The impact of the Access Point power model on the energy-efficient management of infrastructured Wireless LANs

Rosario G. Garroppo[1], Gianfranco Nencioni[1], Gregorio Procissi[1], Luca Tavanti[1,*]

Abstract

The reduction of the energy footprint of large and mid-sized IEEE 802.11 access networks is gaining momentum. When operating at the network management level, the availability of an accurate power model of the APs becomes of paramount importance, because different detail levels have a non-negligible impact on the performance of the optimisation algorithms. The literature is plentiful of AP power models, and choosing the right one is not an easy task. In this paper we report the outcome of a thorough study on the impact that various inflections of the AP power model have when minimising the energy consumption of the infrastructure side of an enterprise wireless LAN. Our study, performed on several network scenarios and for various device energy profiles, reveals that simple one- and two-component models can provide excellent results in practically all cases. Conversely, employing accurate and detailed power models rarely offers substantial advantages in terms of power reduction, but, on the other hand, makes the solving algorithms much slower to execute.

Keywords: Wireless LAN, Optimization, Power Consumption Model, Energy Efficiency, Network Management, Resource Allocation

1. Introduction

The energy saving issue in wireless networks is currently the focus of many research activities. For example, there is a plethora of works dealing with the analysis and reduction of the power consumption in cellular networks [1, 2, 3], wireless sensor networks [4, 5], wireless mesh networks [6, 7, 8], and also wireless Local Area Networks (WLANs) [9, 10, 11].

With specific focus on IEEE 802.11-based networks, there is an increasing interest in the design of efficient reconfiguration algorithms to reduce the power consumption of the infrastructure-side when the load is scarce [9, 12, 13]. Indeed, by turning some access points (APs) off and adjusting the power radiated by the active APs, it is possible to achieve considerable energy savings with respect to the currently widespread technique of continuously operating the WLAN at full power. Obviously, this energy gain shall

*Corresponding author.
Email address: luca.tavanti@iet.unipi.it (Luca Tavanti)
[1]Dipartimento di Ingegneria dell'Informazione, Università di Pisa, Italy.

not be obtained at the expenses of the coverage nor the quality of service levels provided when the trasmission power of all APs is set to the maximum.

In designing such reconfiguration algorithms it is often necessary to first define a power model of the AP. On the basis of this model it is then possible to study and perform the optimisation of the system from an energy-aware perspective.

The assumptions on the AP power model have, in general, a non-negligible impact on the output of the energy-management algorithm, especially because the optimisation is often tailored on the features of the model itself. If an inappropriate power model is employed, it might occur that the planned or expected energy improvement is reduced or even nullified. Consequently, the choice of an appropriate power model is crucial for the valid outcome of any reconfiguration algorithm. However, given the plethora of models proposed over the years, it is not easy to understand which is the most suitable.

In this paper, we specifically address the last point, i.e. our goal is providing some insights and indications to help choosing the appropriate AP power model for some common and future network scenarios. To this aim, we perform a study on the effectiveness and implications that various AP power models have in minimising the energy consumption of an enterprise WLAN system. We first define a general model of the WLAN and of the AP power consumption. We then build a mathematical programming model to minimise the total power consumption (while guaranteeing that the whole traffic demand is met). Finally we solve it to optimality for various "realisations" of the AP power model, under different network compositions and device energy profiles. At the end of this process, we are able to extract valuable information on the usefulness and impact of the AP power model details.

Going in more detail, we basically build our AP power model on the one defined by Garcia-Saavedra *et al.* [14], which can be regarded as the most detailed and reliable appeared so far in the literature. In our model, four major elements contribute to the power consumption of the AP: baseline (due to circuitry powering), the radio frontend, the airtime, and the traffic processing cost (power drain of CPU and memory). Then, by selectively excluding one or more of these elements, we obtain less complete models down to the simplest on/off one.

Then, we characterise all the features of the WLAN system in their most general form, without performing rough approximations nor simplifications. Indeed, while such approximations and/or simplifications might, on the one hand, lead to a simpler mathematical programming model, on the other hand they might undermine the effectiveness of our study, e.g. by leading to solutions that are not applicable or unsatisfactory for the original problem.

To achieve the maximum energy-saving of the system, we operate through a mathematical program on two decision aspects at the network management level: *(i)* associating each user terminal to one of the available APs, and *(ii)* setting the transmission power level of each AP.

The mathematical program we devised is linear (notwithstanding the non-linearity of some functions, as it will be detailed in Sections 3.2 and 4.2) and optimised for fast solving times, so that we can analyse non-trivial network scenarios in acceptable times. The program is solved to optimality by means of a general-purpose Mixed-Integer Linear Programming (MILP) solver, for a wide range of network scenarios and for four different classes of devices. In fact, current (and future) AP equipment is characterised by different ratios among the power drained by its major elements. Consequently, the application of the power model(s) to diverse device classes might lead to different optimisation strategies and resource allocations.

In particular, we distinguish the cases of homogeneous and heterogeneous networks. While the former is undoubtedly the most utilised in the literature, and also quite common in practice (e.g. brand new deployments), it is becoming not so unfrequent for large WLANs to be composed of different types of APs (e.g. due to replacement of malfunctioning equipment, upgrades of old apparatuses, network densification after the initial deployment). Indeed, our work unveils interesting findings about heterogeneous networks which have often been neglected in the literature under the reasoning that passing from an homogeneous to an heterogeneous network is just a matter of more complex notation.

1.1. Contribution

The main contribution of the paper can be summarised as follows.

- We provide an extensive analysis of the impact that the various elements of the AP power model have in optimising the energy efficiency of an enterprise-grade WLAN. This is achieved by means of a general integer linear program of the WLAN which accounts for an accurate and modular power model of the AP and for non-simplistic network features.
- On the basis of the analysis, we delineate the best strategy to minimise the energy consumption in current and future WLANs. We show that accounting for traffic processing at the APs is detrimental, because it hardly brings any improvements in terms of energy savings but makes the problem much harder to execute. We also demonstrate that resource consolidation is often the best strategy. We find that the presence of heterogeneous devices might be exploited to increase the energy efficiency of the system.

The rest of the paper is structured as follows. In the next Section we give a brief summary of the related literature and works. Then, in Section 3 we illustrate the analytical model of the WLAN system, with particular emphasis on the power model of the AP, and sketch the mathematical formulation of the problem. Section 4 describes the framework under which we lead our analysis, whose results are reported and commented in Section 5. Finally, the concluding remarks are drawn in Section 6.

2. Related work

Over the years, several AP power models have been proposed, with diverse assumptions and varying degrees of detail. For example, simple on/off models, in which the AP has a constant power drain, have been and are still widely used. A more sophisticated and yet quite popular model ascribes the energy consumption to two elements: a baseline one, plus a term that depends – often linearly – from the activity of the radio interface, the so-called airtime [15]. Then, various measurement campaigns have led to characterise the power consumption as a (variably complex) function of the traffic load, antenna settings (especially for MIMO devices), datagram size, transmission / reception data rate, encryption, number of connected clients [16, 17, 18, 19].

Recently, a very detailed AP power model has been described by Garcia-Saavedra *et al.* in [14]. The model is extracted from a series of accurate measurements on various real APs. It comprises, in addition to the "classic" baseline and airtime elements, a factor that weights the energy cost of processing the traffic.

