for these devices might be critical. In particular, frequent changes in powering state might threaten the robustness of some components of the BS and finally lead to higher failure rates of the device or parts of it. Depending on the kind of BS management algorithm, the BS might need to enter a number of sleep modes that varies between one per day (with switches between a couple of possible network configurations) to frequencies of the order of more than one per hour, when decisions about network configurations are taken after dynamic traffic estimates or measurements. A quite interesting issue is, thus, the assessment of the robustness of a BS to frequent power state changes, where with robustness we mean the lifetime of the equipment, the number and type of maintenance and repair events. For an evaluation of the feasibility of a BS management algorithm, these values should be related to the normal lifetime of a technology, that is the typical time between the technology deployment and the next generation of devices to be designed, produced and deployed in operative networks. In the past, this technological turnover was around 10-15 years.

The practical implementation of a BS management algorithm requires that the BS can be remotely controlled, so that a decider engine can, at any instant of time, command the proper BS power mode. This, in turn, implies some additional control processes of the BS hardware, some software or hardware interfaces to manage the BS power state changes. Similarly, other parameters of the BS might need to be remotely controlled. For example, the RF transmission power or the antenna tilt, might need to be adjusted depending on the configuration of the network, i.e., depending on the power state of the BS and of the adjacent BSs.

Last but not least, BS on/off strategies bring an extra cost to the network. The cost consists in the overhead of software/hardware updates, as well as the side effect brought by cell wilting Switching off a BS and using a larger cell might lead to a lower signal strength, and thus ”kill the user battery”, as pointed out in [22]. The energy saved by the on/off switching is thus traded off for more energy being drained from the user end.

B. Dense WLANs

The huge popularity of WLAN radio access networks stems from its flexibility and ease of installation, low cost of ownership and reasonably high data rate. Therefore, WLANs cover a wide range of deployment scenarios, from residential areas to big enterprises and university campuses. While an increase in the size of a WLAN leads to the installation of a large number of APs, the explosion of the user demand for capacity calls for dense deployments of the APs. More specifically, as demonstrated, e.g., in [23], the AP density nowadays reaches thousands of APs per square kilometer in campus and enterprise networks. Whereas such a dense configuration is adopted to carry the close-to-the-peak user traffic volume, it has been shown that the APs handle the peak traffic only for small portions of time. In fact, a considerable number of APs is idle when the user traffic is marginal, e.g., at nights, leading to a significant power wastage and thus requiring the adoption of more energy-efficient solutions.

In order to compare the power saving potential of approaches aiming at designing either more energy-efficient APs *(micro-level approaches)* or energy-aware WLAN architectures *(macro-level approaches)*, it is first crucial to look more closely at the power consumption of an AP. A rigorous study presented in [24] provides a comprehensive database for the power consumption of APs by various manufacturers and operating in different power-operation modes: transmitting, receiving and idle. The main conclusion drawn from this study is that however a considerable amount of power can be saved by switching off the transceivers of APs (operating in the idle mode), the final power saving result is still not satisfactory.

Numerous research studies have analyzed the power consumption of an AP in detail, presenting various power consumption models. Most typically, these models consist of two elements[1]: (1) the baseline power consumption, showing the power consumed by the processing and cooling circuits when the AP is idle, and (2) a linear function representing the power consumption of the transceiver (similarly to what has been described in Eq. (1); for more discussion see the use case presented in Section II-C). Regarding the latter element, it has been shown that the APs supporting MIMO techniques tend to consume more power than other types of APs [24, 25]. Moreover, supporting dual-band (i.e., 2.4 and 5 GHz) operation, in particular when both radios are switched on, leads to an increase in power consumption of an AP [24, 25]. The measurements conducted to verify the accuracy of these models, reveal that the baseline power consumption accounts for at least 70% of the total power consumption in WLAN [25]. Due to technological constraints, the baseline power consumption can be hardly modified to save power. Therefore, the micro-level approaches have limited potential for power saving. On the other hand, macro-level approaches, and more precisely, the on/off switching strategies that are advocated in this paper, deal with the power wastage problem by switching the idle APs off and re-configuring accordingly the RAN [14]. Although macro-level approaches outperform the micro-level ones in terms of power saving, limitations of the control and the management scheme of WLANs impose new requirements on manufacturers to provide more flexible and reliable schemes, facilitating a fast and accurate network re-configuration.

The current migration of dense WLANs from autonomous to centralized architectures [26], reflects the importance of a change in control and management schemes that can be beneficial for achieving an energy-efficient operation. While the APs perform all WLAN management and control functions in an autonomous architecture, in a centralized architecture (typical nowadays for campus and enterprise environments) either (1) only time-critical functions, e.g., exchange of management frames, is executed by the AP, whereas all the control and data traffic is routed to/from a central controller or, more commonly, (2) the central controller does not take complete control over all AP traffic, having however precise knowledge about the state of the network (the AP notifies the controller via a separate protocol), i.e., used channel, number of users associated, etc. Hence, the central controller has a global view of the network, and can decide about the actual network configuration according to the traffic handled by the APs. Despite the fact that this type of management scheme, especially in the variant where the AP handles all non-time critical functions, has some obvious shortcomings, e.g., single point of failure, processing latency, etc., on/off switching strategies seem highly beneficial in typical dense WLAN scenarios.

C. Case study

Common analysis method of sleep mode can be applied in both cellular networks and dense WLANs [27]. In order to quantify the relevance of sleep modes for power saving, consider a very simple and idealized scenario, where all cells (or APs) in a given area are identical and traffic is the same in all cells (or APs). Assume also that, when a given fraction of BSs (or APs) is put to sleep, the other BSs (APs) can provide radio coverage. When the aggregate traffic of k cells (APs) is so low that one cell (or AP) only can carry all of it,

---

[1] There are also power consumption models of a WLAN AP in the literature that only consist of the baseline element, e.g., [25].

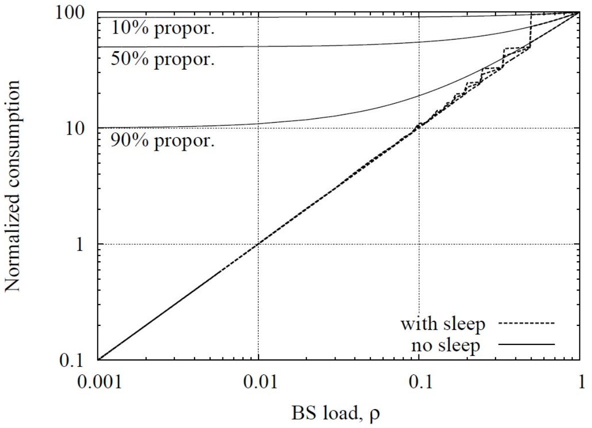


Fig. 3. Effect of the use of sleep modes, for different degrees of load proportionality

$k-1$ BSs (APs) can be put to sleep and one only can remain active. Then, given that the power versus load model is given by Eq. (1), and assuming that the consumption in sleep mode is given by P_{sleep}, the average energy consumption of the k BSs (APs) is:

$$P_S(\rho) = \frac{a + kb\rho}{k} + \frac{(k-1)P_{sleep}}{k} \qquad \rho < \frac{1}{k} \tag{3}$$

Indeed, if the fraction of load in a cell (AP) is less than $\frac{1}{k}$, then just one BS (AP) out of k is capable of carrying the traffic, and the other $k-1$ can be put to sleep. Sleeping BSs (APs) consume a power P_{sleep}, while the only BS (AP) that remains powered on consumes the energy resulting from the traffic of k BSs (APs). Assuming that we can always keep on the smallest number of BSs (APs), Eq. (3) holds as far as $\frac{1}{k+1} < \rho < \frac{1}{k}$: when the load goes below $\frac{1}{k+1}$, a number k of BSs (APs) can be switched off out of $k+1$.

Fig. 3 shows the behavior of the average consumption, $P_S(\rho)$, from Eq. (3), versus the load per cell, ρ, in a cellular network. Different degrees of load proportionality are considered, where with load proportionality we intend the fraction of consumption that is proportional to load, that is $b/(a+b)$ according to the notation in Eq. (1). We consider 3 values: i) 10% of load proportionality, as is the case of most of the BSs that are deployed today, ii) 50% of load proportionality, as in the top devices that are under development, and iii) 90% of proportionality, that is not realistic with today technologies and knowledge, but we consider it as a reference extreme case. Sleep modes have beneficial effects even when the devices are highly performing in terms of load proportionality.

III. TAXONOMY

In order to classify the techniques for energy-efficient dynamic provisioning in wireless access networks that appeared so far in the literature, we developed a taxonomy. The main goal of this taxonomy is to provide a framework that helps to point out all relevant design aspects, evaluate the advantages and shortcomings of the proposed solutions, and analyze their energy saving potential. Fig. 4 illustrates the structure of our taxonomy that consists of five non-overlapping branches, corresponding to the main characteristics of the analyzed proposals: (1) scope of application, (2) metrics, (3) type of the algorithm, (4) control scheme, and (5) evaluation method. Each branch is then subdivided into several categories, specific for each branch, as it will be further explained in what follows.

A. Scope

We first classify proposed approaches regarding their scope of application into two categories: (1) strategies developed to save energy in cellular networks, and (2) algorithms designed to reduce the energy expenditure in WLANs. Due to the different architectures of both networks, we need to further refine these categories. The first refinement is related to the radio access technologies that are being used. Power saving strategies developed for either cellular or WLANs may require availability of either single or multiple radio access technologies, thus naturally leading to creation of two subcategories: *homogeneous* and *heterogeneous* in each case (subcategories are not shown in Fig. 4 for simplicity). The second refinement is cellular network specific: in the context of power saving strategies it is also important to distinguish different deployment structures of the network, namely *flat* (where a single type of BSs is assumed, i.e., macro/micro-cells only) and *multi-tier* (macro-micro co-exist, and can cooperate with small cells) network architectures.

Our classification does not account for the type of services offered by the wireless network, or for the characteristics of the carried traffic. This is due to the consideration that in wireless networks today multimedia services and traffic are largely predominant, and expected to keep growing, both in share and in volume. As a consequence, all of the techniques for energy-efficient dynamic resource provisioning in wireless networks that will be discussed in the later sections of this paper must be able to cope with the requirements of multimedia traffic, such as the need to deliver to the end user an uninterrupted data stream that meets playout deadlines of video frames and/or audio segments. This implies stringent requirements on the BS and AP deactivation procedures, as well as on the handover algorithms. Some recent papers do specifically address energy efficiency in resource management for cellular and WLAN multimedia communications, e.g., [28–37].

B. Metrics

Strategies for energy-efficient dynamic provisioning in wireless access networks use different metric(s) to evaluate whether there is a need to power on additional radios (i.e., cellular BSs or WLAN APs) or the unnecessary radios may be switched off. There are four most common metrics used in this context: user demand, coverage, QoS and energy efficiency. Depending on the complexity of the algorithm, either only one of the specified metrics is chosen, or a combination of them.

1) User demand: User demand is the most common metric used for adjusting the number of active BSs/APs. User demand can be measured or estimated based on the observation of the number of users associated per given radio, the aggregate traffic generated by the active users, etc.

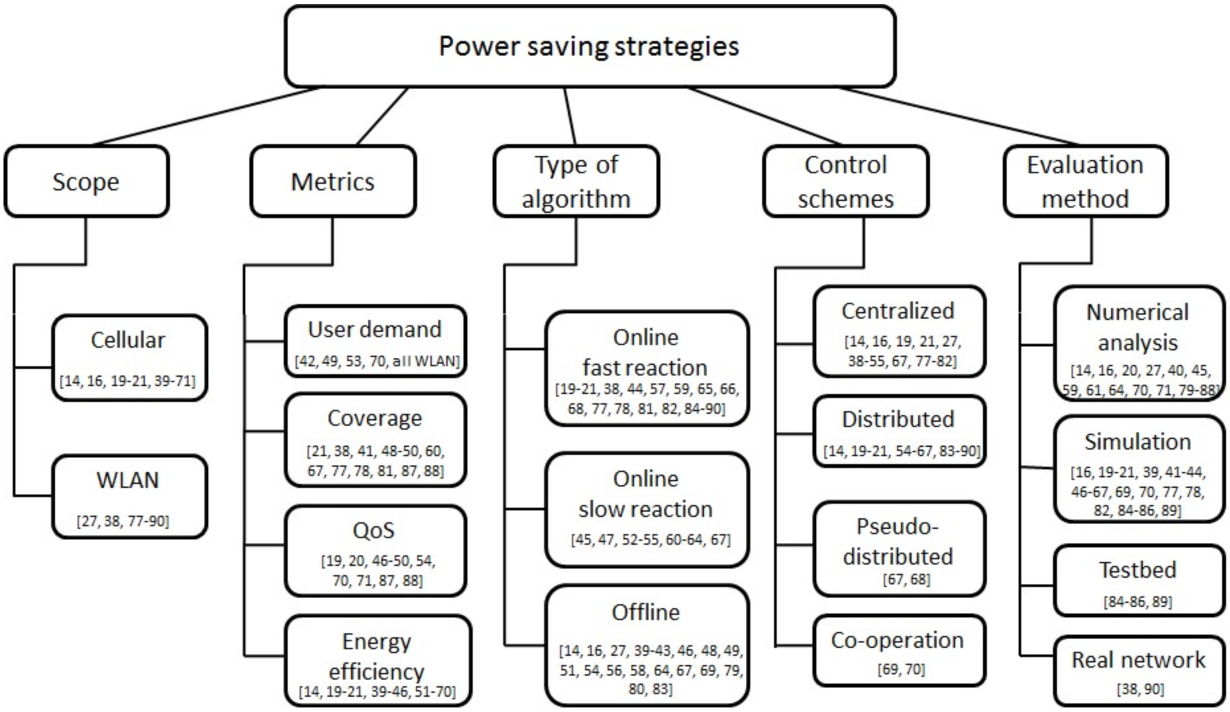


Fig. 4. Proposed taxonomy (for each category, references to all relevant articles are provided).

2) Coverage: Assessment of the RAN coverage is a key metric used by the on/off switching algorithms. Ideally, when reducing the number of inactive radios to save power, the coverage over the entire radio access network area should remain guaranteed. Hence, a power saving mechanism should be adopted to either calculate or estimate the coverage provided by the reduced set of radios. The calculation of the coverage was proven NP-hard, if the AP/BS are placed irregularly or in random fashion. In such cases tile-based approximations and mathematical optimization algorithms are deployed to estimate coverage.

3) QoS: As another option, the on/off switching strategies may also examine whether the quality of service offered by the reduced set of radios fulfills the minimum requirements on QoS. This is done based on the observation of one of the following parameters:

- **Throughput**: it has been shown that when a radio has to handle high traffic volumes, the user throughput declines dramatically due to the high probability of interference and frame losses. Therefore, this metric can provide a good insight into the fulfillment of the QoS requirements.
- **Delay**: this parameter usually refers to the end-to-end delay, i.e., the time that the user needs to transfer a frame to a desired destination. The end-to-end delay contains the transmission delay, the propagation delay and the processing delay. It is clear that under heavy traffic conditions, the end-to-end delay increases drastically, leading to a decrease of QoS.
- **Others**: other parameters, e.g., jitter, the number of dropped frames, etc., can also be good QoS indicators, although most of the research papers categorized according to our taxonomy use either coverage, throughput, delay or the combination of these to evaluate performance.

