The Energy Saving Potential of Static and Adaptive Resource Provisioning in Dense Cellular Networks

Gianluca A. Rizzo
HES-SO Valais
Switzerland
gianluca.rizzo@hevs.ch

Marco Ajmone Marsan
Institute IMDEA Networks, Spain &
Politecnico di Torino, Italy
marco.ajmone@polito.it

***Abstract*—In this paper we study static and dynamic approaches to energy efficiency in dense cellular networks, where interference is one of the main limiting factors. We consider the two main approaches to energy efficiency through adaptive management of the network capacity: Base station (BS) sleeping and cell zooming. We propose an analytic framework for the assessment of the energy efficiency potential of several joint planning and management strategies. Our approach is based on stochastic geometry tools, on an approximate but accurate model of interference, and on a detailed, measurement-driven power model. For a given user density, we show how to derive the optimal BS density, and the BS transmit power which minimizes the mean power consumption of the network, while achieving a target QoS level. Through numerical evaluations, we show the potential savings enabled by joint (and disjoint) optimization of transmit power and density of active BSs. For a realistic network scenario, our approach suggests that huge energy savings are achievable by combining sleeping and zooming. In addition, we show that a static strategy, based on carefully planning the density of installed BS and their transmit power, can achieve most of the benefits of capacity tuning achievable through either sleeping or zooming. This result has a very high relevance for network operators, since it allows avoiding the feared decrease in operational lifetime which the daily switching of BS entails.**

I. Introduction

Energy efficiency in wireless access networks is improving, mostly thanks to the development of more energy parsimonious base station (BS) generations [1]. With the soon-to-come heterogeneous network architectures of 5G, and the consequent increase in capacity through network densification in high traffic areas, many small cells will be deployed, with limited individual energy consumption, but huge total energy demand [2]. This environment offers interesting opportunities for energy efficiency through network management, since the dense small cell deployments will be used to support traffic peaks only in some periods of a day. The two main approaches to energy efficiency through adaptive management of the network capacity are known as base station sleeping and cell zooming [1], [3]. With the term base station sleeping we mean the possibility of activating and deactivating BSs on demand, while cell zooming consists in adaptively selecting the transmission power of BSs, so as to modify the BS coverage. Of course, both these approaches have a significant impact of the interference pattern.

Sleep modes have been proposed in many variants, each of which has shown to be able to deliver savings of at least $20-30\%$ in realistic settings [1], [4]. However, the validity of such estimations of energy savings are limited to the specific algorithm, and to the setting considered. A few works [4] have proposed an analytic approach for estimating the overall energy saving potential of sleep modes. However, such estimates do not take into account the effects of interference, which play a crucial role in determining the overall network performance, especially in future ultra-dense scenarios. Moreover, they are based on a simplified power model, which depends only on BS utilization, and does not take into account the complex interplay between transmit power and utilization.

Cell zooming algorithms have been proposed [3], as a way to adapt to localized increases in traffic demand, as well as a possible approach to compensate for the loss in coverage and capacity due to BS switch off.

Energy efficient network planning has also been proposed [1], [5], typically as a complement to dynamic network management strategies. However, the vast majority of existing results are heuristics based on simple, linear power models which typically neglect the complex interplay between these parameters, the effects of interference, and the total power consumed. Existing estimates of energy savings are closely tied to the details of the given algorithm, and of the setup chosen for the assessment.

This issue becomes even more important in future 5G small cell environments, which are typically dominated by interference. Overall, which dynamic BS management strategy holds the largest potential for energy savings is still an open issue.

In this paper we propose an analytic framework for the assessment of the potential for energy efficiency of several joint network planning and either static or dynamic management strategies, based on stochastic geometry tools, on an approximate but accurate model of interference, and on a detailed, measurement-driven power model. More specifically, for a given user density, we propose an approach for deriving the optimal BS density and the BS transmit power which minimize the mean power consumption of the network, while achieving a target QoS level. Through numerical evaluations, we show the potential savings enabled by joint (and disjoint) optimization of transmit power and density of active BSs. On a realistic setting with measurement-based traffic profiles, we show that huge energy savings are achievable by combining the sleeping and zooming approaches. In addition, we also show that a static strategy, based on carefully planning the density of installed BSs and their transmit power alone reaps most of the benefits of capacity tuning through either sleeping or zooming. This is a result which can have very high relevance for network operators, since it allows them to avoid the feared decrease in operational lifetime (and thus a reduced network availability) which the daily switching of the operational state of BSs entails [6]. Our results are bounds with respect to what can be actually achieved in real cellular networks, since we assume that any base station density is achievable, although this is clearly not possible in practice.

The paper is organized as follows. In Section II we present the system model, and in Section III we derive a model for user perceived performance. In Section IV we formulate an optimization problem for the derivation of the energy optimal network configuration. In Section V we assess numerically our results, and in Section VI we conclude the paper.

II. Model and Assumptions

We consider the downlink information transfer in a cellular access network, which, with respect to the uplink, typically accounts for the bulk of the energy consumed by a BS [7].