In parallel to AP power modelling works, several studies have been produced on the optimisation of the WLAN power consumption. Each of these have assumed the APs to be characterised by a specific power model. For example, Jardosh et al. [20] proposed a strategy to dynamically turn APs on/off to follow the resource demand of the users. This approach, which has been translated into a working testbed, was based on empirical considerations, including the simple on/off AP power model.

On the other hand, a more rigorous optimisation approach based on integer linear programming (ILP) has been followed by Lorincz et al. [9] and Gendron et al. [13]. In both works, the AP power consumption is split in two components, fixed and variable. The latter, in particular, depends on the radiated power. Zhang et al. [21] also employed a very similar model in investigating both the power allocation and the placement of an energy-harvesting AP in a single cell WLAN with cooperative users.

The simple on/off power model is again at the basis of the work by Couto da Silva et al. [22], who exploited a queuing model to decide the assignment of the users in a portion of a dense WLAN with co-located APs. At last, we mention the work of Garcia-Saavedra et al. [23], who studied the trade-off between energy and throughput optimisation in case of heterogeneous user devices. An exact, but quite complex energy model, was also derived. Simplifying, it ascribed the power consumption to a fixed term plus the radiated power and the airtime.

Even from this short survey, it emerges that many AP power models have been employed in the past. However, to the best of our knowledge, no prior work exists that have studied and evaluated the properties and effects of the various AP power models in the context of energy-saving optimisation in wireless LANs. Our work aims at filling this gap.

3. AP power model and problem formulation

3.1. Wireless LAN model

We model the wireless LAN system as follows.

There is a set of deployed access points (APs) that must serve a set of user terminals (UTs). For each AP there exists a set of different transmission power levels (PLs), but at most one PL must be chosen for each AP. Each AP can also be powered off. The UTs are static, and their positions are known. This is a rather common abstraction in network design and resource allocation, where each UT in fact represents the barycenter of an area that contains a quantum of demand [24]. For example, one such UT may aggregate the traffic of all the physical devices present in a given office or room. Thanks to this abstraction, it is also possible to build a stationary traffic model of a mobile population. Then, each UT has a traffic demand that must be satisfied, and each UT must be assigned to exactly one AP.

Let $\mathcal{I}$ be the set of UTs, $\mathcal{J}$ the set of deployed APs, and $\mathcal{K}$ the set of PLs; let i, j, and k be the indexes for such sets.

The power P_j consumed by the generic AP j can be ascribed to several elements:

$$P_j = b_j + A_j\, w_j + t_j. \tag{1}$$

At first there is a constant part, say b_j, which is bound to the mere fact that the device is powered on, and therefore encompasses AC/DC conversion, basic circuitry powering, dispersion, etc. Then, we find a first variable part, say w_j, which is generated by the wireless interface. In turn, w_j can be split into the transmission (w_j^t) and reception (w_j^r) parts. w_j^t essentially depends on the radiated power p_j through an efficiency factor η_j that accounts e.g. for the electrical model of the device; w_j^r derives from the frame reception operations. A variable factor, say A_j, accounts for the so-called "airtime", i.e. the fraction of time the device is either transmitting or receiving frames. A_j can in fact be split into the two directions: $A_j = (a_j^t + a_j^r)\, A_j$, with $a_j^t, a_j^r \in [0,1]$, and $a_j^t + a_j^r = 1$. The last variable part, say t_j, weights the traffic processing operation, and depends on the amount of traffic handled by the AP, say T_j, and the traffic processing cost μ_j.

An expanded form of (1) can be written to make all elements contributing to the power consumption explicit:

$$P_j = b_j + (a_j^t\, \eta_j\, p_j + a_j^r\, w_j^r)\, A_j + \mu_j\, T_j. \tag{2}$$

From (1) it is easy to identify the four components that sums up to build the power model of the AP: the baseline consumption (b_j), the airtime (A_j), the radio operations (w_j), and the processing toll (t_j). Accordingly, we call this characterisation "the four-component power consumption model“, in short the 4C model. The 4C model is currently the most complete characterisation of the AP power consumption [14]. With respect to the model proposed in [14], however, we have not made any distinction between the processing toll of incoming and outgoing traffic, because, as a matter of fact, they require the same energy.

Eq. (1) and (2) address the general case of heterogeneous devices, for which all terms are dependent on the AP index j. However, in practical circumstances, it may occur that some of the elements (such as b_j, η_j, and μ_j) do not vary among the APs, thus allowing to simplify the model.

With regard to the radiated power p_j, note that the vast majority of the commercial APs have a set of preset power values to choose among (see e.g. [25]), and these values are pretty standardised among all vendors and devices. Consequently we can assume that p_j can take a value in the set $\{p_{jk}\}, k \in \mathcal{K}$, but also that these values are not a function of the specific AP j, and therefore $p_{jk} = p_k, \forall j \in \mathcal{J}$.

Finally, to complete the description of the problem, we introduce the following elements:

- d_i, the traffic demand of UT i;
- L, the average packet length;
- r_{ij}, the capacity of link (i,j), i.e. the data rate available between AP j and UT i; r_{ij} is function of the power p_j radiated by AP j; this relationship can be arbitrarily complex, because it depends on various factors (such as modulation and coding scheme, rate adaptation algorithms, overhead), in a nonlinear way;

- r_{ijk}, the capacity of link (i, j) when AP j transmits with PL k, i.e. when $p_j = p_k$;
- $\rho \in [0, 1]$, an AP "utilisation" factor, which can be empoyed to limit the AP airtime to values smaller than 1;
- $\mathcal{I}'_{jk} = \{i \in \mathcal{I} : r_{ijk} \geq \frac{d_i}{\rho}\}$, the set of UTs whose traffic demand can be carried by AP j when it is using PL k.

Throughout our work, we assume that the wireless links are symmetric, which implies that $r_{ij} = r_{ji}$, and consequently that the ratios of the downlink/uplink airtimes are equal to those of the downlink/uplink traffic demand. This assumption does not limit neither the generality nor the validity of the WLAN model, but allows to keep the notation simpler. For example, it is not necessary to split the traffic demand d_i in the downlink and uplink directions, since they contribute to the airtime in the same manner.

3.2. Mathematical programming model

The objective of our study is to minimise the overall power consumption of the APs while satisfying the traffic demand of the users. It must be decided whether to use or not each AP, which PL to assign to each (used) AP, and to which powered-on AP to assign each UT. Therefore, the problem can be seen as a discrete location problem, where the capacity to assign to each location also has to be decided (this is the design part of the problem). Hence, we see this problem as a particular case of a broader class of location-design problems, where both the location and capacity dimensioning decisions must be taken.

To formulate the mathematical programming model, we define the following sets of binary variables:

- x_{ijk}, which is set to 1 if UT i is assigned to AP j using PL k, 0 otherwise;
- y_{jk}, which is set to 1 if AP j uses PL k, 0 otherwise.