4) Energy Efficiency: Finally, various energy efficiency metrics are used in the literature to provide a quantitative analysis of the power saving potential of the on/off switching strategies. Since these strategies aim at reducing the overall energy/power consumption rather than improving the energy-/power-efficiency of a single radio, we refer to the overall energy-efficiency metrics here. Thus, three main types of energy-efficiency metrics can be distinguished:

- direct presentation of the power/energy saving achieved by means of the strategy (e.g., difference of power/energy consumption before and after adoption of the strategy, percentage of power saving, etc.).
- metrics that relate the energy consumption and the performance of the network, e.g., $bits/Joule$, etc.
- metrics relating the network coverage and the power consumption, e.g., km^2/W, $subscriber/W$, etc.

Metrics of the last type are often used for the comparison of the energy saving potential in different types of networks. The drawback of this approach is, however, that it does not show the capacity provided by each network, being thus less accurate.

C. Type of Algorithm

We further categorize power saving strategies according to one of the most important features, that is type of algorithm. From a general point of view, the problem of scaling down the number of radios while preserving the required QoS and coverage, can be seen as an optimization problem with constraints relevant to the selected metric(s). Based on how the optimization is realized, the following algorithm types can be distinguished:

- **Offline algorithms**: algorithms with a predefined schedule of on/off switching that is based on previously measured and processed information relevant to the given metrics. The most important advantages of using the offline algorithms are their low complexity and low processing overhead. However, it has been stated that these advantages can be achieved at the price of over- or under-estimation of the user demand during unexpected events (e.g., a group of users suddenly attempts to connect to the network) [38]. Due to spatial and temporal dependency of the user demand, different offline algorithms may be required at different locations in the network.
- **Online algorithms**: algorithms that estimate the user demand based on the real measurements of the parameters related to the applied metrics. In contrast to the offline algorithms, they are better suited to deal with unexpected events, although at the expense of more processing overhead attributed to the continuous measurements. Because of the constraints posed by the hardware, there is a need to further divide the online algorithms, depending on the time that is needed to make the adjustments to the network:
  - **Slow reaction**, the network reconfiguration process is relatively time-consuming, permitting only long-term changes (several tens of seconds to minutes). In this case the management algorithms can react, at best, based on average traffic measures or pre-computed statistics.
  - **Fast reaction**, the timescale of the reaction (few seconds at most) is short enough to allow quick adjustments of the network to sudden changes in the user demand. In this case it is possible to track the variations of traffic almost in real time (in case of ultra-fast reaction) or to turn-on/off cells in a way not perceptible by the user.

In some cases, algorithms presented in the literature can be implemented in both online and offline fashion. The type of algorithm used depends on the network dimensions, the control scheme and the applicability of the solution. For example, a centralized control scheme over a large network is difficult to be implemented in an online manner due to the potential delays that might be incorporated in the system. For that case, offline management might be the most practical scheme. On the other hand, distributed control techniques are more suitable for online management where decisions can be made locally and in real time.

D. Control Schemes

Another important aspect that influences the design of the analyzed power saving techniques is the control scheme based on which the on/off switching algorithm is developed. The control scheme in the network depends on the architecture of the network and the different application scenarios. Apart from the typical centralized and distributed control schemes, two approaches are worth mentioning: pseudo-distributed and co-operative.

1) Centralized: In this control scheme, the central controller collects the information about the metrics discussed in the last section, processes them, and makes and executes a decision about the status (on/off) of any radio. Although using a central controller simplifies network configuration and management, the most challenging problem is that the central controller becomes a single point of failure. Furthermore, when the central controller processes the collected information, the processing overhead is significantly high in large-scale radio access networks.

2) Distributed: Here, in contrast to the centralized approach, each radio can individually decide to be powered on or off. Therefore, the processing of the information is done in a distributed fashion. The most critical drawback of this control scheme is that each radio may not have a global view of the network, thus leading to a traffic management problem.

3) Pseudo-Distributed: The pseudo-distributed control scheme assumes that critical stations (BSs that are necessary for coverage) can control flexible stations (BSs usually deployed for capacity) that fall under their administrative domain. Critical to flexible station association is usually performed according to cell overlap criteria. With this approach, a smoother transition of QoS is achieved since critical stations have a better view of the network conditions. In general, the performance of the pseudo-distributed approach falls in between the centralized and the distributed case.

4) Cooperation as a special case: All control schemes mentioned so far are widely applied in homogeneous networks, where the co-existing radio networks do not cooperate in providing coverage and capacity. Conversely, various network operators may coordinate their efforts to reduce energy costs. Cooperation between at least two co-existing radio networks requires modified control schemes that have the ability to manage and re-configure various (types of) networks.

E. Evaluation methods

Last, but not least of the aspects in our classification is the method applied to evaluate the performance of a strategy in terms of metrics that were discussed earlier in this section. This dimension of the taxonomy is essential, as it validates possible comparisons of different strategies. Purely-theoretical evaluation methods (e.g., numerical analysis and simulation) provide an insight into the performance of an on/off switching algorithm, whereas experimental verification can reveal more detailed information about the implementation of an algorithm and possible difficulties that can be encountered in practice. We categorize the evaluation methods available in the literature into the following groups:

- **Numerical analysis**: this type of evaluation methods covers the algorithms developed to solve the optimization problems, formulated to find the minimum number of powered-on radios providing the coverage and the required capacity. Usually, there is a trade-off between the computational complexity and the scalability of the numerical analysis.
- **Simulation**: simulation techniques have been widely applied to reflect the operation a RAN. Due to their

simplicity, flexibility and sufficiently realistic models, the discrete-event simulators (e.g., ns-2, openWNS, etc.) are widely used to study the behavior of a RAN, and thus to evaluate the on/off switching strategies.
- **Testbed**: a testbed is a platform for conducting experiment and can be used to verify the practical feasibility of an on/off switching strategy in a controlled environment. This means that the generalization of results achieved through this type of experiments may not be valid, but a testbed enables the rigorous and repeatable testing of the functionality and the implementation of an on/off switching strategy.
- **Real network**: the performance of a on/off switching strategy can be evaluated by implementing the proposed algorithm into a real network. The most realistic evaluation can be performed in this manner; however, uncontrolled changes in RAN environment (e.g., traffic) limits the repeatability of the experiments.

IV. ANALYSIS OF EXISTING MANAGEMENT ALGORITHMS

Now, using the taxonomy presented in Section III, we survey the BS/AP management schemes presented so far in the literature that can provide significant gains in terms of energy efficiency. The main goal for all schemes is to keep QoS at acceptable levels, and provide a self-balanced and self-organized operation of the critical nodes of the network, while saving energy, and providing traffic-proportional power consumption of the network, as argued in Section II-C. We separately describe the most important solutions that were devised for *cellular networks* and *WLANs*. Tables II and III contain all the surveyed articles, summarizing shortly the nature of the examined algorithms, collecting the energy saving numbers reported in papers, as well as providing an overview of the applicability of the proposed solutions that are further commented and compared in Section IV-C.

A. Cellular Networks

For cellular networks, three different case studies are investigated: flat and multi-tier architectures belonging to one operator, as well as co-operation between mobile operators aiming at, as the main objective, switching off the redundant infrastructure that co-exists in the same geographical area, during periods of low traffic. All identified articles related to cellular networks are shown in Table II and are further sorted there according to the control algorithm applied.

1) Flat Network: Flat network architectures, as mentioned in Section III-A, consist of one layer of cells, e.g., typically macrocell only, or microcell only, or a combination of the two. Usually, such an infrastructure is under the administrative domain of one mobile operator. This means that the electricity costs related to the operation of the BSs are assigned to the OPEX of the operator. For the discussion of the power saving schemes, we further distinguish two different types of sub-architectures, i.e., *non-overlapping* architecture and *overlapping* architecture.

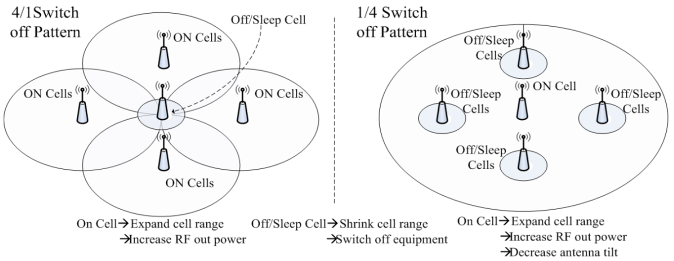


Fig. 5. Switching off of BSs in non-overlapping architectures.

a) Non-overlapping architectures: are such that the coverage areas of the BSs do not overlap. This assumption is usually met in microcell- or macrocell-only deployments. In that case BS switch on/off schemes can be achieved by *X/Y* switch off pattern. Parameter *X* denotes the number of BSs that remains on, while *Y* BSs can be switched off (set in sleep mode) [39, 40]. Since there is no overlapping between the BS cells, the BSs that are in the *on* state should increase their cell range to provide the additional geographical coverage. There are two options to follow. In the first case, the BSs surrounding the cell that will be switched off increase slightly their cell range to fill the coverage hole that appears in the middle. In the second case, the peripheral BSs are switched off, while the central BS that remains *on* needs to substantially increase its cell range, so as to fill the coverage hole from the surrounding stations [54]. Both approaches are illustrated in Fig. 5.

The decision on which approach to follow is a function of the architecture of the network and the available cell range of the BSs. To be more precise, in the latter case, where the central BS is supposed to fill the coverage hole created by the surrounding BS, an antenna tilt is required together with the increase of the RF output power. In general, when $X > Y$ a slight increase of the RF output power of the BSs that remain in the *on* state is enough to provide coverage. Otherwise($X \leq Y$), an increase of the RF output power of the central BS and a reduction of the antenna tilt is required to provide full coverage. In the *non-overlapping* architecture, the most commonly used control scheme that decides on the state of operation of a BS (on/off) or the antenna tilt, can be centralized or distributed. In the centralized approach, a central controller that has under its administrative domain all BSs, sends the required commands to the BSs according to traffic/load criteria in each cell of the network. In the distributed case, each *X/Y* group of BSs in the network decides to change its state of operation. The grouping of BSs in the *X/Y* scheme is made by the central controller of the network or by the mobile network operator [41, 54]. For real application scenarios, a centralized control scheme is usually preferred

TABLE II
SUMMARY OF POWER SAVING ALGORITHMS DEVISED FOR CELLULAR ENVIRONMENTS (ALGORITHMS WITH CENTRALIZED CONTROL SCHEME ONLY).

ARTICLE	MAIN CONTRIBUTION	SCOPE	METRICS	TYPE OF ALGORITHM	CONTROL SCHEME	EVALUATION METHOD	POWER SAVING POTENTIAL
[16]	Use renewable energy to power the BS to achieve zero grid electricity utilization.	Cellular Homogeneous Flat	1. Energy Efficiency	Offline	Centralized	Numerical Analysis, Simulation	up to 40%
[39]	Investigation of energy saving of switching off multiple BSs in both homogeneous and heterogeneous scenario.	Cellular Hom. and Het. Flat	1. Energy Efficiency	Offline	Centralized	Simulation	around 40% to 50%
[40]	Calculation of energy saving rates according to traffic variation during the day, with different sleeping schemes for different network topologies.	Cellular Homogeneous Flat	1. Energy Efficiency	Offline	Centralized	Numerical Analysis	around 30% on average
[41]	Investigation of energy efficient planning and management of BS networks and combination of the above.	Cellular Homogeneous Multi	1. Energy Efficiency 2. Coverage	Offline	Centralized	Simulation	N/A
[42]	Stochastic Analysis of theoretically optimal Energy savings and SINR.	Cellular Homogeneous Flat	1. Energy Efficiency 2. User Demand	Offline	Centralized	Simulation	N/A
[43]	The energy efficient effects of a joint deployment of macrocells and femtocells are investigated.	Cellular Heterogeneous Multi	1. Energy Efficiency	Offline	Centralized	Simulation	up to 60%
[44]	Dynamically adjusting cell sizes to help reduce energy consumption.	Cellular Heterogeneous Multi	1. Energy Efficiency	Online Fast reaction	Centralized	Simulation	up to 40%
[45]	Analyze practical issues when implementing sleep mode in base stations. Define a hysteresis when switch on/off BS to avoid pingpong effect.	Cellular Homogeneous Flat	1. Energy Efficiency	Online Slow reaction	Centralized	Numerical Analysis	up to 60%
[46]	Propose an heuristic, based on optimizing user association in order to turn off as many BS as possible. It shows that it has little impact on QoS.	Cellular Homogeneous Flat	1. Energy Efficiency 2. QoS (Throughput, Delay)	Offline	Centralized	Simulation	around 25-50%
[47]	Automatically switching off unnecessary cells, modifying the radio topology, and reducing the radiated power with methods such as bandwidth shrinking and cell micro-sleep.	Cellular Homogeneous Flat	1. QoS (Traffic)	Online	Centralized	Simulation	N/A
[48]	Introducing a new parameter for traffic estimation, deploying smaller but more cells to provide coverage.	Cellular Homogeneous Multi	1. QoS (traffic) 2. Coverage	Offline	Centralized	Simulation	N/A
[49]	Deployment of sleeping strategies and small cells.	Cellular Homogeneous Multi	1. User Demand 2. Coverage	Offline	Centralized	Simulation	N/A
[50]	Approach based on traffic conditions in neighboring cells to reduce power consumption of the network at the microcell or picocell level.	Cellular Homogeneous Flat	1. QoS (Traffic) 2. Coverage	Online	Centralized	Simulation	up to 75%
[51]	Optimal energy savings in cellular access networks. Investigation of centralized control schemes.	Cellular Homogeneous Flat	1. Energy efficiency	Offline	Centralized	Simulation	N/A
[52]	Proposal of several sleep modes algorithms based on heuristics.	Cellular Homogeneous Flat	1. Energy Efficiency	Online	Centralized	Simulation	up to 50%
[53]	A heuristic to establish a baseline of active base station fractions in macrocell and femtocell networks.	Cellular Homogeneous Multi	1. User demand 2. Energy efficiency	Online	Centralized	Simulation	25-50% for macro and over 80% for femto

TABLE II
SUMMARY OF POWER SAVING ALGORITHMS DEVISED FOR CELLULAR ENVIRONMENTS - CONTINUATION (ALGORITHMS APPLICABLE WITH EITHER CENTRALIZED OR DISTRIBUTED AND PURELY DISTRIBUTED CONTROL SCHEME).