User terminals are assumed to be distributed in space according to a homogeneous planar Poisson point process, Π_u, with intensity λ_u users per km^2, while BSs are assumed to be distributed in space according to a homogeneous planar Poisson point process, Π_b, with density λ_b BSs per km^2. Such distribution reflects the result of real life constraints on BS locations [1].

We assume that all BS densities are feasible. In the homogeneous Poisson process layout of BSs, if each BS independently makes a decision to either turn off, or stay on, according to some probability (probabilistic sleep modes), the resulting point process of BSs is a thinned homogeneous Poisson process, and all BS densities are indeed achievable.
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 short-term 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 the BS time across all associated users. The network performance metric is the average of the per-bit delay. The performance constraint that is enforced is as follows: if the average per-bit delay experienced by a *typical* user, $\bar{\tau}$, is less than a predefined threshold $\bar{\tau}^0$, then users are said to perceive satisfactory performance, and the corresponding BS distribution is feasible. Here, $\bar{\tau}$ is computed as the expectation of τ with respect to the Palm distribution associated with Π_u.

We assume that the network serves a mix of best effort traffic and constant bit rate traffic (the latter can be voice, or voice-like traffic, or video), that is served at strictly higher priority than best-effort data traffic. We consider that a fraction γ of the users makes voice-like calls with mean call holding time μ_H^{-1} and mean inter-call waiting time μ_W^{-1}. The rate requirement for an active call is R_0 bits per second. The remaining fraction $(1-\gamma)$ of the users requests best-effort service. BSs serve calls for the fraction of time that ensures that the user achieves exactly the target bit rate, a fraction of which is consumed by voice-like calls, while the rest is filled by best-effort data traffic. The active BSs in the network must be capable of providing a user-perceived average per-bit delay of at most $\bar{\tau}^0$, while prioritizing voice-like traffic. We assume that, due to the necessity of providing adequate performance to best effort users, voice-like traffic consumes a small fraction of the cell bandwidth, so that the resulting blocking probability for voice calls is negligible. We assume best effort users are in saturation, i.e. they are always receiving content.

A. Channel and Service Model

In this paper, we do not consider the effect of fading and shadowing, and only take into account distance-dependent path loss. We assume *random frequency reuse* is in place, with reuse factor k. That is, every BS is assigned one out of k frequency bands with equal probability. [1]
We assume users are served by the BS which results in the largest SINR at the user location. We consider urban scenarios, where the high capacity demand justifies the use of strategies for energy efficient network planning and management, and where the assumption of large attenuation (with exponent $\alpha \geq 3$) typically holds. In these settings, as no fading is considered and BSs are assumed to all have the same transmit power, assuming that users associate to the closest BS is a reasonable approximation [8].
Denote by $S(x)$ the location of the BS that is closest to a user located at x. We denote the capacity to a user located at a distance r from the BS by $C(r)$ bit/s per Hertz. In what follows, we model $C(r)$ using Shannon's capacity law, given by

$$C(r) = (B/k)\log_2\left(1 + \frac{P_T r^{-\alpha}}{N_0 + I(r,k)}\right)$$

where α is the attenuation coefficient, N_0 the power spectral density of the additive white Gaussian noise, and $I(r,k)$ the total received interfering power. With ρ^V we denote the fraction of BS time that is required, on average, to serve voice-like traffic. In order to serve a call originating from a user at a distance $|x|$, the BS has to devote a fraction of time equal to $\frac{R_0}{C(|x|)}$. For the BS to which the user located at the origin is associated,

$$\rho^V = \sum_{x\in\mathcal{X}} \frac{R_0}{C(|x|)} \cdot \frac{\mu_H^{-1}}{\mu_H^{-1}+\mu_W^{-1}} \mathbf{1}_{S(x)=S(0)}, \tag{1}$$

where $\frac{\mu_H^{-1}}{\mu_H^{-1}+\mu_W^{-1}}$ is the average fraction of time that a user requires voice service, and $\mathcal{X}$ is the set of voice user locations. $\mathbf{1}_{S(x)=S(0)}$ is the indicator function of the event that a user at location x is served by the same BS that serves the user at

[1] In realistic settings, better frequency management mechanisms can be implemented, which minimize the probability of two neighboring cells sharing the same frequency band. However, with probabilistic sleep modes random reuse allows keeping such probability independent of the density of active BSs, and hence it allows estimating the impact of interference on energy saving strategies independently of any frequency management mechanism.

the origin. BSs devote only the resources (time) that remain after serving the voice calls to best effort users. We assume that BSs use a processor sharing mechanism to divide capacity among all the connected best-effort users.

B. Energy Consumption Model

The power consumed by a BS depends on a number of factors, which vary according to the BS type (e.g. macro, micro, femto) and the implementation technology, among others. In what follows, we refer to the BS energy models proposed in [7], [9], [10]. With minor differences, they all propose a power model which consider only the power consumed by downlink communications, as it constitutes the bulk of the total power consumed by a BS. The structure of the power model we consider is $P_{tot} = P_{PA} + P_{BB} + P_{RF} + P_{OH}$, where:

- P_{PA} is the power consumed by the power amplifier. It naturally scales down when the transmit power p is decreased, when the number of occupied subcarriers is reduced in idle mode operation, and/or when there are subframes not carrying data.
- P_{BB} is the power consumed by the baseband signal processing units. Typically, it does not depend on load, unless micro-sleep modes are implemented [7], [10].
- P_{RF} is the power consumed by the radio frequency transceivers for uplink and downlink. Its dependency on load is similar to P_{BB} [9].
- P_{OH} is the power consumed by active cooling, and the losses in DC-DC conversion, and main supply.