The objective is to minimise the total power consumption, as described by:

$$z = \min \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}} \Big\{ b_j\, y_{jk} + (a^t_j\, \eta_{jk}\, p_k + a^r_j\, w^r_j) \sum_{i \in \mathcal{I}'_{jk}} \frac{d_i}{r_{ijk}}\, x_{ijk} + \mu_j \sum_{i \in \mathcal{I}'_{jk}} \frac{d_i}{L}\, x_{ijk} \Big\}, \quad (3)$$

The minimisation is subject to the following constraints:

$$\sum_{j \in \mathcal{J}, k \in \mathcal{K} : i \in \mathcal{I}'_{jk}} x_{ijk} = 1, \quad i \in \mathcal{I}, \quad (4)$$

$$\sum_{k \in \mathcal{K}} y_{jk} \leq 1, \quad j \in \mathcal{J}, \quad (5)$$

$$\sum_{i \in \mathcal{I}'_{jk}} \frac{d_i}{r_{ijk}}\, x_{ijk} \leq \rho\, y_{jk}, \quad j \in \mathcal{J}, k \in \mathcal{K}. \quad (6)$$

$$x_{ijk} \in \{0, 1\}, \quad j \in \mathcal{J}, k \in \mathcal{K}, i \in \mathcal{I}'_{jk}, \tag{7}$$

$$y_{jk} \in \{0, 1\}, \quad j \in \mathcal{J}, k \in \mathcal{K}. \tag{8}$$

Equations (4) are the single assignment constraints that impose that each UT must be assigned to exactly one AP and one PL. Equations (5) impose that at most one PL can be selected for each AP. Equations (6) are the capacity constraints for each AP, which include the utilisation factor ρ. The joint enforcement of (4) and (6) also ensures that the PL assignments are coherent among the x and y variables and that no UT is assigned to powered-off APs. Finally, relations (7) and (8) define the integrality of the variables.

A few noteworthy remarks follow. An AP j is turned off if no PL is selected, i.e. if $\sum_{k \in \mathcal{K}} y_{jk} = 0$. By defining and using the set $\mathcal{I}'_{jk}$ we arranged the programming model so that the x_{ijk} variables exist only when $r_{ijk} \geq \frac{d_i}{\rho}$. This allows for a faster resolution of the programming model, but has no impact on its generality and correctness. The presence of the data rate at the denominator in (3) and (6) generally leads to a non-linear problem, because the rate depends on the radiated power p_j, which is an unknown of the problem (specifically, via the y_{jk} variables: $p_j = \sum_{k \in \mathcal{K}} p_k \, y_{jk}$). To overcome this hurdle, in both the objective function and the constraints the rate function r_{ij} is always employed in its "sampled" version r_{ijk}. This allows to build a linear programming model (which can be fed directly to general-purpose solvers) which includes a non-linear function (see Section 4.2 for a realistic example).

3.3. Variants to the power model of the AP

The just outlined model accounts for all aspects of the AP power consumption, i.e. it minimises the total power consumption according the 4C model defined in (1). However, in many studies, simplifications of this model have been (and are still) employed. The most meaningful model variants are the following:

$$\text{1C} \qquad P_j = b_j, \tag{9}$$

$$\text{2C} \qquad P_j = b_j + \zeta \, A_j, \tag{10}$$

$$\text{3Cw} \qquad P_j = b_j + w_j \, A_j, \tag{11}$$

$$\text{3Ct} \qquad P_j = b_j + \zeta \, A_j + t_j, \tag{12}$$

In detail, (9) represents the simplest characterisation, which is a basic on/off model. Eq. (10) adds a variable part that depends on the airtime. Yet, differently from (1), this is not weighted by a variable "radio" factor, but by a constant term (ζ). The power consumption of the radio frontend is added by (11), whereas (12) adds the traffic processing cost.

4. Computational analysis

4.1. Identification and characterisation of the device classes.

Following to the work of Garcia-Saavedra et al. [14], we have indentified three classes of devices, as a function of the relation among the addends of the AP power model (1). A fourth class has been added to account for the future trends of energy efficient devices, in which the baseline consumption should be drastically reduced. Table 1 illustrates how the maximum power (say P_{max}) has been divided among the three addends. The

"Radio" element includes both w_j and the airtime A_j (which, in fact, has been assumed to be 1 for this operation). Devices belonging to class D3 can be taken as an example of the majority of current carrier-grade devices, in which the baseline consumption amounts to 75% of the total [19]. In contrast, class D4 is representative of future energy-efficient devices, which should scale the power with the usage. Classes D1 and D2 are in between these two extremes, representing, to some extent, two cases of single chip low power solutions [18].

Table 1: Classes of devices and related power distribution (in Watt).

Class	Baseline	Radio	Processing
D1	4.8	2.4	4.8
D2	6	3	3
D3	9.6	1.2	1.2
D4	1.2	4.8	6

Note how P_{max} has been normalised to the same value (12 W) for all classes, in order to eliminate any bias due to unbalancement among the classes. This normalisation has been achieved by scaling proportionally each component, so as to keep the ratios among the components of each class fixed (and in line with the numbers extracted from [14]). We also re-normalised the components when assessing the performance of the non-4C models in the heterogeneous tests. This was necessary to keep the maximum power to the same value (P_{max}=12 W) for all the devices, given that the lack of one or more components leads, for non-4C models, to an unbalancement in the power consumption among devices of different classes. For example, when assessing the 3Cw model, there is a huge disparity in the maximum consumption between classes D3 and D4 due to the lack of the processing term, and therefore we scaled b and η parameters so that the 12 W value is reached by the sum of the sole baseline and radio components.

4.2. Parameters of the optimisation model.

The first aspect to specify for the computational analysis is the function that binds the rates r_{ij} to the transmitted power p_j. To this purpose we can start from this simple formula that defines the rate r_{ij} available above the MAC layer:

$$r_{ij} = \frac{10^6 \cdot L}{\tau_{ij}}, \tag{13}$$

where τ_{ij} is the average global time (in μs) for delivering a single frame. This time includes the overhead created by the MAC and physical layers, such as headers, control frames, and various protocol procedures. In the hypothesis of ideal channel access, there exists a formula that allows to compute τ_{ij} for the IEEE 802.11g standard[2] (see [26, 27]):

$$\tau_{ij} = \tau_{proto} + 4 \lceil \frac{L_{h,t} + L}{N_{DBPS}} \rceil, \tag{14}$$

where τ_{proto} is the protocol delay (e.g. back-off, SIFS, DIFS) plus the physical frame delimiters (preamble and sync fields), $L_{h,t}$ is the length of headers and trailers plus the ACK frame, and N_{DBPS} is the number of data bits per OFDM symbol. In turn, N_{DBPS}

can be approximated as $N_{DBPS} = 4\,\tilde{r}_{ij}$, with $\tilde{r}_{ij}$ being the raw bit rate available at the physical layer (in Mbps). In case of non-ideal channel, we may assume that $\tilde{r}_{ij}$ is the average raw bit rate resulting from the rate adaptation policies aimed at keeping the packet error rate at a roughly constant value.

The raw bit rate $\tilde{r}_{ij}$ can be related to the transmitted power p_j by the classic signal propagation rules. To this purpose, we employed a simplified version of the COST-231 multi-wall path loss model for indoor, non-LOS environments [28]. This allows to compute the path loss α as a function of the number and type of walls, columns, and other building elements. Then, as reported in several experimental studies, such as [29], it is possible to bind the signal-to-noise ratio expressed in dB ($\mathrm{SNR}_{ij}^{[dB]}$) to the data rate by means of a linear function, where β and δ are two suitable "linearisation" factors. A further aspect to be considered is that, when the received power falls below a given sensitivity threshold γ, we must assume $\tilde{r}_{ij} = 0$. Similarly, we must also cap $\tilde{r}_{ij}$ to the maximum rate allowed by the specific technology, say $\tilde{r}_{max}$. Thus, we can summarise the relationship between $\tilde{r}_{ij}$ and p_j with this unique nonlinear expression:

$$\tilde{r}_{ij} = \begin{cases} \min\{\beta \cdot \mathrm{SNR}_{ij}^{[dB]} + \delta, \tilde{r}_{max}\}, \text{ if } p_j + \alpha_{ij} > \gamma, \\ 0, \text{ otherwise,} \end{cases} \tag{15}$$

where p_j, α_{ij}, and γ are all expressed in dB. As for the specific parameter values, we have set L to 700 bytes [30], $\tau_{proto} = 157.5\,\mu$s, $L_{h,t} = 428$ bits, $\beta = 1.76$ and $\delta = -7.48$ [29], $\gamma =$ -121 dB [25], and $\tilde{r}_{max} = 54$ Mbps.