ARTICLE	MAIN CONTRIBUTION	SCOPE	METRICS	TYPE OF ALGORITHM	CONTROL SCHEME	EVALUATION METHOD	POWER SAVING POTENTIAL
[14]	Address three questions concerning research in the field of energy-efficient networking.	Cellular Homogeneous Flat	1. Energy Efficiency	Offline	Centralized and Distributed	Numerical Analysis	N/A
[19]	Investigation of the tradeoff between energy saving and coverage.	Cellular Homogeneous Flat	1. QoS (Outage probability)	Online Fast reaction	Centralized and Distributed	Simulation	N/A
[21]	Adaptive cell reconfiguration based on traffic load.	Cellular Homogeneous Flat	1. Energy Efficiency 2. Coverage	Online Fast (Distr.) and Slow (Centr.)	Centralized and Distributed	Simulation	up to 50%
[54]	Proposal of a simple yet robust sleep mode algorithm.	Cellular Homogeneous Flat	1. Energy Efficiency 2. QoS (blocking probability) 3. Coverage	Both Online and Offline	Centralized and Distributed	Simulation	up to 28%
[55]	Femtocell management.	Cellular Homogeneous Flat and Multi	1. Energy savings	Online	Centralized and Distributed	Simulation	N/A
[20]	Focus on the design of base station sleep and wake-up transients, also known as cell wilting and blossoming. The results show that sleep and wake-up transients are short, lasting at most thirty seconds.	Cellular Hom. and Het. Flat	1. QoS (Switch-off time) 2. Energy Efficiency	Online Fast reaction	Distributed	Numerical Analysis, Simulation	N/A
[56]	An algorithm that reduces the power consumption of BS and reach the global optimum.	Cellular Homogeneous Flat	1. Energy Efficiency 2. Coverage	Offline	Distributed	Simulation	up to 70%
[57]	A novel energy saving procedure which allows the femtocell base station (BS) to completely switch off its radio transmissions and associated processing when not involved in an active call.	Cellular Homogeneous Flat	1. Energy Efficiency	Online Fast reaction	Distributed	Simulation	up to 37%
[58]	Minimizing the cost with a flexible tradeoff between delay and energy.	Cellular Homogeneous Flat	1. Energy Efficiency	Offline	Distributed	Simulation	up to 45%
[59]	Study the switch-off transients for one cell, investigating the amount of time necessary to implement the switch-off.	Cellular Homogeneous Flat	1. Energy Efficiency	Online Fast reaction	Distributed	Numerical Analysis, Simulation	N/A
[60]	Holistic view and comparison of different metrics and presentation of measurements.	Cellular Homogeneous Flat	1. Energy savings 2. Coverage	Online	Distributed	Simulation	around 50%
[61]	Mainly review paper focusing on real traffic data on UK networks with dynamic BS management.	Cellular Homogeneous Flat	1. Energy savings	Online	Distributed	Numerical Analysis, Simulation	N/A
[62]	Cooperation between heterogenous networks.	Cellular Heterogeneous Flat	1. Energy savings 2. Coverage	Online	Distributed	Simulation	N/A
[63]	Proposal of a QoS-aware algorithm for sleep modes.	Cellular Homogeneous Flat	1. Energy savings	Online	Distributed	Simulation	up to 90%
[64]	Derivation of a scheme In which each BS tunes its operating point without coordinating with other BSs.	Cellular Homogeneous Flat	1. Energy Efficiency	Both Offline and Online	Distributed	Numerical Analysis, Simulation	up to 20%
[65]	Performance analysis of integrated PA enhancement + smart scheduling: compares time vs. frequency alternatives	Cellular Homogeneous Flat (LTE)	1. Energy efficiency	Online Ultra-Fast	Distributed	Simulation	13 to 24%
[66]	Proposal of cell-DTX / micro-DTX to reduce the energy consumption of active cells. Analysis on a single-cell case	Cellular Homogeneous Flat (LTE)	1. Energy efficiency	Online Ultra-Fast	Distributed	Simulation	up to 61% if LTE compliant, 89% otherwise

TABLE II

SUMMARY OF POWER SAVING ALGORITHMS DEVISED FOR CELLULAR ENVIRONMENTS - CONTINUATION (ALGORITHMS WITH PSEUDO-DISTRIBUTED CONTROL SCHEME AND ALGORITHMS FOR MNO COOPERATION).

ARTICLE	MAIN CONTRIBUTION	SCOPE	METRICS	TYPE OF ALGORITHM	CONTROL SCHEME	EVALUATION METHOD	POWER SAVING POTENTIAL
[67]	Introduction of pseudo distributed management.	Cellular Homogeneous Flat	1. Energy Efficiency 2. Coverage	Both Online and Offline	Centralized, Distributed and Pseudo-distributed	Simulation	N/A
[68]	Innovative architecture that separates always-on signalling cells from on-demand data cells. Analysis of theoretical savings	Cellular Heterogeneous Flat	1. Energy efficiency	Online Fast or Ultra-Fast	Pseudo-Distributed	Analysis	up to 80% with new architecture
[69]	Cooperation between two mobile operators to reduce energy consumption based on redundant BS operation.	Cellular Homogeneous Flat	1. Energy savings	Offline	Cooperation	Simulation	N/A
[70]	Game-theory based framework for cooperation among the MNOs.	Cellular Heterogeneous Flat	1. Energy Efficiency (revenue,costs) 2. User Demand 3. QoS (capacity)	Offline	Cooperation	Numerical Analysis, Simulation	N/A
[71]	Estimation of energy savings for sleep modes which guarantee a given QoS constraint is met.	Cellular Homogeneous Flat	1. QoS (delay)	Both Online and Offline	N/A	Numerical Analysis	up to 80%

in order to eliminate the risk of creating coverage holes in the network. Of course, this increases the complexity of the algorithm, but guarantees a reliable QoS.

b) Overlapping architectures: are such that the BSs of the network might generate overlapping cells. This is a typical case of a macro-micro topology in an area with a non-uniform traffic distribution, where micro BSs are used to provide the required capacity under the coverage of umbrella macro BSs. In such topology two types of BSs are distinguished. The *critical* stations, that are usually macrocells which cannot be set into sleep mode due to the coverage issues, and the *flexible* stations that can be set into sleep mode. The flexible stations are usually microcells that are deployed under the macrocell coverage. Since there is an overlap between critical and flexible stations in the network, there is no need to increase the cell range of the BSs that remain in the *on* state. This makes the entire procedure of energy management much easier and more reliable with respect to the network QoS, when compared to the non-overlapping case. The reason is that there is no possibility to create a coverage hole in the network. The control scheme can be implemented in this case either in centralized, distributed or pseudo-distributed fashion. Due to the negligible probability of creating a coverage hole in the network, centralized control should be avoided, in order to keep complexity low. With distributed control, each BS can decide on its state of operation (on/off), independent of the conditions of neighboring cells. In the pseudo-distributed case, flexible stations are assigned to critical stations, typically using cell overlap criteria [67]. Moreover, self-organization can be also considered, as presented in [21]. The critical stations, under traffic load criteria within their cells, as well as according to the traffic conditions in the flexible cells, decide on the state of operation of the flexible stations. With the pseudo-distributed control, an online management algorithm is preferred, in order to provide smooth transitions during the on/off switch, as well as acceptable energy savings and QoS. In [68], the critical BSs are devoted to signaling, while the flexible BSs are used for data transmission, thus defining an innovative cellular architecture in which an always-on set of BSs provides full coverage and manages signaling information, while the network data plane is handled by a different set of BSs, which can be activated on demand, when and where traffic is present.

c) Implementation aspects: cell management schemes can be implemented in either slow or fast reaction procedures. Using slow reaction, the BS on/off transition is implemented in a near real-time basis, and is usually performed assuming a specific control window of the order of minutes, up to an hour. Within this time window, BSs are measuring traffic for a small portion of time, and the decision about the on or off state is implemented for the remaining portion of time. A typical example is given in [67]. Using fast reaction, the on/off management is performed at time periods of the order of seconds or even at the frame level (ultra fast). Ultra-fast online algorithms combine component-level improvements of the power amplifiers with the smart scheduling strategies exploiting such improvements. Two main approaches can be identified [65]: (1) scheduling policies adapting the actual used bandwidth, combined with power amplifiers with adaptive operating point, and (2) scheduling policies creating micro-sleep periods, using power amplifiers with deactivation mode [66, 72]. The latter technique is also known as micro- or cell-DTX (Discontinuous Transmission), because during the micro-sleep periods the BS suspends its RF transmission. In [73] the micro-DTX is enhanced with traffic shaping, to enable more frequent micro-sleep periods at the BS, while still satisfying the QoS requirements in terms of delay. It must be noted that these algorithms apply only to OFDM-based networks, such as LTE or LTE-A, and cannot be used in other networks, e.g., WCDMA-based.

2) Multi-Tier Network or offloading: Multi-tier network architectures, as defined in Section III-A, include cells of different sizes: macro-micro cells can co-exist with small cell networks [43, 55, 57, 74]. The small cell networks are usually not related to the OPEX of the mobile operator. A typical example is provided by the femtocell and/or WLAN layers. In both cases, the user can roam over the multi-tier network under user association rules or multipath TCP connections [75]. This procedure is similar to the heterogeneous case discussed in the WLAN section (see Section IV-B2). The main objective is to provide an online user association algorithm (or offloading solution) that can reach the goals related to energy efficiency and load control of the network that falls under the administrative domain of the mobile operator. One delicate aspect regards the fairness of the offloading strategy. Since the mobile network operator offloads the traffic to a femtocell or WLAN, an important issue is the "migration "of the electricity costs between the administrative domains of the operator and the owner of the small cell. Based on the current femtocell and WLAN technologies, the power consumption of small cell networks is almost flat, and independent of the traffic served at the node. For that reason, such an offloading scheme can be considered fair in terms of electricity cost "migration "and the only issue to be addressed is that of QoS at both sites. The major benefit for the mobile operator are longer BSs switch off intervals, but also the support of switch off deactivation even in case of peak traffic. This characteristic is crucial since the network can be capable of providing load (power) control and support efficient RES operation.

In case of cooperation with the femtocell layer, the main constraint is that femtocells are supposed to be in open access mode and, thus, able to absorb traffic from the cellular network. Open access can be managed by the mobile operator (since femtocells are connected to his network) or it can be left for the decision of the femtocell owner. The second issue that must be addressed is the user-to-femtocell association rule. There are two strategies to follow. The *no priority* user association rule treats BSs and femtocells as equal priority stations, and users are assigned according to the best server criterion. The second approach, called *femto priority*, assumes that users are always connected to a femtocell in case there is an adequate coverage, independently of the channel condition in the macro-micro layer. With this latter approach, more users are connected to femtocells, and thus greater savings are expected. The most practical control and coordination scheme is the pseudo-distributed case, where critical stations should be able to control the flexible stations, but also the status of the access to the femtocells that fall under their administrative domain. In that case, the mobile operator is relieved from managing a large number of femtocells. The controller of the critical stations decides when to initiate the open access at the femtocell layer, according to the external conditions that can be related to RES capacity, traffic load, electricity price, etc.

For the cooperation with a WLAN network, the situation is slightly different. The mobile operator does not usually have the administrative rights to the WLAN APs. The cooperation between the macro-micro cellular network and WLAN APs is a user association rule in multi-Radio Access Technology (multi-RAT) case. To support such connectivity, mobile devices that are capable of using both WLAN and cellular technologies can automatically, or under the mobile operator command flow, migrate data to the preferred network. Furthermore, with the development of multipath-TCP techniques, it is feasible to transmit data in both networks, equally or unequally weighted along the two paths (WLAN and cellular), provided there are no severe differences between paths characteristics. Similar to the femtocell case, offloading is achieved at the administrative domain of the mobile operator with a negligible increase of electricity costs for WLAN owners. Regarding the control scheme, if the mobile operator is willing to manage the data route, then it should broadcast it to the mobile terminals that can set the preferred RAT. Following a similar procedure with the femtocell layer cooperation, *other RAT priority* or *no priority* procedures should be implemented, so as to decide the user-to-technology association. For practical implementation of such solutions, online implementations of offloading algorithms under pseudo-distributed control schemes seem to be the most promising ones, keeping complexity low, but also providing energy savings and holding QoS above the given thresholds.

Multi-tier networks can provide more degrees of freedom to the mobile network operator (MNO) to migrate traffic when necessary and reduce power consumption of the network. This can be very useful when the network is powered by Renewable Energy Sources (RES), as they provide a specific time variant power capacity in the network. In this context, offloading techniques allow BS management even on the occasion of high traffic periods. In other words, multi-tier networks and offloading can not only significantly reduce the energy consumption of the network, they can also provide the ability to adapt the peak power of the network to time variant RES capacity [76].

3) Mobile Operator Cooperation: With the objective of reducing their CAPEX, and possibly also their OPEX, in particular the portion spent to power their infrastructure, MNOs have recently introduced the concept of *network sharing*. The main idea is that MNOs should cooperate and share their infrastructures, including their approaches for implementing sleep modes, in order to adapt the active capacity to the current traffic needs, and thus save energy.

Consider an area served by n MNOs, which operate separate networks. As it was observed in Section II, due to the end user behavior (i.e., the combination of user activity and mobility patterns), traffic fluctuates significantly during a day. Thus, a network which is dimensioned to meet a given QoS constraint at the peak traffic load, offers a capacity which is underutilized for the long periods of time, during which traffic is lower (and possibly, much lower) than the peak value. Since n MNOs coexist in the same service area, the underutilization of the access networks capacity occurs for all access networks at roughly the same time, due to the similar average customer behavior. In terms of consumed energy, most of networking devices, including BSs, consume about the same quantity of energy regardless of the amount of carried traffic; thus, the network consumes about the same amount of energy at the full load (under the peak traffic) as when it is underutilized.

The network sharing approach allows MNOs to take advantage of this situation and save energy, by modulating the active capacity to follow the traffic demand. The key idea underlying the energy efficiency of network sharing is that network capacity supply can be modulated by switching off some networks for the time periods in which traffic is low over the service area, so that a subset of the access networks is sufficient to provide the capacity necessary to achieve the desired QoS. Of course, while the network of an MNO is off, its customers must be allowed to roam to the networks of the MNOs that are active. The core of the scheme is then the decision of the switch-off pattern, meaning the sequence according to which the networks are switched off, together with the switch-off and switch-on instants. Switch-off and switch-on instants can be rather accurately determined by the analysis of the historical traffic traces, which exhibit a remarkable periodicity, adding margins to account for both unpredictable local traffic variations, and transient delays. The decision of the switch-off patterns can be taken targeting various objectives. For example, one objective could be to balance the roaming traffic, i.e., to establish that the MNOs switch-off their network at such a frequency that, on average, the MNOs carry more or less the same amount of roaming traffic from the other MNOs that switch off their networks. Another possibility could be to balance the average number of times that a network switches off in a given period, or the estimated energy saving. A positive side-effect of network sharing is that, considering globally the amount of deployed resources, active resources are more effectively used than in traditional scenarios without sharing. Indeed, network sharing aims at reducing energy wastage that derives from daily periods of over-provisioning by making the available capacity more closely follow the traffic profile.

In [69] it was estimated that in scenarios like those of the European countries with the largest networks, network sharing can lead to saving that ranges from 30 to 40%. Despite the great energy (and cost) saving that can be achieved in operating the networks, many difficulties still exist in the path to network sharing, that relate to both operational problems and commercial sensitivity of information. The first, and probably one of the major difficulties, is that MNOs are reluctant to allow their subscribers to roam through their competitor's network. The competitor MNOs might capture users' profiles and try to attract them with suitable offers. Similarly, MNOs are concerned about QoS and the fact that the other operators might not offer the same degree of QoS to their users. This is particularly true for the operators with a dominant position in the market. Moreover, the difference in the QoS levels adopted by MNOs in terms of both performance and coverage is often used as a service differentiator to attract customers. Other difficulties are related to the high initial cost incurred to allow the sharing of the networks, due to the complexity of the control of several parallel networks operated as a pool; and, the need for extended roaming and billing procedures to allow the seamless transfer of services from one network to another, and to share revenues between the involved MNOs.

Furthermore, a game-theory based framework based on the cooperation of the network operators has been suggested in [70]. In this work, several techniques to allocate the cost and the benefits of the cooperation, e.g., power saving, are presented. Nevertheless, the main conclusion drawn from the analysis is that each allocation method is tailored for a given scenario and cannot be generalized in other scenarios.