To complete the parameter list, we set p_k taking values in the range from $p^{max} = 0.1$ W to $p^{min} = (\frac{1}{2})^{|\mathcal{K}|-1}\,p^{max}$, with

$$p_{k+1} = \tfrac{1}{2}\,p_k, \qquad k = 1, \ldots, |\mathcal{K}| - 1, \tag{16}$$

where, clearly, $p_1 = p^{max}$ and $p_{|\mathcal{K}|} = p^{min}$.

4.3. Scenarios and instance generation method.

We assessed the performance of the power consumption models over five networks composed of a different mix of devices. Four of them are homogeneus, in which all the APs belong to the same class, with the class varying from D1 to D4. In the fifth, the APs of all classes are mixed in the same proportion, i.e. each AP belongs to any given class with probability 0.25.

We then defined a set of 13 scenarios. The features of each scenario are determined by several parameters: the number of APs, UTs, and PLs, the amount of traffic demand per UT, the spatial density of the APs[3] (Υ_{AP}, measured in number of AP per m^2), and the ratio between the downlink and uplink traffic (which, given that the links are symmetric, is also the ratio between the trasmission and reception airtimes, a^t and a^r). Table 2 reports the values of each parameter. Each scenario is generated by changing the value of one of the parameters from "reference" to "higher" or "lower"; in this way it is possible to estimate their impact on the model performance. Table 3 details the paramater values for each scenario.

[2]This formula could be easily adapted for the IEEE 802.11a standard, and, with some more work, extended to the IEEE 802.11n standard. Yet, this is well beyond the purpose of our work, since the rate

Table 2: Parameter values for the tested scenarios.

Parameter	Reference	Lower	Higher
Number of APs, $\lvert\mathcal{J}\rvert$	16	8	32
Number of UTs, $\lvert\mathcal{I}\rvert$	96	48	192
Number of PLs, $\lvert\mathcal{K}\rvert$	3	2	4
Mean traffic demand, $\bar{d}_i$	300 kbps	150 kbps	600 kbps
Traffic variation, Δd_i	67%	10%	–
AP density, Υ_{AP}	$0.003/\text{m}^2$	$0.001/\text{m}^2$	$0.01/\text{m}^2$
Downlink fraction, a^t	0.75	0.25	–

Table 3: Expanded list of the scenarios with the related parameters.

Scenario	$\lvert\mathcal{J}\rvert$	$\lvert\mathcal{I}\rvert$	$\lvert\mathcal{K}\rvert$	$\bar{d}_i$ [kbps]	Δd_i	Υ_{AP} [m^{-2}]	a^t
1	16	96	3	300	67%	0.003	0.75
2	8	96	3	300	67%	0.003	0.75
3	32	96	3	300	67%	0.003	0.75
4	16	48	3	300	67%	0.003	0.75
5	16	192	3	300	67%	0.003	0.75
6	16	96	2	300	67%	0.003	0.75
7	16	96	4	300	67%	0.003	0.75
8	16	96	3	150	67%	0.003	0.75
9	16	96	3	600	67%	0.003	0.75
10	16	96	3	300	67%	0.001	0.75
11	16	96	3	300	67%	0.01	0.75
12	16	96	3	300	67%	0.003	0.25
13	16	96	3	300	10%	0.003	0.75

5. Computational results

The total power consumption for each network composition (i.e. D1-only, D2-only, D3-only, D4-only, and a mix all classes) and for each AP power model is reported in Figure 1. Then, Figure 2 shows the time necessary to find the optimal solution normalised to the time of the simplest 1C model. The absolute values, obtained on a PC equipped with a 2.27 GHz 64-bit processor, can be found in Table 4. Finally, Figure 3 summarises the airtime values and the number of active APs yielded by the best possible solutions. The bars in the figures refer to the average computed over all instances (10 per scenario)

function serves just as an example in the computational analysis.

[3]Note that $\lvert\mathcal{J}\rvert$ and Υ_{AP}, i.e. the number and density of APs, affect the tested scenarios in different ways. An increased/decreased $\lvert\mathcal{J}\rvert$ (with constant Υ_{AP}) implies a larger/smaller test area but the same degree of freedom in associating the UTs to the APs (i.e. the average number and distance of available APs per UT is the same). Conversely, a higher/lower Υ_{AP} (with constant $\lvert\mathcal{J}\rvert$) determines more/less association possibilities per each UT.

and all scenarios (for a total of 130 instances). The markers for the 95% confidence intervals are also shown. In general, we have registered pretty similar performance across all scenarios. The scenarios for which the numbers differ sensibly from the average are highlighted and discussed in the text. Detailed comments on the figures are in the following subsections.

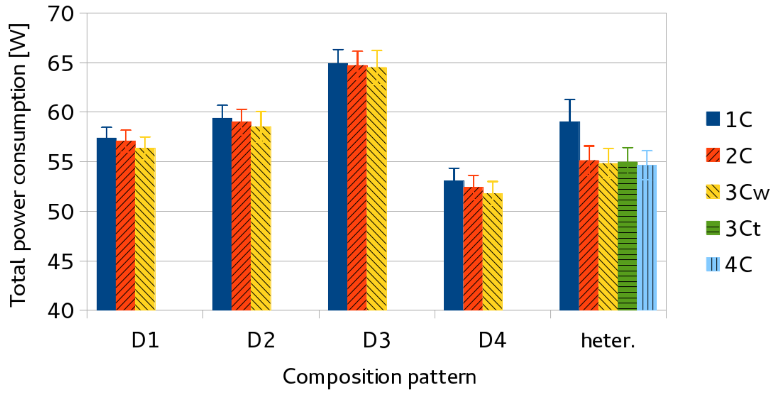

Figure 1: Power consumption vs. network composition. The vertical axis starts at 40 W to better emphasize the differences among the power consumption models.

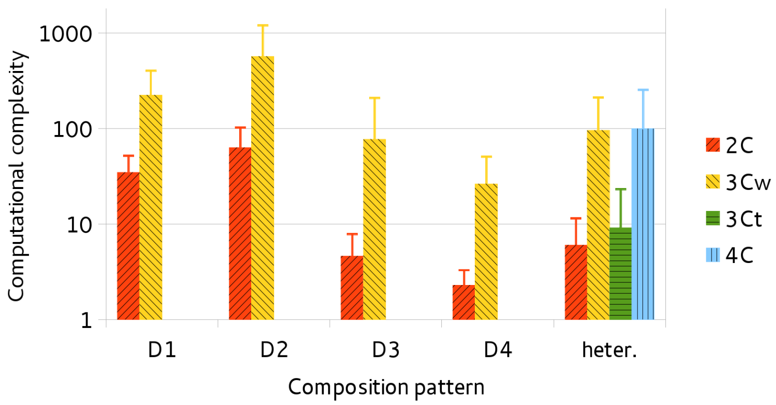

Figure 2: Normalised computational times vs. network composition. The vertical axis is in logarithmic scale.

5.1. Homogeneous networks.

5.1.1. Traffic processing is uninfluential.

Starting the analysis with the homogeneous patterns, it can be immediately noted that for these network scenarios we have reported the power consumption for the 1C, 2C and 3Cw models only. In fact, following to the definition of the objective function (3), the traffic processing term becomes a constant, because all μ_j are the same (say μ) and

Table 4: Average CPU times (in seconds) vs. network compositions and power models.

	D1	D2	D3	D4	heter.
1C	8.43	3.28	56.9	5.54	5.06
2C	292	207	261	12.6	30.2
3Cw	1881	1859	4377	145	483
3Ct					46.2
4C					504

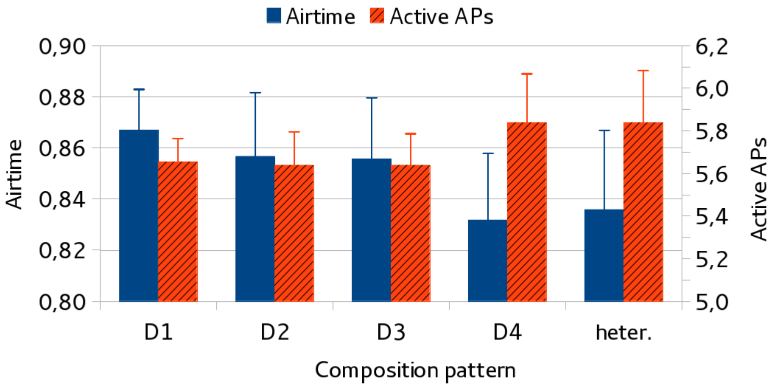

Figure 3: Airtime and number of active APs vs. network composition for the best possible allocation (i.e. based on model 4C).

all the traffic (say D) must be processed (as a result of constraints (4)):

$$\mu_j \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \sum_{k\in\mathcal{K}} \frac{d_i}{L} x_{ijk} = \mu \sum_{i\in\mathcal{I}} \frac{d_i}{L} = \mu \frac{D}{L}.$$

Therefore the traffic processing term is not relevant for the solution of the problem (there is no point in miminising a constant). Given that all APs consume the same energy to process the traffic, it makes no difference on which AP the traffic is processed. Therefore, in homogeneous networks, model 4C is equivalent to 3Cw, and model 3Ct to 2C.

Thus, a first remarkable point is that in homogeneous deployments there is no use in accounting for the power consumption that arises from traffic processing.

5.1.2. Resource consolidation fits all

The second aspect that emerges from Figure 1 is that in all homogeneous networks the gains of the 2C and 3Cw models are marginal with respect to model 1C. Indeed, among all scenarios and network compositions, the highest difference we observed between 1C and 3Cw is 5.1%. This occurred when all APs belongs to class D4 and are very densely deployed (scenario 11). On average, however, employing the most complete 3Cw model leads to a power efficiency gain of about 1.6% with respect to the simple 1C model. On the other hand, solving 3Cw to optimality requires roughly 100 times greater computational resources than 1C (see Figure 2).

Therefore, unless even minimal energy reductions are valuable, it seems clear that in homogeneous networks employing the simple on/off power model leads to good results without requiring much computational effort. Note that employing the 1C model implies that the optimal allocation implements the resource consolidation strategy, i.e. it concentrates the traffic on the least number of APs, and all these APs operate at the maximum transmission power. Also, 1C does not distinguish among the classes of devices, as proven in Figure 4, where it is manifest that the solution is almost the same for all network compositions. Nevertheless, even in cases where the most complete 3Cw model might make some difference in terms of number of active APs (see the D4 bars in Figure 3 and Figure 4), the power gain is still minimal (2.4%). In addition, since the solving times of 1C are very short (see Table 4), it can lend itself for quasi-real time resource allocation techniques.

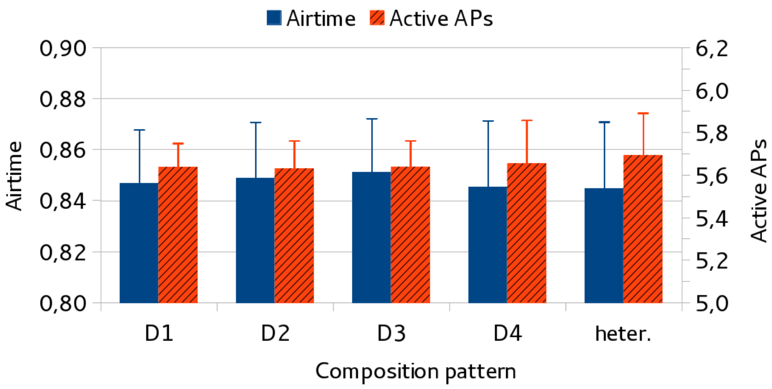

Figure 4: Airtime and number of active APs vs. network composition for the resource consolidation strategy implemented by model 1C.

As for the 2C model, it lies somewhere in between 1C and 3Cw, but it provides neither short solving times (10 times slower than 1C, on average), nor good power gains (just a 0.7% better than 1C). Therefore, the presence of the airtime in the AP power model does not bring substantial benefits. The same can be said for the PLs. The difference between 2C and 3Cw in terms of power efficiency is also minimal (0.9%), but has a notable impact in terms of solving time (roughly 11 times higher).

5.1.3. Better to pay as you go

In terms of absolute power consumption, it is apparent (see again the groups of bars in Figure 1) that using devices with a low baseline consumption (i.e. class D4) is definitely beneficial with respect to devices with a high baseline consumption (i.e. class D3). The average power gain of D4 over D3 is around 19%, with a peak of 29.3% for scenario 10 (and model 3Cw). Since in such a scenario the APs are less densely distributed, there is the need of keeping more APs active, but less loaded (figures varying from 7.1% to 12.6%), in order to cover the whole service area. As a consequence, the more the APs allow to scale the power, the more efficient the network becomes.

More in general, we can see from Figure 3 that when D4-based devices are employed (and model 3Cw is used for computing the solution), the optimal allocation provides for a few more active APs (+4%), but with slightly less occupancy (smaller airtimes, -3%), than employing devices belonging to class D3. The reason is that in scenario D4 the power consumption model of the APs has a very low baseline figure, and therefore it is beneficial, in terms of overall power consumption, to have more active APs than in the other scenarios. However, since the density of the UTs is constant and the offered traffic is roughly the same, it follows that the UTs are closer to the APs, and consequently the service data rate is higher and the airtime is smaller. In other words, the use of class D4 tends to spread the load over more APs, whereas class D3 tends to consolidate the traffic over less APs. Nevertheless, D4-based APs allow to save considerably more power (19.7%).

The last comment is about the effect of the components of the AP power model on the diverse device classes. The addition of the airtime (model 2C) allows for a peak power improvement of 1.2% in the D4-based network with respect to the average (over all network patterns) 0.7%, and a poor 0.3% for class D3. Similarly, enriching the model with the radio frontend power (model 3Cw) provides a further 0.9% gain on average, with peaks of 1.2% both in D1 and D4, and a minimum (0.3%) in D3. Thus the D4 class of devices allows for greater system optimisations when more complete power models are employed, whereas class D3 is almost model-agnostic (as it could have been expected given the numbers in Table 1).