B. Dense WLANs

For dense WLANs, two different use cases are discussed: homogeneous and heterogeneous application scenarios. All identified articles related to WLAN are shown in Table III and are also sorted according to the control scheme applied.

1) Homogeneous scenarios: Strategies devised for the homogeneous scenarios rely entirely on the WLAN technology that is supposed to provide the coverage required for the Internet connectivity over the entire area in question. In contrast to the strategies designed for heterogeneous networks, these strategies have less complexity in terms of deployment, control and management, due to their purely-WLAN-based nature. These characteristics can be beneficial to define a subset of WLAN APs that would suffice to provide the coverage during the marginal (low) traffic conditions, whereas additional APs would be only switched on upon detecting an increase in the user demand.

To this end, authors of [77] suggest that APs could form *clusters*, based on the Euclidean distance between them. Within a cluster, WLAN APs can communicate with each other with a given RSSI (received signal strength indicator). It is assumed that due to the close distance between neighboring APs, which is the case in dense WLAN deployments, only one AP from each cluster, called a *cluster-head*, is switched on to provide the coverage and required capacity for the entire cluster. When the number of users associated with the cluster-head exceeds the maximum allowed number of associated users per AP, all the APs in the cluster will be simultaneously powered on to provide the additional capacity. The proposed strategy relies on the online-measured number of associated users per AP, collected and processed by the central controller. The decision on the status of each AP (on/off) is also made and executed at the central controller and the algorithm pertains to the fast reacting solutions. It has been reported that using this strategy would yield a 20 % to 50 % power saving in less dense scenarios, whereas in more dense WLANs the total power saving would increase to 50 % to 80 %. The main limitation of this simple approach is owed to an incorrect cluster definition. Using just the Euclidean distance to form the clusters may lead to a situation, in which the coverage is not provided over the entire WLAN area. This is due to the non-ideal channel conditions, such as presence of interference and fading, that may cause two, even adjacent, APs to have significantly different performance, and consequently resulting in coverage holes appearing between clusters. To address this issue this work has been extended in [38], where the discovery of the neighboring APs is based on number and RSSI of

TABLE III

Summary of WLAN power saving algorithms.

Article	Main contribution	Scope	Metrics	Type of algorithm	Control scheme	Evaluation method	Power saving potential
[27]	A cluster-based power saving scheme applicable in dense WLANs.	WLAN Homogeneous	1. User demand (aggregated traffic, num.assoc.users)	Offline	Centralized	Numerical analysis	up to 41%
[38]	Traffic measurements (once a day) are used to ensure the coverage and sufficient bandwidth. The neighboring APs can increase their transmission power to strengthen their coverage.	WLAN Homogeneous	1. Coverage 2. User demand (Channel busy time)	Online Fast reaction	Centralized	Real network	53% for the low traffic case and 16% in the high traffic case
[77]	Based on Euclidean distance between APs, the cluster of APs is made. In each cluster, only one of the AP is powered on to provide the coverage and capacity.	WLAN Homogeneous	1. Coverage 2. User demand	Online Fast reaction	Centralized	Simulation	20-50% in less dense WLANs 50-80% in dense WLANs
[78]	Group of closely placed APs forms cluster. Either all APs or all but one AP in the cluster are turned off, without shrinking the coverage.	WLAN Homogeneous	1. Coverage 2. User demand (num.assoc.users)	Online Fast reaction	Centralized	Simulation	up to 60%
[79]	Optimized management and power control of the APs according to temporal dependency of traffic.	WLAN Homogeneous	1. User demand (aggregated traffic, num.assoc.users)	Offline	Centralized	Numerical analysis	up to 63%
[80]	Heuristic algorithm based on greedy methods and local search to deal with high computational complexity of ILP.	WLAN Homogeneous	1. User demand (aggregated traffic, num.assoc.users)	Offline	Centralized	Numerical analysis	N/A
[81]	Central scheduler based on maximum coverage problem	WLAN Homogeneous	1. Coverage 2. User demand (num.assoc.users)	Online Fast reaction	Centralized	Numerical analysis	up to 80%
[82]	Paging in cellular network is used to detect the user presence in WLAN.	WLAN Heterogeneous	1. User demand (aggregated traffic)	Online Fast reaction	Centralized	Numerical analysis and Simulation	N/A
[83]	Optimization framework for network management based on temporal dependency of traffic.	WLAN Homogeneous	1. User demand (aggregated traffic, num.assoc.users)	Offline	Distributed	Numerical analysis	up to 40%
[84–86]	An additional wake-up receiver activates the powered-off APs on-demand. Wakeup ID and protocol designed based on the ESSID of a WLAN.	WLAN Homogeneous	1. User demand	Online Fast reaction	Distributed	Testbed, Numerical analysis and Simulation	N/A
[87, 88]	Aggressive switching off approach based on user presence detection.	WLAN Homogeneous	1. Coverage 2. QoS (delay, prob. of user presence discovery)	Online Fast reaction	Distributed	Numerical analysis	up to 98%
[89]	Using Bluetooth to meet marginal demand, whereas the WLAN is activated to provide the higher capacity	WLAN Heterogeneous	1. User demand (aggregated traffic)	Online Fast reaction	Distributed	Testbed and Simulation	48%
[90]	IEEE 802.15.4 is used as a supporting technology to power the APs on/off.	WLAN Heterogeneous	1. User demand	Online Fast reaction	Distributed	Real network	up to 91%

received beacons instead of Euclidean distance. For any two APs, if the number and RSSI of the received beacons exceed pre-defined thresholds, it is assumed that the APs are located in close proximity, and thus can form a cluster.

Another drawback of the simple strategy proposed in [77] is the overly simplistic metric of the user demand. The number of users associated with an AP may not reflect the real traffic needs of different applications. To overcome that, [38] proposes channel-busy fraction, i.e., the percentage of time the channel is busy due to the transmission and inter-frame spacing, as a more reliable indicator of the user demand. Similar to [77], information about channel busy fraction is collected from the APs by the central controller to make and execute decisions about status of the APs. Based on the measured channel busy fraction, if the activity of the users in a cluster is marginal, only the cluster head remains in operation. On the other hand, when the channel busy fraction exceeds a given threshold, additional AP(s) belonging to the cluster are switched on to offer the required capacity. It is reported that up to 53% and 16% power saving can be achieved for the low and high traffic conditions, respectively.

Another cluster-based approach has been proposed in [78].

It aims at switching off the inactive APs to the extent that in some clusters, even all the APs are switched off, when the coverage area of a cluster can be covered by neighboring clusters. Similar to [38], the clusters are built based on the measurement results collected and processed by a central controller. The user demand is indicated by the number of users associated with an AP. If it exceeds a given threshold, additional AP(s) are switched on to provide the required capacity. The time-scale of the reaction is fast. It has been shown that power consumption can be reduced by 60%, when the proposed approach is used. However, all aforementioned cluster-based strategies have a common shortcoming: the frequent online measurements may degrade the user performance.

Conversely, the cluster-based on/off switching strategy devised in [27] needs no online measurement and is based on a usage pattern, derived from offline-measured and processed information about the user demand. The number of associated users and the bandwidth per connection are used to reflect the user demand for the capacity. It is assumed that all potential users have full coverage. It is worth noting here that without any assessment of the coverage, such an assumption seems fairly unrealistic. The required capacity is estimated by means of the usage pattern that is fed into two continuous-time Markov chain (CTMC) models of a cluster of APs to calculate the required number of powered-on APs. Using this approach, it has been shown that 40% of the power can be saved. Rather than using a model for clustering the APs, in [79] an ILP (Integer Linear Program) optimization model is developed to scale up/down the density of the powered-on APs. Under the assumption that possible positions of the users are known (for a given prediction period), the coverage is provided only over these positions by switching on a subset of APs. Again, similar to [27], the coverage provisioning problem is not considered in this study. In addition to this subset, redundant APs are switched on to meet the user demand for the capacity indicated by the number of users associated with an AP and traffic generated by them, known by the central controller. The proposed approach offers up to 63% power saving. In another effort, Lorincz et al. present a heuristic approach in [80] to reduce the computational complexity of the optimization algorithm introduced in [79]. The usage pattern is approximated by a discrete function fed into the optimization algorithm. By applying the proposed heuristic algorithm, composed of a greedy approach finding a feasible solution and a local search iteratively moving to the best solution, the computation time is reduced considerably. However, the power saving is reduced by 10% in comparison to [80]. The approach proposed in [79] is further modified to be applied also in a decentralized system, as shown in [83].

An interesting power saving scheme has been proposed in [81]. Its design has been based on the maximum coverage problem, which is a maximization of the set cover problem (see [91] for more information). The proposed algorithm runs on a central controller, which collects the number of associated users with each AP and their data rates, and switches APs on and off dynamically (in the fast reaction scale), while maintaining the coverage and guaranteeing user performance. Although 80% power saving is reported by this strategy, the definition of the cost function of the algorithm may result in frequent handoffs and excessive delays. The cost function is defined using the measured data rate of the user terminals that mainly depends on channel conditions. Any change in the measurement values leads to a network reconfiguration, handoff, and consequently delay.

In the approaches proposed in [84], its modified [85] and extended version [86], all inactive APs can be switched off regardless of the coverage. To react to user presence, an auxiliary low-power (no more than 1 mW) wake-up receiver is attached to the AP, in order to switch it on, when the special wake-up signal is received. In order to reduce the probability of false negatives (the user presence is not discovered by the auxiliary receiver) and false positives (the AP is waken up unnecessarily), rigorous theoretical and experimental studies are performed.

An alternative approach to the power wastage problem in dense WLANs was recently presented in [87] (further results in [88]). Authors claim that the density of the WLAN APs can be reduced drastically, to the extent that the APs remaining in the operation will only provide the coverage required to discover the user presence. They make the following observation: to detect the user presence (with a given probability) it is actually sufficient that just one out of several Probe Request frames transmitted with the lowest bit rate is received within a desired delay. Once the user presence is discovered, additional APs can be switched on to provide the additional capacity required by the user. One of the main characteristics of the proposed algorithm is that the status of each AP can be determined in a distributed fashion. By conducting numerical evaluations, it is demonstrated that up to 98% of inactive APs can be switched off by means of the proposed online fast-reacting algorithm. To even further improve the possible power saving potential of this algorithm, the impact of change in the coverage percentage, probability of user detection and acceptable user delay is also studied.

To summarize, homogeneous strategies have little complexity in terms of deployment, control and management, but their main drawback is that they are not transparent to the users, meaning that the users may experience a slight performance degradation (e.g., delay) during on/off switching phases. Furthermore, they require software and hardware modifications from the operator side. And last but not least, these solutions cannot reduce the number of inactive APs as aggressively as the strategies that involve additional radio technologies that are described next.

2) Heterogeneous scenarios: In case there is more than one radio technology available, i.e., the target area is covered by a WLAN and at least one more radio network, e.g., Bluetooth, sensor, cellular network, etc., there is a possibility to deploy heterogeneous strategies to save energy. All the algorithms presented here are based on online measurement and the time scale for the reaction is fast. Hence, when the demand for capacity is marginal, all WLAN-APs can be aggressively switched off to avoid power wastage. In this manner, the coverage and the required (low) capacity are provided by the other co-existing network(s). Among these technologies, cellular is the most commonly used, due to its widespread deployment. For instance, in [82], it is assumed that the mobile station has all the time the Internet connectivity via its cellular interface and its position is roughly known. When there is no user demand for the WLAN capacity, all the APs can be switched off, and if the user demand increases, a WLAN AP placed in the vicinity of the mobile station will be switched on. In this approach, the coverage and the capacity are entirely provided by the cellular network and the control scheme can therefore be seen as a cooperation between the WLAN and the cellular operator (as already discussed in Section IV-A3). It has been reported that up to 50% of the power consumed by the WLAN can be saved by applying this strategy.

Another strategy, devised in [89], presents an example of cooperation between WLAN and Bluetooth. In this case, both mobile stations and APs have to support the Bluetooth technology. It is assumed that the user mobile device can provide multi-hop connectivity, via its Bluetooth interface, to the APs equipped with Bluetooth radios. In this way, the capacity required for low-bit-rate applications is provided by this multi-hop Bluetooth connectivity, and either an increase in the user demand or a degradation of the quality of the Bluetooth link will lead to the activation of the WLAN radio. Due to the limited coverage range of the Bluetooth technology (typically 10 m), the Bluetooth-coverage and capacity can be offered only to the mobile stations located in the proximity of the multi-radio APs. It has been demonstrated that up to 48% power saving can be achieved when this strategy is used.

Authors of [90] propose a strategy in which the IEEE 802.15.4 narrow-band radios are used to discover the user presence. The WLAN coverage and capacity is solely provided when the user presence can be discovered by receiving the active scanning packets from mobile stations via the IEEE 802.15.4 interface. Due to the fact that IEEE 802.15.4 narrow-band radio consumes less power than the WLAN AP, the power saving is dramatically high, of the order of 91%. Such a huge power saving is achieved at the expense of numerous false positives (detection of a non-WLAN transmission) that can be explained by the fact that IEEE 802.15.4 narrow-band radios are not WLAN-technology-selective (any radiation in that frequency range will be treated as WLAN transmission).

Summarizing, strategies developed for heterogeneous scenarios have potential to provide huge power savings. The main requirements limiting their application lie in the availability of multiple radio technologies and interfaces at both mobile stations and WLAN APs, as well as in providing the signaling between different technologies. This latter limitation imposes further requirements on operators and standardization bodies to develop a framework for such an operation. Furthermore, the delay, caused by the handoff between two technologies, and threshold of the capacity, at which such a handoff should occur, should be carefully studied in order not to degrade the user performance.

C. Comparison of Techniques

Looking more closely at Tables II and III that summarize the BS/AP management strategies identified in the literature, we can draw the following conclusions. As for cellular networks, the majority of the algorithms have been devised for technologically homogeneous scenarios, with only few proposals that take into account different types of access network. The main limiting factor is the fact that the different access technologies usually belong to different operators, and thus sharing electricity costs (and users) is a real challenge due to the lack of interest from the MNO side in sharing commercially sensitive information. This is not the case for WLANs, where having multiple radio technologies permits more aggressive energy savings. Nevertheless, also for WLANs, homogeneous proposals seem to be more common due to the decreased complexity, despite the potential problems with coverage provisioning.

In terms of metrics that are used to evaluate whether there are enough radios powered on, the majority of the algorithms for cellular networks is based on a combination of various metrics, with the most popular being QoS assessment and energy-efficiency. Furthermore, in terms of how optimization is performed, online algorithms seem to be more common than offline, however no clear trend can be identified. This can be explained by the fact that the choice between online and offline algorithms essentially depends on the scenario where the power saving strategy is applied. In a scenario where the user demand can be estimated based on a pre-defined schedule, an offline algorithm appears to be more suitable due to its low processing overhead. On the other hand, QoS can be guaranteed even in case of an unexpected variation in the user demand, when an online algorithm with higher processing overhead is applied (see Sec. III-C). This is also the case for WLAN. As for the metrics, user demand seems the most common choice for WLAN strategies, mainly due to the simplicity in obtaining this data.

As for the applied control scheme, the strategies with centralized control outnumber distributed control strategies for both access network types. This can be related to the evolution of centrally-managed networks, stimulated by the need for more simplified configuration and management schemes.

Not surprisingly, for both wireless access networks, the vast majority of the evaluation is done by means of simulation or numerical analysis. This is mainly caused by the relative difficulty of the current hardware to support the proposed mechanisms as well as by deployment restrictions that pose additional constraints reducing thus the expected energy gains. Possible enablers that would help to overcome this problem are discussed in more detail in Section V.

The numbers for the estimated energy savings reported in the literature refer exclusively to the specific algorithms, which differ greatly in assumptions relative to traffic, in configuration of the wireless network, and by the way in which energy efficiency is measured. As such they are hardly comparable, and they do not give a precise idea of what are the possibilities and the limitations of the proposed algorithms. Therefore in Section VI, we propose a performance assessment that would help to objectively evaluate the expected potential of the on/off switching algorithms.

V. PRACTICAL ASPECTS OF IMPLEMENTATION OF MANAGEMENT SCHEMES

As stated in Section IV-C very few of the presented BS management strategies are actually implemented in real networks. Therefore, it is interesting to discuss possible enablers for equipment and networks that would help to deal with the numerous constraints and limitations reducing the expected energy gains (or in the worst case, completely preventing commercial deployment). The major constraints come from standards rigidities, deployment restrictions and architectural aspects of network equipments, in particular BSs and APs.

A. Standards

Standards are defined by well-known standardization fora (mainly 3GPP, IETF and IEEE) and detail how wireless systems should work in order to allow inter-operation of equipment (BS or APs, mobile terminals, core networks) developed by different vendors, under precise constraints on network performance. All new BS management proposals need to smoothly integrate within the existing standards, in order to bring effective energy gains into actual networks. Standard modifications are possible, but at the price of a high standardization effort and longer time-to-products.

For modern cellular networks (e.g., WCDMA/UMTS, LTE, LTE-Advanced) 3GPP is the main standardization forum. 3GPP already started working on energy savings aspects, and some initial technical specifications have been approved. Until now, the majority of the work focused on solutions at network management layer, between the radio access (BSs) and the OA&M system. Of particular interest for the BS management algorithms are the modifications defined in [92, 93], which introduce the basic signaling to switch-on/off a BS via its backhaul interface, for UMTS (using the I_{ub} interface between the NodeB and the RNC) and LTE (using the X2 interface between the peer eNodeBs), respectively. It must be noted however that the solution standardized on X2 does not contain any explicit switch-off command: the decision to turn-off (or enter some stand-by mode) is expected to be taken autonomously by the concerned eNB. BS management solutions needing an explicit switch-off signal are not natively supported through such interfaces. However they may exploit Network Management System (OA&M) commands (largely available) to request a BS to turn-off, e.g., for maintenance reasons.

BS management algorithms exploiting the presence of multiple carriers on a given BS (e.g., by turning off some unnecessary carrier at low load) can be easily integrated into existing equipments, because the controlled hand-over (or relocation) of mobiles from one carrier to another is a feature supported by all cellular standards (since GSM). And carrier switch-off does not require any standard specification. In a similar way, BS on/off algorithms can exploit the Carrier Aggregation feature of LTE-Advanced [94, 95]. In this case the on/off algorithm applies to the different Carrier Components.

X2 natively supports the exchange between peer LTE eNBs of cell load information [93]. This is useful for the implementation of some of the algorithms presented before (e.g. online, co-operation) when applied to an LTE or LTE-Advanced network.

Inter-RAT scenarios (i.e. interworking between 3GPP technologies: LTE-A, LTE, GSM and UMTS layers) are also partially supported, as inter-RAT mobility is a native feature between 3GPP technologies. However, 3GPP is still working in defining the exchange of load and energy-related information (e.g. switch-on/off commands) between different RATs [96]. 3GPP has also recently initiated to work to a tighter integration of WLAN and 3GPP access networks, but the finalization of such specifications is still uncertain and not expected before long time. In the meantime, the best opportunity to “link”the two worlds come from the ANDSF (Access Network Discovery and Selection Function) specified in [97]. This function is used to inform a mobile terminal about the presence of alternative coverage technologies in a given place (e.g. a WLAN hotspot). It could be extended to integrate some energy-saving information and exploited to turn-on/off alternative coverage cells according to the presence/absence of users.

Some BS Management algorithms need or may benefit of modifications of the air interface (i.e. the physical layer). For example, the reduction of emitted pilots (or of the periodic signaling) may allow longer “off”periods (and thus higher energy savings) in algorithms like e.g. micro-DTX [66], that work at symbol, sub-frame or frame time-scale (between $100\,\mu$s and $10\,$ms). Unfortunately, from the standardization point of view the modification of the physical layer is very difficult to achieve. In fact, the need to be able to serve legacy mobiles (i.e., compliant to former 3GPP releases) significantly limits the possibility of introducing new energy-saving functions on the air-interface. Attempts to reduce the number of emitted pilots, e.g., [98], have been rejected in the past. Currently, the best opportunity comes from the MBMS-Blanking technique [96], which consists in substituting a regular empty LTE sub-frame (i.e. not carrying any user data) with an MBMS (Multimedia Broadcast Multicast Services) one, containing fewer pilots and signaling, with an estimated energy reduction of around 45%. Similarly, Almost-Blank Sub-frames (ABS, defined for LTE-Advanced in the scope of eICIC feature [99]) can be used to reduce the amount of transmitted control signaling (including pilots) in specific sub-frames and desynchronize the emissions of interfering cells.

3GPP is still working on energy saving features, and new mechanisms can be expected to come in the near future. Concerning the PHY level, the new work item New Carrier Type offers a concrete opportunity to change the air-interface structure (e.g., pilots reduction, signaling minimization) on such new carriers.

In the context of WLANs, the entire standardization work has been recently summarized in the most recent draft of the IEEE 802.11 standard (IEEE 802.11-2012) [100]. A good overview of the IEEE 802.11 standard structure is provided in [101], for readers interested in further details. This newest version of the IEEE 802.11 standard incorporated the following amendments that are important in the context of energy-efficient operation of WLANs: IEEE 802.11e (QoS-aware MAC) and IEEE 802.11h (spectrum management in

5 GHz band), both already included in the IEEE 802.11-2007 standard, as well as IEEE 802.11k (radio resource management enhancements) and IEEE 802.11v (wireless network management extension). Especially the last two amendments are of the utmost importance in the context of the development of the power management schemes discussed in this article, providing means for the collection of the detailed information about the network state by WLAN APs.

B. Commercial Support

Manufacturers are beginning to develop commercial software/hardware solutions for the management of the energy consumption of network components, including WLAN APs, femto BSs, and possibly longer range equipment. For instance, the Cisco EnergyWise software includes an intelligent energy management solution that can be used for monitoring, controlling, and reporting the energy usage of information technology and facilities equipment. This product is able to administer the energy requirements of new and existing Power over Ethernet (PoE, IEEE 802.3 at/af) devices; to extend power management to desktop and laptop PCs, IP phones, etc.; and to manage energy-saving efforts through interactive GUIs. Specifically in the access networks, the Cisco EnergyWise protocol has been implemented on new-model Cisco switches, e.g., the Catalyst 4500 series, for managing the power consumption of WLAN access points and femtocells connected to them. With the EnergyWise protocol, Cisco Switches are able to set different power levels for the devices connected to them by PoE interfaces. The EnergyWise protocol has the capability of setting 11 different power levels, based on the premise that the connected WLAN access points or femto BSs support multiple operating modes (e.g., full power, standby, sleep, off, etc.). For the connected devices that do not have multiple working modes, the protocol is still applicable to directly switch them on/off. The EnergyWise protocol is also functional to schedule different power levels to the connected devices at different times of the day, and different days of the week, and then make the device to follow this schedule recurrently [102].

C. Deployment in Real Environments

BS Management solutions are also impacted by the real network architectures and deployments cases. The most important limitation impacts algorithms aiming at turning-off base stations in single-layer deployments. In fact it is extremely difficult to ensure that such action will not create a coverage hole somewhere in the area surrounding the concerned cell (e.g. in deep indoor). That applies even in case of home femtocells and access points, that may be responsible for providing coverage (and not only capacity) in some specific places (e.g., basements). Such risk is unacceptable by mobile operators, and this drastically reduces the chances to see such algorithms deployed in commercial networks. On the contrary, algorithms exploiting the co-existence of several layers are expected to be more and more pertinent in future networks, as the HetNet (Heterogeneous Networks, i.e., an umbrella macrocell with several smaller cells under its coverage), and multi-RAT approaches gain momentum, as a valid way to increase the mobile network capacity.

D. Integration on Existing Equipment

Finally, BS Management algorithms need to take into account real equipment hardware and software limitations. The most important limitations concern the ability of a BS to actually enter into some energy-saving modes, the amount of saving achievable by these modes, and the transient time for moving from these modes to a fully operational state (active state). Currently deployed BSs (of any technology) do not implement such power-states, and can usually only work in two states: fully active or completely switched off (except for a backhaul network interface that remains active to receive switch-on commands).

Algorithms based on long time-scales (i.e., tolerating off-on transients in the order of minutes) are not affected by this problem: they can be implemented by completely turning off a current base station. This can be achieved even on already deployed base stations, using the OA&M interface. Algorithms requiring shorter wake-up delays cannot be based on a complete switch-off of the base station. In fact, the delay necessary to move from a power-saving mode to the fully active mode depends on the number (and type) of components that are switched off (see, e.g., [103] for an analysis on 3G femtocell, and [104] for WLAN APs). A cold-restart requires several minutes to complete, with the delay increasing with the size of the base station. Modern base station power profiles (e.g., [105]) usually include such limitations and should be used when validating algorithm performance. When considering faster algorithms, it can be noted that the new generation of power amplifiers (implementing dynamic gating and drain voltage control) will enable symbol-time-scale algorithms (tens to hundred of μs) to work correctly and efficiently.

Improvement of PA efficiency is always under the research spot, resulting in continuous improvement of performances [106]. The current state of art for high-power BSs (macros) is represented by Doherty Power Amplifiers (DPA), which contain one main (carrier) amplifier always active and an auxiliary one (peaking) active only when signal peaks occur [107, 108]. This aspect makes DPA particularly adapted to signals with high PAPR (like in LTE). DPAs show an efficiency of around 40 %. PA efficiency can be further improved by two new emerging concepts: Envelope Tracking PAs (ETPA, [109]) and Switch-Mode Power Amplifier (SMPA, [110]). In ETPA the supply voltage is constantly varied to track the fluctuations of the transmitted input. ETPAs allow to drastically reduce the heat dissipation, and thus, to increase the energy efficiency to around 50-55 %. ETPA technology is well suited for low power systems (small cells) where the instantaneous required pass band is lower. The Switch-Mode Power Amplifier (SMPA) architecture promises even higher gains. In SMPA the active component (transistor) is operated in an *on/off* way: when *on*, the transistor acts as a very low resistance, when *off* it acts as an open circuit. Since the transistor shows a virtually zero *on* resistance (thus, virtually no voltage) and an infinite *off* resistance (thus, no current), a close to 100 % efficiency is theoretically achievable. In reality, efficiencies of around 60 % are expected.

Additional attention must be brought to the internal hardware architecture of a BS. For example, modern multi-carrier and multi-technology BSs (recently hitting the market) tend to use the same RF head for several carriers or several technologies (e.g., one RF head for one sector, offering two 3G carriers and one LTE carrier). In that case, algorithms based on carrier or RAT switch-off do not bring significant savings, as the RF transceiver and power amplifier must remain active for serving the remaining carriers/technologies.

E. Issues to Be Addressed for Future Equipment

When looking to the predictions on technology, equipment and components evolution, it is possible to identify some risks of limiting the effectiveness of BS Management solutions. For example, the tendency to integrate more and more components onto a single chip, such as in the case of Radio-on-Chip (RoC), Digital-on-Chip (DoC) and System-on-Chip (SoC), presents the risk of limiting the flexibility in turning off some of such components. In fact, while allowing more efficient circuits, the integration of several components implies the inability to turn off some of them, without turning off all others. This may results on components that can be no longer turned off for energy saving. Similar problems come from the centralization of the baseband processing, like in the BBU pooling concept. In such cases, the granularity of a BBU component can be the one of several cells, thus making more difficult to turn off some of the cells in a fine-tuned way. The system impact of longer term PA architecture evolutions, like ETPA, [109] and SMPA [110], requires additional analysis.

VI. PERFORMANCE ASSESSMENT (LONG TERM SAVINGS)

In this section, we estimate the potential for energy savings of sleep modes, both in cellular networks and in WLANs. Underlying the whole research on sleep modes is the widespread belief that the savings achievable are in general high (say 30% or more, as reported in Tables II-III), and therefore interesting enough for network operators, especially considering the ever increasing energy prices. This belief came from the general observation that cellular and WLAN networks rarely operate at their full capacity, and therefore techniques which make their consumed energy proportional to the traffic load could bring substantial savings.

However, estimating the actual energy saving potential of sleep modes is not trivial, given the great variety of the proposed approaches that were presented in Section IV. As stated in Section IV-C the results reported in the literature are hardly comparable, and they do not give a precise idea of what are the possibilities and the limitations of sleep modes. Therefore, in this section we estimate the energy savings that can be achieved in cellular and in WLAN access networks by using sleep modes in periods of low traffic loads, through the determination of the energy-optimal BS/AP densities as a function of user density. By taking into account the QoS perceived by end users, we derive realistic estimates that can be used to evaluate the effectiveness of sleep modes. The analysis can be applied to several BS/AP configurations, and to different energy models. This analysis has been mainly developed in [71].

We consider a system in which users form a homogeneous planar Poisson point process, with the intensity λ_u users per km^2, while BSs/APs form a planar point process, with the density λ_b BSs/APs per km^2. While the methodology introduced here is quite general, and can be extended to many different BSs/APs configurations, we restrict ourselves to Poisson layouts, as they model with good accuracy the effects of real life constraints on the BS/AP locations [71]. We assume that all BSs/APs densities are feasible. Indeed, in the homogeneous Poisson process layout of BSs/APs, if each BS/AP independently makes a decision to either turn off, or stay on, according to some probability, the resulting point process of BS/AP is a thinned homogeneous Poisson process, and all BS/AP densities are indeed achievable. But for other BS/AP layouts (e.g. Manhattan, hexagonal) this is not the case, and our analysis brings to energy savings estimates which are more optimistic. For cellular networks, our analysis considers downlink communications. The end user performance metric that we use is the *per-bit delay* τ of best effort data transfers, defined as the inverse of the user throughput, i.e., the actual rate at which the user is served, taking into account the capacity to the user as well as the sharing of BS/AP time across all associated users. The performance constraint that is enforced is as follows: if the per-bit delay experienced by a *typical* user, $\bar{\tau}$, is less than a predefined threshold $\bar{\tau}^0$ seconds, then users are said to perceive satisfactory performance, and the corresponding BS/AP distribution is feasible. Here we do not consider the effect of shadowing and only take into account distance-dependent path loss. We assume that users are served by the BS/AP that is closest to them, i.e., by the one that corresponds to the strongest received signal, as it normally happens in reality. The capacity can be modeled, for example, using Shannon's capacity law or other models such as a quantized set of achievable rates. Here we assume the network only serves best-effort traffic. For cellular networks, we assume that BSs use a processor sharing mechanism to divide capacity among all the connected best-effort users. By doing so, a notion of fairness is imposed, since all best effort users associated with a particular base station are served for an identical fraction of time. We assume that BSs/APs always transmit at a fixed transmit power. When the BS/AP density is higher than that required to achieve the threshold expected per-bit delay $\bar{\tau}^0$, we assume that base stations only serve users for the fraction of time required to satisfy the performance constraint, and remain idle (i.e., not transmitting to any user) for the rest.