5.2. Heterogeneous network

5.2.1. Simple is not enough

In this case, differently from all homogeneous networks, the simple 1C power model shows its weakness. The total consumption (see Figure 1) is definitely higher than the other models, with a peak value of about 13% recorded for scenarios 3 and 4 (when the ratio $|\mathcal{I}|/|\mathcal{J}|$ is the smallest). In particular, the 4C model can yield a tangible advantage in those scenarios where there are more degrees of freedom in allocating the resources. For example, in cases in which the ratio between the number of users and the number of APs is low, when the traffic is scarce and with little variation among the users, and when the APs are densely deployed, employing the 4C model yields the largest gap with respect to simple models. In fact, to verify this concept, we have run further computational experiments on a set of instances with the just mentioned features (i.e. with $|\mathcal{J}| = 32$, $|\mathcal{I}| = 96$, $\bar{d}_i = 150\,\text{kbps}$, $\Delta d_i = 10\%$, and $\Upsilon_{AP} = 0.01$), and the outcome was a 18.4% of power saving with respect to the 1C model. Therefore, even though the 1C model runs in much shorter times than 4C (two orders of magnitude, on average), it also performs definitely worse.

5.2.2. Airtime is the key

Still from Figure 1, it can be seen how the greatest jump is between the 1C and 2C models. This means that the sole introduction of the airtime in the power model allows for notable energy savings. In detail, the 2C model achieves on average 7% lower power consumption than 1C, with six times slower solving times. The difference between 2C and 4C is, in terms of power consumption, less than 1%. Hence, the 2C power model yields very good results in reasonable times, thus being an interesting tradeoff between the complete, but complex to solve 4C model, and the fast-executing, but simplistic 1C model.

Further, but limited gains, can be obtained by adding to the 2C model the wireless operations related power (3Cw), the traffic processing term (3Ct), or both (4C). The first leads to an improvement of a miserable 0.5% with respect to 2C, but it also needs sixteen times more computational resources. The second reduces the consumption by a negligible 0.3%, delivered in 1.5 times as 2C. Finally, 4C improves over 2C by a 0.9% in roughly the same time as 3Cw.

5.2.3. The traffic processing term, again

A further insight into the results can be achieved by analysing the impact of the traffic processing term, which, in the homogeneous networks, was not relevant, as discussed in Sec. 5.1.1.

From a comparison between the 3Ct and 2C models, and between the 3Cw and 4C models, which differ by the traffic processing term only, it appears (see Figure 1) that the results are almost identical. Going to the numbers, 3Ct and 2C yield, respectively, 55.11 W and 54.96 W, whereas 3Cw and 4C yield 54.82 W and 54.64 W. The gains allowed by considering the traffic processing term are indeed minimal (no more than 0.3%), but the time to obtain them might be increased by up to four times (see Figure 2).

5.2.4. The mix is better than the average

The last information we extract from Figure 1 is about the power consumption of the heterogeneous pattern. Note, at first, that the average total power consumption of all the homogeneous networks, when computed by means of the 3Cw model, is 57.78 W. For the heterogeneous this is 54.64 W (4C model), which implies that some efficiency can be obtained also from having a mix of different devices. Indeed, having a diversity of AP classes to choose among is a benefit that the optimisation program can use to best match the AP selection in function of the specific scenario.

However, it must also be pointed out that this saving can be achieved only when the more complete models are employed. The 1C model is not that smart. For example, in comparing the D1-based network and the heterogeneous case, it can even provide worse results (since it chooses the APs irrispectively of their energy profile).

Therefore, a simple AP power model can be deemed suitable for homogeneous networks, but for heterogeneous deployments at least the airtime should be considered to obtain an acceptable power efficiency.

5.2.5. Adjusting the trasmission power

Figure 5 shows how the AP power levels are allocated when using the 4C power model in the heterogeneous case. We recall that three PLs are available in all scenarios except for scenario 6 (two PLs) and 7 (four PLs), and that PL1 is the highest level and PL4 the lowest. From the figure, it appears that there is no uniformity across the scenarios, as in some the highest PL dominates, whereas in a few the lowest is most used.

From a deeper analysis, it emerges that the highest power level is typically employed when the network is scarcely loaded (e.g. scenarios 3, 4, 8, 12). In such cases, the resource consolidation approach is followed by the optimisation model, and the strategy is to keep active as few APs as possible, but with the highest power, in order to accommodate the demand of many distant users. Most notably, all APs are set to PL1 in scenario 12, and this scenario corresponds to the case $a^t = 0.25$, i.e. the uplink traffic is dominant.

Conversely, when the network is heavily loaded, the optimal approach tends to keep more APs active, but with lower power levels. When dealing with much more traffic per user (or with more users), it is necessary to employ more APs, but, given that the users are closer, it is not necessary to radiate at full power. This applies, for example, to scenarios 2, 5, and 9.

Note that this last remark was not as obvious as it might look. In fact, decreasing the transmission power implies that lower data rates might be available at the user terminals, and consequently longer airtimes might be necessary to transfer the data. Hence, transmission power and airtime are inversely proportional. Since they are combined by multiplication (see (2) and (3)), spotting which element is dominating is not straightforward.

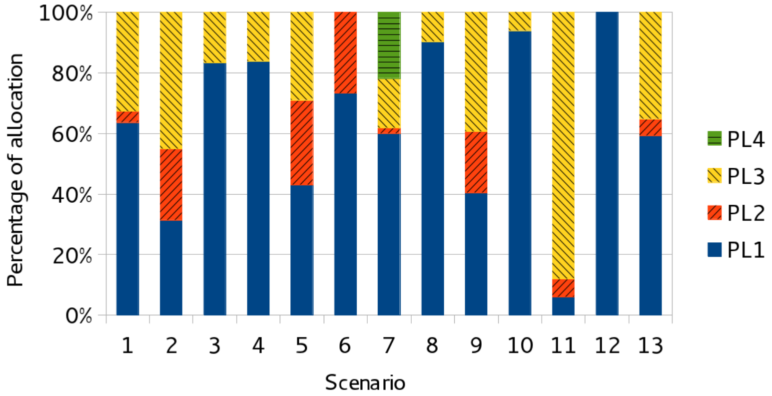

Figure 5: Cumulative percentage of allocated power levels for the various scenarios when the 4C model is used in the heterogeneous network.

A common aspect to almost all scenarios is the bimodal distribution of the power levels. In most cases, the middle power lever is very seldomly chosen (if any). Only in the three scenarios in which the network is heavily loaded (i.e. 2, 5, and 9), does PL2 have some utility.

From a comparison with the homogeneous networks, see Figure 6-9, it emerges how such a bimodal distribution characterizes the heteroneous and, even more, the D4-based homogeneous networks only, whereas all other homogeneous networks presents a smoother PL distribution.

A last remark that is common to all network types, is the fact that in scenario 11 almost all APs employs PL3. This is the consequence of the increased AP density, which implies that the UTs are closer to the APs, and therefore radiating at the lowest power is sufficient to reach all UTs and accommodate their traffic. On the other hand, in the sparse scenario 10, there is a predominance of PL1, which compensates for the longer distances between APs and UTs.