We denote with U the utilization of BS/AP, i.e., U is the average fraction of time in which the BS/AP is transmitting. We model the power consumed by a BS/AP as $k_1 + k_2U$, where k_1 is the power consumed by keeping the device on with no traffic, and k_2 is the rate at which the power consumed by the BS/AP increases with the utilization. The first energy model that we study reflects the current hardware design, and assumes that the bulk of the energy consumption at BS/AP is accounted for by just staying on, while the contribution to energy consumption due to utilization is negligible (i.e.,

$k_2 = 0$). We also study energy consumption models with k_1 and k_2 chosen to reflect a more energy-proportional scenario, i.e., $k_1 << k_2$.

We characterize the per-bit delay perceived by a typical best-effort user who is just beginning service, as a function of the density of users and of BS/AP. For cellular networks, the average per-bit delay $\bar{\tau}$ perceived by a typical best-effort user joining the system when the density of BS/AP is λ_b and the density of users is λ_u, is given by:

$$\bar{\tau}_P = \int_0^\infty \left(\int_0^\infty \int_0^{2\pi} e^{-\lambda_b A(r,x,\theta)} \lambda_u x d\theta dx \right) \frac{e^{-\lambda_b \pi r^2} \lambda_b 2\pi r}{C(r)} dr \qquad (4)$$

where $A(r,x,\theta)$ is the area of the circle centered at (x,θ) with radius x that is not overlapped by the circle centered at $(0,-r)$ with radius r, and $C(r)$ is the capacity to a user at distance r from the base station at $(0,-r)$ [71]. For WLANs, the expected per-bit delay is computed as the inverse of the user throughput, computed with the well known Bianchi's formula [111], assuming that all hosts and access points have always data to transmit (saturation condition). In the case of the energy model with $k_2 = 0$, energy consumption is minimized by using the lowest BS/AP density that can achieve the desired user performance. Given λ_u and λ_b, the per-bit delay perceived by a typical user can be evaluated using Eq. (4). $E^0\left[\bar{\tau}\right]$ is decreasing in λ_b. Thus, we can set the expressions equal to the target per-bit delay, $\bar{\tau}^0$, to determine the minimum required BS/AP density λ_b^*.

When $k_1 << k_2$, the utilization of the BSs/APs in the network plays a key role in determining the energy consumed. Again, $\bar{\tau}$ can be evaluated given λ_u and λ_b using Eq. (4). In this case, it is easy to see that the desired user performance can be achieved by the BSs/APs only actively serving best-effort users for a time fraction $\frac{\bar{\tau}}{\bar{\tau}^0}$ of the time originally used, provided that $\bar{\tau} < \bar{\tau}^0$. If, instead, $\bar{\tau} > \bar{\tau}^0$, the BS/AP density λ_b cannot meet the performance constraint. Thus, the BS/AP serving the typical user will be serving actively for a time fraction $\frac{\bar{\tau}}{\bar{\tau}^0}$. From this, we can calculate the energy consumed in order to satisfy the performance constraint at any feasible BS/AP density. By inspection, we can then determine the BS/AP density that minimizes energy consumption.

We have estimated numerically the potential energy savings that can be obtained by turning off BS/APs in periods of low load, while still guaranteeing quality of service. We considered different choices for the parameters of the BSs/APs energy model, while always keeping the total power consumed with utilization 100% at $1500W$ for a cellular BS, and at $10W$ for a WLAN AP. In one setting, the total energy consumption does not vary with utilization. In this setting, we choose $k_2 = 0$ W and k_1 equal to 1500 W for BSs, and 10 W for APs, in accordance with typical values found in the literature [24]. We refer to this setting as the *on-off* setting. This choice of parameters approximately models the behavior of BSs/APs currently deployed, in which the dependency of the energy consumed on load is negligible. Moreover, as current trends in BS/AP design aim at tying power consumption to BS/AP utilization, we considered a few settings in which the energy consumed by a BS/AP depends on its utilization. These *energy proportional* (EP) settings allow us to examine how switch off strategies could evolve in the future. We distinguish them by the ratio $\frac{k_2}{k_1+k_2}$ that we use as a metric for energy proportionality. For instance, a setting with $k_1 = 500$ W and $k_2 = 1000$ W is denoted EP 66.6% and one with $k_1 = 100$ W and $k_2 = 1400$ W is denoted EP 93.4%. For a complete description of the numerical evaluation settings, please refer to [71].

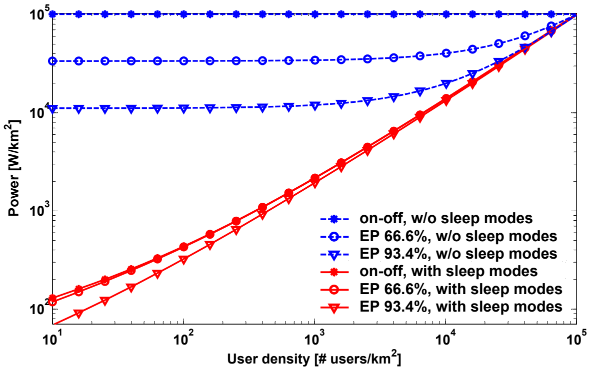


Fig. 6. Minimum power consumed by BSs per km^2, as a function of user density. Base stations layout is Poisson, and $\tau = 10\mu s$.

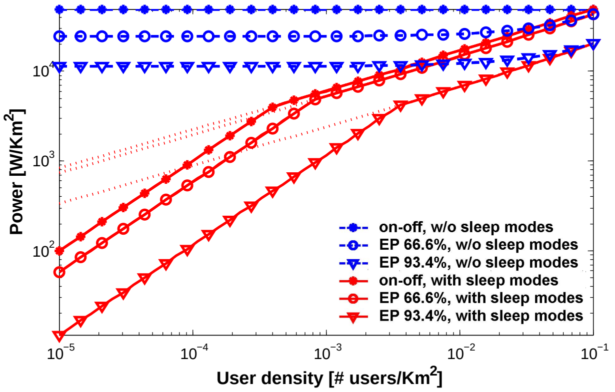


Fig. 7. Minimum power consumed by APs per km^2 in a IEEE 802.11g network, as a function of user density. $\tau = 10\mu s$.

The importance of sleep modes and system level techniques is evident from Fig. 6, where we plot the average power consumed per km^2 for the Poisson layout in two cases: (1) when sleep modes are used to adapt the base station density to load, and (2) when the network is always provisioned for the peak load, so that power savings are only due to the energy proportionality of the BS power consumption.

We observe that in case (1), when sleep modes are used, energy proportional BSs result in a slightly more energy efficient behavior at low user densities, as expected. However, we clearly see that much of the reduction in energy consumption is obtained through the intelligent use of sleep modes to adapt the active BS density to the user population, even in the absence of improved hardware.

On the contrary, in case (2), when sleep modes are not used, and the BS density remains at the level required to support the peak user density, energy proportional BSs do provide large energy savings with respect to current BSs whose power consumption is almost independent of utilization. However, the power consumption at low user densities is up to two orders of magnitude higher in this case with respect to case (1), even under highly optimistic (and probably unrealistic) assumptions on energy proportionality. This highlights the need to tackle the problem of energy consumption in cellular access networks through both improved hardware and system level techniques. It also shows clearly that, even under futuristic assumptions on the energy efficiency of hardware, the intelligent use of sleep modes and other dynamic provisioning techniques can be crucial to achieving maximum energy efficiency.

In Fig. 7 we plot the average power consumed by APs per km^2, when sleep modes are used (in red) and when they are not used (in blue). The WLAN standard considered in our evaluations is the IEEE 802.11g, which is currently one of the most widely adopted. From these plots we see how sleep modes still provide a substantial improvement of the energy efficiency of the system with respect to device level techniques alone. By comparing with the equivalent curves for BSs, we see how the rate at which APs can be turned off when user density decreases is lower for WLAN. The reason behind this difference is in the shape of the capacity curve in function of SNR (distance). In the cellular network scenario, given a transmit power of $10W$ for BSs, the resulting energy optimal BS density is such that the capacity seen at the border of the cell shows very little sensitivity to variations in cell size. Conversely, in a WLAN scenario, where the maximum received power is typically around $-40dBm$, the average SNR at the border of a cell is such that user capacity at the border of the cell is much more sensitive to distance from the AP, so that a decrease in AP density has a strong impact on average throughput. We can also see that the impact of energy proportionality of an AP on the performance of sleep modes is much higher than for BSs, significantly reducing the power consumption of the system at all user densities. Indeed, any increase in the amount of energy proportionality of APs brings to an increase of APs density and a consequent decrease in the average number of users per AP. But in WLAN this also brings to a decrease in the rate of collisions and in the number of retransmissions per packet. And this increase in the efficiency in the utilization of the shared transmission medium further enhances the energy efficiency of the system.

In order to take into account the performance of those techniques for sleep modes in WLANs, which remove the necessity to guarantee the full coverage, e.g., [88] (see Sec. IV-B1), we have assumed that the energy efficient AP density cannot fall below user density. Indeed, this emulates the best possible performance of those sleep schemes, which for very low user densities tend to turn on one AP for each active user, on average. The red continuous curves in Fig. 7 show a knee, which is due to those schemes which do not guarantee the full coverage. We see how these schemes allow to greatly improve the performance of WLAN sleep modes at low user densities. However, a more precise evaluation of the energy efficiency of such schemes should take into account the energy cost of the devices which sense the medium and selectively activate the APs.

Fig. 6 and 7 allow us also to perform a rough comparison of the total power consumed by the WLAN and by the cellular network to offer the same service, with the same QoS guarantee. We can observe that for a same user density, the WLAN consumes a slightly higher amount of energy than the cellular network. Considering that the evaluation for WLAN is more optimistic that the one made for cellular networks because of the high impact of interference on WLANs in practical settings, it turns out that when sleep modes are implemented, offloading traffic to WLANs from cellular networks might result in an increase of the overall energy consumed by the wireless access network.

VII. GENERAL CONCLUSIONS AND LESSONS LEARNED

Cellular networks and WLANs, originally developed to meet the peak of the user demand, are currently facing the forecasted traffic explosion, and are striving for the deployment of more infrastructure. Therefore, energy consumption starts to play a more important role in the overall operational expenditures, and network operators are keen on finding viable solutions to cut their energy bill. It has been shown in this paper that developing more energy-efficient hardware can only partially solve the problem, and thus there is a need to look for better alternatives to more effectively cut the overall energy consumption. As most of the energy consumed in RANs is attributed to the BS/AP operation, the development of energy-efficient BS/AP management schemes is of utmost importance. To this end, this paper has identified the most prominent examples of BS/AP management algorithms that can be applied in the different network architectures that are currently present in the networking landscape. Among many analyzed schemes it seems that these applicable in heterogeneous network scenarios present the biggest potential for both cellular networks and WLANs, despite possible initial implementation difficulties. Furthermore, offloading and cooperation of MNOs seem viable alternatives to make energy-efficient network operation reality, however, especially in the latter case, widely applicable solutions are yet to be developed. Nevertheless, this last research topic presents the biggest potential among all analyzed schemes for energy-efficient dynamic provisioning in the coming years.

To complement this picture, the aspects related to the implementation of BS management schemes in current and future infrastructures have been discussed, to indicate the key drivers that could facilitate the adoption of the proposed energy saving schemes. Fast standardization tracks were identified as one of the key drivers enabling (or, in the opposite case, blocking) broad deployment of the devised solutions. Future perspectives, especially in terms of hardware development, indicate that despite continuously increasing efficiency of PAs, there is still a need for energy-efficient RAN management.

ACKNOWLEDGMENT

This work has been supported by the TREND project (Towards Real Energy-efficient Network Design, grant agreement

n. 257740), a Network of Excellence funded by the European Community's 7th Framework Programme.

S. Lambert is funded by a grant from the Agency for Innovation by Science and Technology in Flanders (IWT).

REFERENCES

[1] Ericsson, "Ericsson mobility report," Jun. 2013. [Online]. Available: http://www.ericsson.com/res/docs/2013/ericsson-mobility-report-june-2013.pdf

[2] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 2013-2018," Feb. 2014. [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html

[3] A. Fehske, G. Fettweis, J. Malmodin, and G. Biczok, "The global footprint of mobile communications: The ecological and economic perspective," *IEEE Communications Magazine*, vol. 49, no. 8, pp. 55–62, Aug. 2011.

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

[5] S. Lambert, W. V. Heddeghem, W. Vereecken, B. Lannoo, D. Colle, and M. Pickavet, "Worldwide electricity consumption of communication networks," *Optics Express*, vol. 20, no. 26, pp. B513–B524, Dec. 2012.

[6] S. Vadgama and M. Hunukumbure, "Trends in green wireless access networks," in *Workshops, IEEE International Conference on Communications (ICC Wkshps '11)*, Jun. 2011.

[7] G. Miao, N. Himayat, Y. Li, and A. Swami, "Cross-layer optimization for energy-efficient wireless communications: a survey," *Wireless Communications and Mobile Computing*, vol. 9, no. 4, pp. 529–542, Apr. 2009.

[8] S. Ehsan and B. Hamdaoui, "A survey on energy-efficient routing techniques with qos assurances for wireless multimedia sensor networks," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 2, pp. 265–278, Second Quarter 2012.

[9] 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 Communications Surveys & Tutorials*, vol. 13, no. 2, pp. 223–244, Second Quarter 2011.

[10] A. Bianzino, C. Chaudet, D. Rossi, and J. Rougier, "A survey of green networking research," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 1, pp. 3–20, First Quarter 2012.

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

[12] L. Shi, P. Chowdhury, and B. Mukherjee, "Saving energy in long-reach broadband access networks: architectural approaches," *IEEE Communications Magazine*, vol. 51, no. 2, pp. S16–S21, Feb. 2013.

[13] M. Deruyck, W. Joseph, B. Lannoo, D. Colle, and L. Martens, "Designing energy-efficient wireless access networks: LTE and LTE-Advanced," *IEEE Internet Computing*, vol. 17, no. 5, pp. 39–45, Sep. 2013.

[14] M. Ajmone Marsan and M. Meo, "Green wireless networking: Three questions," in *10th IEEE IFIP Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net '11)*, Jun. 2011, pp. 41–44.

[15] TREND (Towards Real Energy-efficient Network Design) project. [Online]. Available: http://www.fp7-trend.eu/

[16] M. Ajmone Marsan, G. Bucalo, A. Di Caro, M. Meo, and Y. Zhang, "Towards zero grid electricity networking: Powering bss with renewable energy sources," in *IEEE International Conference on Communications (ICC '13)*, Jun. 2013, pp. 596–601.

[17] W. Vereecken, W. Van Heddeghem, M. Deruyck, B. Puype, B. Lannoo, W. Joseph, D. Colle, L. Martens, and P. Demeester, "Power consumption in telecommunication networks: overview and reduction strategies," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 62–69, Jun. 2011.

[18] EARTH (Energy Aware Radio and network technologies) project, "Energy efficiency analysis of the reference systems, areas of improvements and target breakdown," 2012.