6. Discussion and conclusions

In the paper we have discussed the impact that the various elements of the AP power consumption model have when optimising the power efficiency of an enterprise wireless

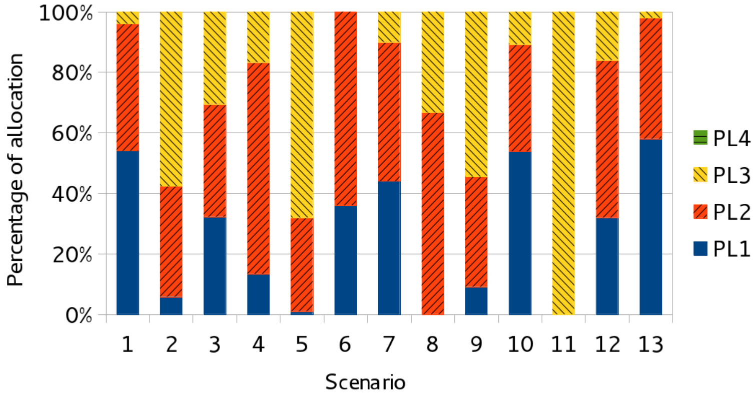

Figure 6: Cumulative percentage of allocated power levels for the various scenarios when the 4C model is used in the D1-based network.

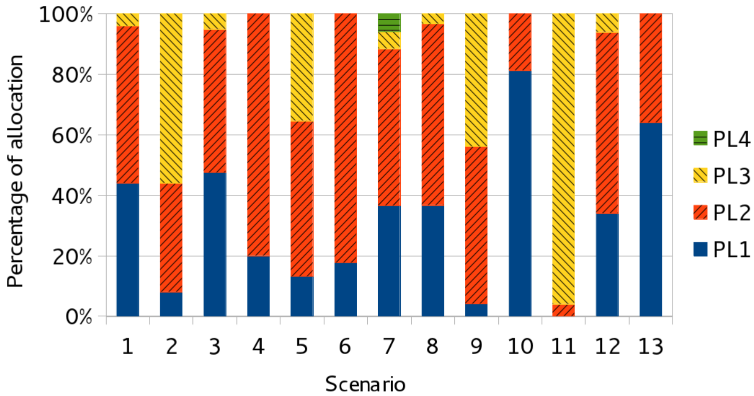

Figure 7: Cumulative percentage of allocated power levels for the various scenarios when the 4C model is used in the D2-based network.

LAN. The performance of the models has been assessed for four classes of devices with different balance of the power components, deployed in homogeneous and heterogeneous networks, and for a variety of operational scenarios. From this extensive analysis, it emerged that:

- The power consumption due to the traffic processing operation is fundamentally irrelevant. This has been mathematically proven for the homogeneous networks, whereas in the heterogeneous case the computational analysis revealed that its impact is well below the 1%.

- In homogeneous networks, the simplest on/off power model is sufficient to provide very good results. Further but marginal energy gains can be achieved with the more sophisticated 2C and 3Cw models, but at the expense of much greater computational complexity.

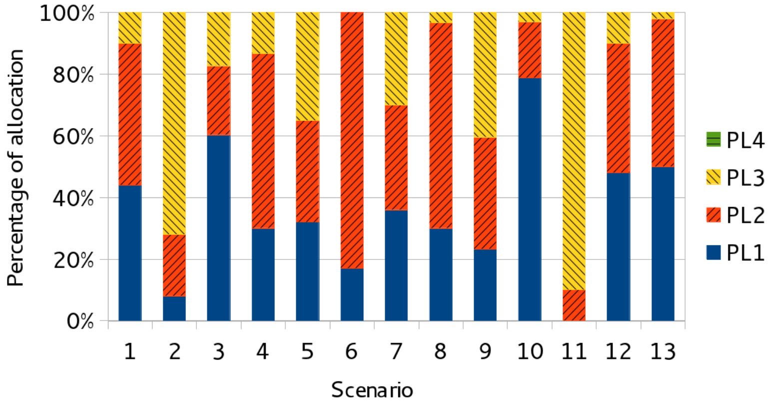

Figure 8: Cumulative percentage of allocated power levels for the various scenarios when the 4C model is used in the D3-based network.

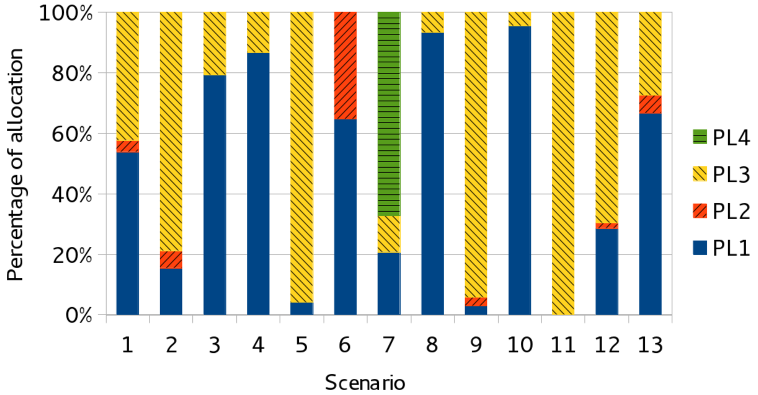

Figure 9: Cumulative percentage of allocated power levels for the various scenarios when the 4C model is used in the D4-based network.

- In heterogeneous networks, the best compromise between energy efficiency and computational complexity is given by the 2C model, which includes the baseline and the airtime components. The fast-executing on/off power model could be regarded as a passable alternative only for heavily loaded networks or in cases with an evenly distributed traffic demand. Conversely, the complete 4C model might produce some energy benefits only for networks where the APs are very densely deployed (but with much longer solving times).

- The "resource consolidation" strategy, i.e. turning off as many APs as possible, tends to be a good solution in the majority of the scenarios. This is especially true for the homogeneous networks, with the exception of the class D4 case. Indeed, when the APs are characterised by a very low baseline consumption and for heterogeneous networks, keeping active more APs with a low transmission power is more energy efficient than applying consolidation. However, to achieve this result, it is necessary to employ the 3Cw or 4C models, which are also the most complex to solve.

- When more power levels (PLs) are available at the APs (and a suitable model is used for the optimisation), for the heterogeneous and D4-based homogeneous networks the optimal PL allocation tends to be bimodal, i.e. either the highest or the lowest PL is chosen. The PLs are more evenly distributed in the other homogeneous networks. In any case, however, the contribution of the PLs to the overall power saving is quite limited, but requires a notable computational effort.

In the light of the above-mentioned findings, for currently deployed networks, which are mostly built with sets of identical (or very similar) devices for which the baseline power consumption is prevailing, the best approach to obtain satisfying energy-efficiency figures is to apply the resource consolidation strategy. This can be easily achieved with a simple on/off power model, which has also the advantage of being quickly solvable.

Nevertheless, as the networks grow and evolve with the addition (or replacement) of new and different devices, as well as for future networks based on more energy-efficient APs, this straightforward strategy might no longer be suitable. In such scenarios, enhancing the power model with a term that weights the power consumption due to the airtime becomes the mandatory upgrade to keep the energy performance of the system close to optimality. However, since this addition comes at the cost of notably increased solving times, the study of very efficient heuristics might be a requisite if real-time network reconfiguration is envisioned.