[19] S. Zhou, J. Gong, Z. Yang, Z. Niu, and P. Yang, "Green mobile access network with dynamic base station energy saving," in *ACM 15th annual international conference on mobile computing and networking (MobiCom '09)*, Sep. 2009.

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

[21] K. Sambanis, D. Kutscher, and M. Bruneer, "Self organized energy efficient networks," in *IEEE 21st International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '10)*, Sep. 2010, pp. 1665–1670.

[22] N. Ding, D. Wagner, X. Chen, A. Pathak, Y. C. Hu, and A. Rice, "Characterizing and modeling the impact of wireless signal strength on smartphone battery drain," in *ACM SIGMETRICS/international conference on Measurement and modeling of computer systems (SIGMETRICS '13)*, Jun. 2013, pp. 29–40.

[23] W. Lemstra, V. Hayes, and J. Groenewegen, *The innovation journey of Wi-Fi: The road to global success*. Cambridge University Press, 2010.

[24] S. Chiaravalloti, F. Idzikowski, and Ł. Budzisz, "Power consumption of WLAN network elements," TKN Group, TU Berlin, Tech. Rep. TKN-11-002, Aug. 2011. [Online]. Available: http://www.tkn.tu-berlin.de/fileadmin/fg112/Papers/TR-WLAN-Power.pdf

[25] A. Murabito, "A comparison of efficiency, throughput, and energy requirements of wireless access points," *University of New Hampshire, InterOperability Laboratory*, 2009.

[26] T. Sridhar, "Wireless lan switches - functions and deployment," *The Internet Protocol Journal*, vol. 9, no. 3, pp. 2–15, Sep. 2006.

[27] M. Ajmone Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "A simple analytical model for the energy-efficient activation of access points in dense WLANs," in *1st International Conference on Energy-Efficient Computing and Networking*, Jun. 2010, pp. 159–168.

[28] N. Sharma, D. K. Krishnappa, D. Irwin, M. Zink, and P. Shenoy, "Greencache: Augmenting off-the-grid cellular towers with multimedia caches," in *4th ACM Multimedia Systems Conference (MMSys '13)*, 2013, pp. 271–280.

[29] M. Hoque, M. Siekkinen, and J. Nurminen, "Energy efficient multimedia streaming to mobile devices - a survey," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 579–597, First Quarter 2014.

[30] P. Kolios, V. Friderikos, and K. Papadaki, "Energy-aware mobile video transmission utilizing mobility," *IEEE Network*, vol. 27, no. 2, pp. 34–40, Mar.-Apr. 2013.

[31] C. Khirallah, D. Vukobratovic, and J. Thompson, "Bandwidth and energy efficiency of video broadcasting services over LTE/LTE-A," in *IEEE Wireless Communications and Networking Conference (WCNC '13)*, Apr. 2013, pp. 1909–1914.

[32] H. Ma and R. Zimmermann, "Energy conservation in 802.11 WLAN for mobile video calls," in *IEEE International Symposium on Multimedia (ISM '12)*, Dec. 2012, pp. 266–273.

[33] R. Ding and G.-M. Muntean, "Device characteristics-based differentiated energy-efficient adaptive solution for video delivery over heterogeneous wireless networks," in *IEEE Wireless Communications and Networking Conference (WCNC '13)*, Apr. 2013, pp. 4588–4593.

[34] Y. Li, M. Reisslein, and C. Chakrabarti, "Energy-efficient

video transmission over a wireless link," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 3, pp. 1229–1244, Mar. 2009.

[35] S.-P. Chuah, Z. Chen, and Y.-P. Tan, "Energy minimization for wireless video transmissions with deadline and reliability constraints," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 3, pp. 467–481, Mar. 2013.

[36] E. Uysal-Biyikoglu, B. Prabhakar, and A. El Gamal, "Energy-efficient packet transmission over a wireless link," *IEEE/ACM Transactions on Networking*, vol. 10, no. 4, pp. 487–499, Aug. 2002.

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

[38] A. P. Jardosh, K. Papagiannaki, E. M. Belding, K. C. Almeroth, G. Iannaccone, and B. Vinnakota, "Green WLANs: On-demand WLAN infrastructures," *Mobile Networks and Applications*, vol. 14, no. 6, pp. 798–814, Dec. 2009.

[39] M. Ajmone Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Multiple daily base station switch-offs in cellular networks," in *4th International Conference on Communications and Electronics (ICCE '12)*, Aug. 2012, pp. 245–250.

[40] ——, "Optimal energy savings in cellular access networks," in *Workshops, IEEE International Conference on Communications (ICC Wkshps '09)*, Jun. 2009.

[41] L. Chiaraviglio, D. Ciullo, G. Koutitas, M. Meo, and L. Tassiulas, "Energy efficient planning and management of cellular networks," in *IEEE Wireless on-demand Network Systems and Services (WONS '12)*, Jan. 2012, pp. 159–166.

[42] D. Tsilimantos, J.-M. Gorce, and E. Altman, "Stochastic analysis of energy savings with sleep mode in OFDMA wirelss networks," in *IEEE International Conference on Computer Communications (Infocom '13)*, Apr. 2013, pp. 1121–1129.

[43] H. Claussen, L. Ho, and F. Pivit, "Effects of joint macrocell and residential picocell deployment on the network energy efficiency," in *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '08)*, Sep. 2008.

[44] S. Bhaumik, G. Narlikar, S. Chattopadhyay, and S. Kanugovi, "Breathe to stay cool: adjusting cell sizes to reduce energy consumption," in *The 1st ACM SIGCOMM workshop on Green Networking*, Sep. 2010, pp. 41–46.

[45] S. E. Elayoubi and L. Saker, "Sleep mode implementation issues in green base stations," in *IEEE 21st International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '10)*, Sep. 2010, pp. 1683–1688.

[46] K. Dufkova, M. Bjelicaz, B. Moony, L. Kencl, and J.-Y. Le Boudecy, "Energy savings for cellular network with evaluation of impact on data traffic performance," in *European Wireless Conference (EW '10)*, Apr. 2010, pp. 916–923.

[47] O. Blume, H. Eckhardt, S. Klein, E. Kuehn, and W. M. Wajda, "Energy savings in mobile networks based on adaptation to traffic statistics," *Bell Labs Technical Journal*, vol. 15, no. 2, pp. 77–94, Aug. 2010.

[48] X. Weng, D. Cao, and Z. Niu, "Energy-efficient cellular network planning under insufficient cell zooming," in *IEEE 73rd Vehicular Technology Conference (VTC '11-Spring)*, May 2011.

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

[50] Z. H. Abbas and F. Y. Li, "A novel teletraffic-based power saving scheme for cellular networks with microcells," in *IEEE 36th Conference on Local Computer Networks (LCN '11)*, Oct. 2011, pp. 681–687.

[51] L. Chiaraviglio, D. Ciullo, M. Meo, and M. Ajmone Marsan, "Energy-efficient management of umts access networks," in *21st International Teletraffic Congress (ITC '09)*, Sep. 2009, pp. 15–17.

[52] 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 on Selected Areas in Communications*, vol. 29, no. 8, pp. 1525–1536, Sep. 2011.

[53] W. Vereecken, M. Deruyck, D. Colle, W. Joseph, M. Pickavet, L. Martens, and P. Demeester, "Evaluation of the potential for energy saving in macrocell and femtocell networks using a heuristic introducing sleep modes in base stations," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 170, pp. 1–14, May 2012.

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

[55] A. De Domenico, R. Gupta, and E. Calvanese Strinati, "Dynamic traffic management for green open access femtocell networks," in *IEEE 75th Vehicular Technology Conference (VTC '12-Spring)*, May 2012.

[56] I. Viering, M. Peltomaki, O. Tirkkonen, M. Alava, and R. Waldhausee, "A distributed power saving algorithm for cellular networks," in *4th IFIP TC 6 International Workshop on Self-Organizing Systems (IWSOS '09)*, Dec. 2009, pp. 232–237.

[57] I. Ashraf, L. Ho, and H. Claussen, "Improving energy efficiency of femtocell base stations via user activity detection," in *IEEE Wireless Communications and Networking Conference (WCNC '10)*, Apr. 2010.

[58] K. Son and B. Krishnamachari, "Speedbalance: Speed-scaling-aware optimal load balancing for green cellular networks," in *Mini-conference, IEEE International Conference on Computer Communications (Infocom '12)*, Mar. 2012, pp. 2816–2820.

[59] M. Ajmone Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Switch-off transients in cellular access networks with sleep modes," in *Workshops, IEEE International Conference on Communications (ICC Wkshps '11)*, Jun. 2011.

[60] C. Peng, S. B. Lee, S. Lu, H. Luo, and H. Li, "Traffic-driven power saving in operational 3G cellular networks," in *ACM 17th annual international conference on mobile computing and networking (MobiCom '11)*, Sep. 2011, pp. 121–132.

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

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

[63] B. Rengarajan, G. Rizzo, and M. Ajmone Marsan, "Qos-aware greening of interference-limited cellular networks," in *IEEE 14th International Symposium and Workshops on a World of Wireless, Mobile and Multimedia Networks (WoWMoM '13)*, Jun. 2013.

[64] L. Saker, S.-E. Elayoubi, and T. Chahed, "Minimizing energy consumption via sleep mode in green base station," in *IEEE Wireless Communications and Networking Conference (WCNC '10)*, Apr. 2010.

[65] A. Ambrosy, O. Blume, H. Klessig, and W. Wajda, "Energy saving potential of integrated hardware and resource management solutions for wireless base stations," in *IEEE 22nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC '11)*, Sep. 2011, pp. 2418–2423.

[66] P. Frenger, P. Moberg, J. Malmodin, Y. Jading, and I. Godor, "Reducing Energy Consumption in LTE with Cell DTX," in *IEEE 73rd Vehicular Technology Conference (VTC '11-Spring)*, May 2011.

[67] S. Kokkinogenis and G. Koutitas, "Dynamic and static base station management schemes for cellular networks," in *IEEE*

Global Communications Conference (GlobeComm '12), Dec. 2012, pp. 3443–3448.

[68] A. Capone, I. Filippini, B. Gloss, and U. Barth, "Rethinking cellular system architecture for breaking current energy efficiency limits," in *The Second IFIP Conference on Sustainable Internet and ICT for Sustainability (SustainIT '12)*, Oct. 2012.

[69] M. Ajmone Marsan and M. Meo, "Energy efficient wireless internet access with cooperative cellular networks," *Computer Networks*, vol. 55, no. 2, pp. 386–398, Feb. 2011.

[70] M. Sereno, "Cooperative game theory framework for energy efficient policies in wireless networks," in *Third International Conference on Future Energy Systems: Where Energy, Computing and Communication Meet (e-Energy '12)*, May 2012.

[71] B. Rengarajan, G. Rizzo, and M. Ajmone Marsan, "Bounds on qos-constrained energy savings in cellular access networks with sleep mode," in *The 23rd International Teletraffic Congress (ITC '11)*, Sep. 2011, pp. 47–54.

[72] R. Gupta, E. Calvanese Strinati, and D. Ktenas, "Energy efficient joint DTX and MIMO in cloud radio access networks," in *IEEE 1st International Conference on Cloud Networking (CLOUDNET '12)*, Nov. 2012, pp. 191–196.

[73] R. Gupta and E. Strinati, "Base-station duty-cycling and traffic buffering as a means to achieve green communications," in *IEEE 76th Vehicular Technology Conference (VTC '12-Fall)*, Sep. 2012.

[74] K. Dufkova, M. Popovic, R. Khalili, J. Boudec, M. Bjelica, and L. Kencl, "Energy consumption comparison between macro-micro and public femto deployment in a plausible lte network," in *2nd International Conference on Energy-Efficient Computing and Networking (e-Energy '11)*, May 2011, pp. 67–76.

[75] C. Pluntke, L. Eggert, and N. Kiukkonen, "Saving mobile device energy with multipath TCP," in *ACM 6th International Workshop on Mobility in the Evolving Internet Architecture (MobiArch '11)*, Jun. 2011.

[76] G. Koutitas and L. Tassiulas, "Smart grid technologies for future radio and data centers networks," *IEEE Communications Magazine*, vol. 52, no. 4, pp. 120–128, Apr. 2014.

[77] A. P. Jardosh, G. Iannaccone, K. Papagiannaki, and B. Vinnakota, "Towards an energy-star WLAN infrastructure," in *8th IEEE Workshop on Mobile Computing Systems and Applications (HotMobile '07)*, Mar. 2007, pp. 85–90.

[78] R. Nagareda, A. Hasegawa, T. Shibata, and S. Obana, "A proposal of power saving scheme for wireless access networks with access point sharing," in *International Conference on Computing, Networking and Communications (ICNC '12)*, Jan. 2012, pp. 1128–1132.

[79] J. Lorincz, A. Capone, and M. Bogarelli, "Energy savings in wireless access networks through optimized network management," in *5th International Symposium on Wireless Pervasive Computing (ISWPC '10)*, May 2010, pp. 449–454.

[80] J. Lorincz, M. Bogarelli, A. Capone, and D. Begusic, "Heuristic approach for optimized energy savings in wireless access networks," in *International Conference on Software, Telecommunications and Computer Networks (SoftCOM '10)*, Sep. 2010, pp. 60–65.

[81] K.-W. Chin, "A green scheduler for enterprise WLANs," in *Australasian Telecommunication Networks and Applications Conference (ATNAC '11)*, Nov. 2011.

[82] S. Lee, S. Seo, and N. Golmie, "An efficient power-saving mechanism for integration of WLAN and cellular networks," *IEEE Communications Letters*, vol. 9, no. 12, pp. 1052–1054, Dec. 2005.

[83] A. Capone, F. Malandra, and B. Sansò, "Energy savings in wireless mesh networks in a time-variable context," *Mobile Networks and Applications*, vol. 17, no. 2, pp. 298–311, Apr. 2012.

[84] S. Tang, H. Yomo, Y. Kondo, and S. Obana, "Exploiting burst transmission and partial correlation for reliable wake-up signaling in Radio-On-Demand WLANs," in *IEEE International Conference on Communications (ICC '12)*, Jun. 2012, pp. 4954–4959.

[85] ——, "Wake-up receiver for radio-on-demand wireless LANs," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 42, pp. 1–13, Feb. 2012.

[86] H. Yomo, Y. Kondo, K. Namba, S. Tang, T. Kimura, and T. Ito, "Wake-up ID and protocol design for radio-on-demand wireless LAN," in *IEEE 23rd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC '12)*, Sep. 2012, pp. 419–424.

[87] F. Ganji, Ł. Budzisz, and A. Wolisz, "Assessment of the power saving potential in dense enterprise WLANs," in *24th IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC '13)*, Sep. 2013, pp. 2835–2840.

[88] ——, "Assessment of the power saving potential in dense enterprise WLANs," TKN Group, TU Berlin, Tech. Rep. TKN-13-003, Apr. 2013. [Online]. Available: http://www.tkn.tu-berlin.de/fileadmin/fg112/Papers/papers_all/TR-final-20130426.pdf

[89] J.-W. Yoo and K. H. Park, "A cooperative clustering protocol for energy saving of mobile devices with WLAN and Bluetooth interfaces," *IEEE Transactions on Mobile Computing*, vol. 10, no. 5, pp. 491–504, Apr. 2011.

[90] N. Mishra, K. Chebrolu, B. Raman, and A. Pathak, "Wake-on-WLAN," in *15th international conference on World Wide Web (WWW '06)*, May 2006, pp. 761–769.