As a final remark, note that both the suggested 1C and 2C models do not account for the availability of multiple transmission power levels at the APs. This implies that, to allocate also the PLs more complex models must be used, but you cannot expect notable energy savings. On the downside, resorting to heuristics to overcome the complexity of these models might not be convenient, because the solutions provided by 2C and 3Cw/4C models are very close, and therefore designing heuristics that are much faster than 3Cw but with better performance than 2C seems to be a very tough job.

Acknowledgements

This work was partially supported by the Italian Ministry of University and Research (MIUR) under the “GreenNet” FIRB research project.

References

[1] X. Wang, A. V. Vasilakos, M. Chen, Y. Liu, T. T. Kwon, A Survey of Green Mobile Networks: Opportunities and Challenges, Mobile Networks and Applications 17 (2012) 4–20.

[2] M. Wildemeersch, T. Quek, C. Slump, A. Rabbachin, Cognitive Small Cell Networks: Energy Efficiency and Trade-Offs, IEEE Transactions on Communications 61 (2013) 4016–4029.

[3] X. Zhang, Z. Su, Z. Yan, W. Wang, Energy-Efficiency Study for Two-tier Heterogeneous Networks (HetNet) Under Coverage Performance Constraints, Mobile Networks and Applications 18 (2013) 567–577.

[4] Y. Mo, B. Wang, W. Liu, L. Yang, A Sink-Oriented Layered Clustering Protocol for Wireless Sensor Networks, Mobile Networks and Applications 18 (2013) 639–650.

[5] N. Tung, P. Vinh, The Energy-Aware Operational Time of Wireless Ad-Hoc Sensor Networks, Mobile Networks and Applications 18 (2013) 454–463.

[6] A. De La Oliva, A. Banchs, P. Serrano, Throughput and energy-aware routing for 802.11 based mesh networks, Computer Communications 35 (2012) 1433–1446.
[7] R. Zhu, J. Wang, Power-Efficient Spatial Reusable Channel Assignment Scheme in WLAN Mesh Networks, Mobile Networks and Applications 17 (2012) 53–63.
[8] C. Luo, S. Guo, S. Guo, L. T. Yang, G. Min, X. Xie, Green Communication in Energy Renewable Wireless Mesh Networks: Routing, Rate Control, and Power Allocation, IEEE Transactions on Parallel and Distributed Systems 25 (2014) 3211–3220.
[9] J. Lorincz, A. Capone, D. Begusic, Optimized network management for energy savings of wireless access networks, Computer Networks 55 (2011) 514–540.
[10] S.-L. Tsao, C.-H. Huang, A survey of energy efficient MAC protocols for IEEE 802.11 WLAN, Computer Communications 34 (2011) 54–67.
[11] W. Ge, S. Chen, H. Ji, X. Li, V. Leung, Green Access Point Selection for Wireless Local Area Networks Enhanced by Cognitive Radio, Mobile Networks and Applications 18 (2013) 553–566.
[12] J. Lorincz, A. Capone, D. Begusic, Heuristic Algorithms for Optimization of Energy Consumption in Wireless Access Networks, KSII Transactions on Internet and Information Systems 5 (2011) 626–648.
[13] B. Gendron, R. G. Garroppo, G. Nencioni, M. G. Scutellà, L. Tavanti, Benders Decomposition for a Location-Design Problem in Green Wireless Local Area Networks, Electronic Notes in Discrete Mathematics 41 (2013) 367–374.
[14] A. Garcia-Saavedra, P. Serrano, A. Banchs, G. Bianchi, Energy consumption anatomy of 802.11 devices and its implication on modeling and design, in: Proceedings of the 8th international conference on Emerging networking experiments and technologies, CoNEXT '12, 2012, pp. 169–180. doi:10.1145/2413176.2413197.
[15] H. Wu, S. Nabar, R. Poovendran, An Energy Framework for the Network Simulator 3 (NS-3), in: Proc. 4th International ICST Conference on Simulation Tools and Techniques (SIMUTools), 2011, pp. 222–230.
[16] Y. Xiao, P. Savolainen, A. Karppanen, M. Siekkinen, A. Ylä-Jääski, Practical Power Modeling of Data Transmission over 802.11G for Wireless Applications, in: Proceedings of the 1st International Conference on Energy-Efficient Computing and Networking, 2010, pp. 75–84. doi:10.1145/1791314.1791326.
[17] K. Gomez, R. Riggio, T. Rasheed, D. Miorandi, F. Granelli, Energino: A hardware and software solution for energy consumption monitoring, in: 10th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2012, pp. 311–317.
[18] G. Palem, S. Tozlu, On energy consumption of Wi-Fi Access Points, in: IEEE Consumer Communications and Networking Conference (CCNC), 2012, pp. 434–438. doi:10.1109/CCNC.2012.6181015.
[19] J. Wu, H. Li, W. Sun, Study on Real Energy Consumption of Large-scale Campus Wireless Network, in: Proc. International Conference on Computing, Networking and Communications (ICNC), 2013, pp. 605–609. doi:10.1109/ICCNC.2013.6504155.
[20] A. P. Jardosh, K. Papagiannaki, E. M. Belding, K. C. Almeroth, G. Iannaccone, B. Vinnakota, Green WLANs: On-Demand WLAN Infrastructures, Mobile Networks and Applications 14 (2009) 798–814.
[21] X. Zhang, Z. Zheng, J. Liu, X. Shen, L.-L. Xie, Optimal power allocation and AP deployment in green wireless cooperative communications, in: IEEE Global Communications Conference (GLOBECOM), 2012, pp. 4000–4005. doi:10.1109/GLOCOM.2012.6503742.
[22] A. P. Couto da Silva, M. Meo, M. A. Marsan, Energy-performance trade-off in dense WLANs: A queuing study, Computer Networks 56 (2012) 2522–2537.
[23] A. Garcia-Saavedra, P. Serrano, A. Banchs, M. Hollick, Energy-efficient fair channel access for IEEE 802.11 WLANs, in: IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2011. doi:10.1109/WoWMoM.2011.5986436.
[24] K. Tutschku, Demand-based radio network planning of cellular mobile communication systems, in: Proc. 17th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM), volume 3, 1998, pp. 1054–1061. doi:10.1109/INFCOM.1998.662915.
[25] Cisco, Aironet 3600 Series Access Point, 2014.
URL: http://www.cisco.com/c/en/us/products/wireless/aironet-3600-series/index.html.
[26] J. Jun, P. Peddabachagari, M. Sichitiu, Theoretical maximum throughput of IEEE 802.11 and its applications, in: Second IEEE International Symposium on Network Computing and Applications (NCA), 2003, pp. 249–256. doi:10.1109/NCA.2003.1201163.
[27] J. Bordim, A. Barbosa, M. Caetano, P. Barreto, IEEE802.11b/g Standard: Theoretical Maximum Throughput, in: First International Conference on Networking and Computing (ICNC), 2010, pp.

197–201. doi:10.1109/IC-NC.2010.40.
[28] European Comission, COST 231 – Digital Mobile Radio Towards Future Generations Systems, Final Report, 1999.
[29] J. Zhang, K. Tan, J. Zhao, H. Wu, Y. Zhang, A Practical SNR-Guided Rate Adaptation, in: Proc. IEEE Conference on Computer Communications (INFOCOM), 2008, pp. 2083–2091. doi:10.1109/INFOCOM.2008.274.
[30] Center for Applied Internet Data Analysis, The CAIDA UCSD Statistical information for the CAIDA Anonymized Internet Traces, 2014.
URL: http://www.caida.org/data/passive/passive_trace_statistics.xml.