[91] V. V. Vazirani, *Approximation algorithms.* Springer-Verlag New York, Inc., 2001.

[92] 3GPP TSG WG R3, "TS 25.433 UTRAN Iub interface Node B Application Part (NBAP) signaling." [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/25433.htm

[93] ——, "TS 36.423 Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 Application Protocol (X2AP)." [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/36423.htm

[94] 3GPP TSG WG R4, "TR 36.808 Evolved Universal Terrestrial Radio Access (E-UTRA); Carrier Aggregation; Base Station (BS) radio transmission and reception." [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/36808.htm

[95] ——, "TR 36.823 Evolved Universal Terrestrial Radio Access (E-UTRA); Carrier Aggregation Enhancements; UE and BS radio transmission and reception." [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/36823.htm

[96] ——, "TR 36.927 Evolved Universal Terrestrial Radio Access (E-UTRA); Potential solutions for energy saving for E-UTRAN." [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/36927.htm

[97] 3GPP TSG WG C1, "TS 24.312 Access Network Discovery and Selection Function (ANDSF) Management Object (MO)." [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/24312.htm

[98] Ericsson, "R1-095011 Extended cell DTX for enhanced energy-efficient network operation," 2009. [Online]. Available: http://www.3gpp.org/ftp/tsg_ran/wg1_rl1/TSGR1_59/Docs/R1-095011.zip

[99] 3GPP TSG WG R2, "TS 36.300 Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2." [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/24312.htm

[100] IEEE 802.11 WG, "IEEE Std. 802.11-2012, IEEE Standard for Information technology - Telecommunications and information exchange between systems Local and metropolitan area networks - Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," pp. 1–2793, Mar. 2012. [Online]. Available: http://standards.ieee.org/getieee802/download/802.11-2012.pdf

[101] G. Hiertz, D. Denteneer, L. Stibor, Y. Zang, X. Costa, and B. Walke, “The IEEE 802.11 universe,” *IEEE Communications Magazine*, vol. 48, no. 1, pp. 62–70, Jan. 2010.

[102] Cisco, “Cisco energywise technology.” [Online]. Available: http://www.cisco.com/en/US/products/ps10195/index.html

[103] I. Haratcherev and A. Conte, “Practical energy-saving in 3G femtocells,” in *IEEE Green Broadband Access (GBA) workshop, IEEE International Conference on Communications (ICC Wkshps ’13)*, Jun. 2013, pp. 602–606.

[104] I. Haratcherev, M. Fiorito, and C. Balageas, “Low-power sleep mode and out-of-band wake-up for indoor access points,” in *Workshops, IEEE Global Communications Conference (GlobeComm Wkshps ’09)*, Nov. 2009.

[105] NTT DOCOMO, Alcatel-Lucent, Alcatel-Lucent Shanghai Bell, Ericsson, Telecom Italia, “[r1-113495], base station power model.” [Online]. Available: http://www.3gpp.org/ftp/tsg_ran/wg1_rl1/TSGR1_66b/Docs/R1-113495.zip

[106] J. Joung, C. K. Ho, and S. Sun, “Green wireless communications: A power amplifier perspective,” in *Asia-Pacific Signal Information Processing Association Annual Summit and Conference (APSIPA ASC ’12)*, Dec. 2012.

[107] W. H. Doherty, “A new high efficiency power amplifier for modulated waves,” *Proceedings of the Institute of Radio Engineers*, vol. 24, no. 9, pp. 1163–1182, Sep. 1936.

[108] B. Kim, J. Kim, I. Kim, and J. Cha, “The Doherty power amplifier,” *IEEE Microwave Magazine*, vol. 7, no. 5, pp. 42–50, Oct. 2006.

[109] J. Kim, J. Son, S. Jee, S. Kim, and B. Kim, “Optimization of envelope tracking power amplifier for base-station applications,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 4, pp. 1620–1627, Apr. 2013.

[110] A. Wentzel, C. Meliani, and W. Heinrich, “RF class-S power amplifiers: State-of-the-art results and potential,” in *IEEE MTT-S International Microwave Symposium Digest (MTT ’10)*, May 2010, pp. 812–815.

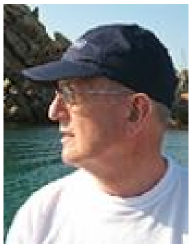
[111] G. Bianchi, “Performance analysis of the IEEE 802.11 distributed coordination function,” *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 535–547, Mar. 2000.

BIOGRAPHIES

Łukasz Budzisz (S’05-M’09) received his M.Eng.Sc. degree in Electronics and Telecommunication from Technical University of Łódź, Poland, in 2003, and a Ph.D. degree in Signal Theory and Communications from the Technical University of Catalunya, Spain, in 2009. Currently he is a postdoctoral researcher at Technical University of Berlin, Germany. His research interests are in the area of (mobile/wireless) network architectures and protocols. He is particularly interested in network congestion control, green networking, and mobility management. He has published over 30 peer-reviewed publications in prestigious journals and international conferences, a book chapter and a patent. He has participated in more than 10 competitive projects (IST, ICT, CELTIC, BMBF, SFI, CYCIT) He is a regular reviewer in a number of international journals and conferences.

Fatemeh Ganji received her Bachelor and Master degrees in Telecommunications, Electrical Engineering in Tehran, Iran. She joined the Telecommunication Networks Group at Technical University of Berlin in 2011 and focused her research activities on energy efficiency in wireless local area networks (WLANs). She is now with Department of Security in Telecommunication, Technical University of Berlin/ Telekom Innovation Laboratories (T-Labs). Her research interests are in the area of wireless network architectures, energy efficiency in wireless networks, machine learning and hardware security. She has published over 10 IEEE conference and journal papers.

Gianluca Rizzo received his degree in electronic engineering from Politecnico di Torino, Italy, in 2001. From September 2001 to December 2003, he was a researcher at Telecom Italia Lab, Torino, Italy. From January 2004 to October 2008, he was at EPFL Lausanne, where he received his Ph.D. in computer science in 2008. From 2009 to 2013 he was a staff researcher at IMDEA Networks Institute. Since April 2013 he is Senior Researcher at HES SO Valais, Switzerland. He was the recipient of a 2010 Marie-Curie Amarout Europe Programme fellowship. He is also the recipient of the best paper award at the 23rd International Teletraffic Congress (ITC), 2011, and the 11th IEEE International Symposium on Network Computing and Applications (NCA 2012). His research interests are in performance evaluation of computer networks, particularly network calculus, and green networking.



Marco Ajmone Marsan holds a double appointment as Full Professor at the Department of Electronics and Telecommunications of the Politecnico di Torino (Italy), and Research Professor at IMDEA Networks Institute (Spain). He earned his graduate degree in Electrical Engineering from the Politecnico di Torino in 1974 and completed his M.Sc. in Electrical Engineering at the University of California at Los Angeles (USA) in 1978. In 2002, he was awarded a honoris causa Ph.D. in Telecommunication Networks from Budapest University of Technology and Economics. From 2003 to 2009 he was Director of the IEIIT-CNR (Institute for Electronics, Information and Telecommunication Engineering of the National Research Council of Italy). From 2005 to 2009 he was Vice-Rector for Research, Innovation and Technology Transfer at Politecnico di Torino. Marco Ajmone Marsan is involved in several national and international scientific groups: he was Chair of the Italian Group of Telecommunication Professors (GTTI); the Italian Delegate in the ICT Committee and in the ERC Committee of the EC’s 7^{th} Framework Programme. He is a Fellow of the IEEE and he is listed by Thomson-ISI amongst the highly-cited researchers in Computer Science. He has been principle investigator for a large number of research contracts with industries, and coordinator of several national and international research projects.

Michela Meo received the Laurea degree in Electronic Engineering in 1993, and the Ph.D. degree in Electronic and Telecommunications Engineering in 1997, both from the Politecnico di Torino, Italy. Since November 2006, she is associate professor at the Politecnico di Torino. She co-authored almost 200 papers, about 60 of which are in international journals. She edited six special issues of international journals, including ACM Monet, Performance Evaluation, and Computer Networks. She was program co-chair of two editions of ACM MSWiM, general chair of another edition of ACM MSWiM and of IEEE Online GreenComm, program co-chair of the IEEE QoS-IP, IEEE MoVeNet 2007, and IEEE ISCC 2009, IEEE Online GreenComm 2012, IEEE Infocom Miniconference 2013, and she was in the program committee of about 100 international conferences, including SIGMETRICS, INFOCOM, ICC, and GLOBECOM. Her research interests include the field of performance evaluation and modeling, green networking and traffic classification and characterization.

Yi Zhang received his B.E. and Ph.D. degree both from Department of Electronic Engineering, Tsinghua University, Beijing, China, in 2007 and 2012, respectively. He was a visiting Ph.D. student at University of Califorinia, Davis, U.S.A. from 2008 to 2010. He is currently a postdoctoral researcher at Politecnico di Torino, Italy. His research interest includes energy-efficient wireless networks and optical networks. He has co-authored over 10 IEEE conference and journal papers, and served as a TPC member for over 5 IEEE/ACM conferences.

George Koutitas (M05) was born in Thessaloniki, Greece. He received the B.Sc. degree in Physics from Aristotle University of Thessaloniki Greece, 2002 and the M.Sc. degree (with distinction) in Mobile and Satellite Communications from the University of Surrey UK, 2003. During his studies, he received the Nokia Prize and Advisory Board Prize 2003 for the best overall performance and best MSc Thesis. He defended his PhD in radio channel modeling from the Centre for Communications Systems Research (CCSR) of the University of Surrey in 2007 under a full scholarship. His main research interests are in the area of Wireless Communications (modeling and optimization), Energy Efficient Networking and Smart Grids. He is involved in research activities concerning energy efficient network deployments and design, Green IT and sensor networks/actuators for smart grid applications. Currently, he is a member of the academic and research staff at the School of Science and Technology of the International Hellenic University, Greece where he also works at the Smart IHU project (rad.ihu.edu.gr). Finally, he is a post doc at the University of Thessaly (Dept. Computer Engineering and Telecommunications).

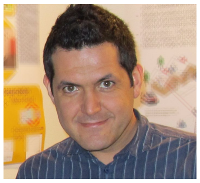
Leandros Tassiulas (S89, M91, SM/05 F/07) is Professor of Telecommunication Networks in the Department of Computer Engineering and Communications at the University of Thessaly Greece since 2002 and Associate Director of the Informatics and Telematics Institute of the Center for Research and Technology Hellas (CERTH). He holds a Diploma in Electrical Engineering from the Aristotelian Univ. of Thessaloniki, Greece, in 1987, and a Ph.D. degree in Electrical Engineering from the Univ. of Maryland, College Park in 1991. He has held positions as Assistant Professor at Polytechnic University New York (1991-95), Assistant and Associate Professor Univ. of Maryland College Park (1995-2001) and Professor Univ. of Ioannina Greece (1999-2001). He has been visiting researcher at IBM T.J.Watson research center in 1999 and in Flarion Technologies in 2003. His research interests are in the field of computer and communication networks with emphasis on fundamental mathematical models and algorithms, architectures and protocols of wireless systems, sensor networks, novel internet architectures and satellite communications. His contributions include foundational models and algorithms for network resource allocation as well as architectures and protocols for specific wireless technologies. His research has been recognized by several awards including the inaugural INFOCOM 2007 Achievement Award For fundamental contributions to resource allocation in communication networks, the INFOCOM 1994 best paper award, a National Science Foundation (NSF) Research Initiation Award in 1992, an NSF CAREER Award in 1995, an Office of Naval Research Young Investigator Award in 1997 and a Bodosaki Foundation award in 1999.

Bart Lannoo received a M.Sc. degree in electrotechnical engineering and a Ph.D. degree from Ghent University (Belgium) in July 2002 and May 2008, respectively. Since August 2002, he has been working at the Internet Based Communication Networks and Services (IBCN) research group of the Department of Information Technology (INTEC) of Ghent University, where he is currently a postdoctoral researcher. As a member of the IBCN research group, he is also affiliated with the research institute iMinds. Since September 2011, he is coordinating the Green ICT research at IBCN. His main research interests are in the field of fixed and wireless access networks, focusing on MAC protocols, Green ICT and techno-economics. He has been involved in various European research projects like the European FP7 projects ALPHA (Architectures for fLexible Photonic Home and Access Networks), OASE (Optical Access Seamless Evolution) and TREND (Towards Real Energy-efficient Network Design). He is author or co-author of more than 100 international publications, both in journals and in proceedings of conferences.

Sofie Lambert received her M.Sc. degree in Photonics in 2011 from Ghent University (UGent, Belgium). In February 2012, she joined the Information Technology department (INTEC) of Ghent University to work on Green ICT. There she studied the worldwide electricity consumption in ICT infrastructure, specifically in communication networks. Her current research is focused on the energy efficiency of various future optical and wireless access network architectures, and the impact of energy saving strategies such as sleep modes.

Mario Pickavet received an M.Sc. and Ph.D. degree in electrical engineering, specialized in telecommunications in 1996 and 1999, respectively. Since 2000, he is professor at Ghent University (Belgium) where he is teaching courses on discrete mathematics, multimedia networks and network modeling. He is co-leading the research cluster on Network Modeling, Design and Evaluation (NetMoDeL) covering 4 research topics: Fixed internet architectures and optical networks, techno-economic studies, green ICT and design of network algorithms. He has published about 300 international publications, both in journals (IEEE JSAC, IEEE Comm. Mag., Journal of Lightwave Technology, Proceedings of the IEEE, ...) and in proceedings of conferences. He is co-author of the book 'Network Recovery: Protection and Restoration of Optical, SONET-SDH, IP, and MPLS'.

Alberto Conte is a research team manager in the Wireless domain at Alcatel-Lucent Bell Labs. After receiving M.Sc. degrees in corporate communications and telecommunication engineering from Institut Eurcom (France) in 1996 and Politecnico di Torino (Italy) in 1997 respectively, he joined the software department of Bell Labs, initially working on embedded systems and VoIP. He then shifted his research interests to wireless technologies for both enterprises (wifi) and public/operated cellular systems. Mr. Conte and his team pioneered the definition of end-to-end mobile solutions based on OFDMA (WiMAX, LTE). His current research focuses on 4G/5G networks and their energy efficiency. He participates to several collaborative EU projects and is an active lecturer on LTE technology at Telecom ParisTech, INSA-Lyon and Supelec.



Ivaylo Haratcherev is a Research Engineer in Wireless Domain at Alcatel-Lucent Bell Labs France. He received his M.Sc. degree in Systems and Control Engineering from the Technical University of Sofia, Bulgaria in 2000. He obtained a Ph.D. degree in Computer Science in the Parallel and Distributed Systems group at Delft University of Technology, the Netherlands in 2006. His research interests include energy efficiency and low-power design of wireless networks and network devices, self-configuration and self-optimization in wireless communications, embedded systems, testbed design and implementation, RTOS, and real-time control systems.

Adam Wolisz received his degrees (Diploma 1972, Ph.D. 1976, Habil. 1983) from Silesian Unversity of Technology, Gliwice, Poland. He joined TU Berlin in 1993, where he is a chaired professor in telecommunication networks and executive director of the Institute for Telecommunication Systems. He is also an adjunct professor at the Department of Electrical Engineering and Computer Science, University of California, Berkeley. His research interests are in architectures and protocols of communication networks. Recently he has been focusing mainly on wireless/mobile networking and sensor networks.