The resulting overall consumed power shows a dependency on transmit power (keeping load constant) which is well approximated by a linear function [9], [10].
A measure of system load (or utilization) of a BS is typically the fraction of BS resources (channels, or BS time) which are active. In what follows we model utilization as the average fraction of time in which the BS is transmitting. Coherently with these observations, in this paper we adopt the following analytical model for BS consumed power:

$$k_1 + U[k_2 + k_3(p - p_{min})] \tag{2}$$

where p is the transmit power, which we assume can be varied within an interval $[p_{min}, p_{max}]$. The component k_1 models the total power consumed when the BS is idle, which does not depend on load or transmit power (e.g. part of cooling, power amplifier consumption in idle state).
U is the utilization of BSs. In what follows, we model utilization as the average fraction of time in which the BS is transmitting, and we assume that, when transmitting, a BS is using the whole bandwidth at its disposal. When transmit power is higher than that required to achieve the threshold expected per-bit delay $\bar{\tau}^0$, we assume that BSs 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. Conversely, when transmit power is lower than that required to achieve $\bar{\tau}^0$, BSs increase their utilization (the fraction of active resources) in order to satisfy the user-level QoS requirements.
k_2U models that fraction of consumed power which depends on utilization, but not on transmit power (such as part of baseband and RF processing, when micro-sleep techniques are in use [10]). The third component, $k_3U(P_T - P_{min})$, models that fraction of consumed power due to the power amplifier which depends, at the same time, on the transmit power P_T and on the amount of active resources at the BS (such as the energy consumed by the power amplifier). Finally, as P_{OH} is, by its nature, directly proportional to the total consumed power, it is taken into account as a (constant) factor in all coefficients k_1, k_2 and k_3.

III. Modeling User Perceived Performance

In this section, we characterize the per-bit delay perceived by a typical best-effort user who is just beginning service, as a function of the main system parameters. In settings with dense small cells, it becomes crucial to model accurately the effect of interference on user perceived performance. In order to do so, it is essential to model its dependency on traffic patterns. In what follows, we propose a result which relates the main performance parameter of the system, i.e., the expected per-bit delay seen by a typical user, to the main system parameters, such as user density, BS density, and the target QoS value for the network, in terms of maximum expected per-bit delay.

Theorem 1 (Average BS utilization [4])**.** For a given density of users, of base stations, and a given share of voice users γ, the average base station utilization in the network $U(\bar{\tau})$ is given by

$$U(\bar{\tau}) = \left[1 + \gamma\left(R_0 \frac{\mu_H^{-1}}{\mu_H^{-1} + \mu_W^{-1}}\bar{\tau}^0 - 1\right)\right]\frac{\bar{\tau}}{\bar{\tau}^0} \tag{3}$$

where $\bar{\tau}$ is the per-bit delay perceived by a typical best effort user which is beginning service.

In order to take into account realistically the effects of interference on power consumption in a cellular access network, we present a result which relates the average interfering power for the typical user which is just beginning service.

Lemma 1. The average interfering power for the typical user arriving in the system at a distance r from the serving BS, can be approximated as

$$\bar{I}(r, k, \bar{\tau}) = U(\bar{\tau})\frac{p\lambda_b 2\pi r^{2-\alpha}}{k(\alpha - 2)} \tag{4}$$

where $\bar{\tau}$ is the expected per-bit delay.

Proof. The total received interfering power $I(r, k, t)$ at time t for a user at a distance r from its serving BS is a random variable, depending on the location of all other BSs in the plane, as well as on the specific traffic pattern at each interfering BS. Its expression can be written as

$$I(r, k, t) = \sum_{j \in \Phi(k)} p r_j^{-\alpha} u_j(t) \tag{5}$$

where $\Phi(k)$ is the set of interfering BSs in the plane. $u_j(t)$ is equal to 1 if the $j-th$ BS is active at time t. Given the random reuse policy we have assumed, $\Phi(k)$ is a Poisson point process, derived by thinning the process of BSs by a factor k, and by removing the BS serving the given user. The computation of the Palm expectation of $I(r,k,t)$, $\bar{I}(r,k,t)$, is complex, as $u_j(t)$ is also function of the specific BS layout. In what follows, we approximate such computation by assuming all BSs have the same mean utilization $U(\bar{\tau})$. Moreover, note that as a result of interference, the BS with the strongest SINR at the user location is not necessarily the closest one. However, in regimes of high attenuation due to propagation ($\alpha \geq 3$ or more, typical of urban settings) this is still a good approximation. Hence we assume users associate to the closest BS. As a result of these approximations, we can write $\bar{I}(r,k,t)$ with

$$\bar{I}(r,k,\bar{\tau}) = U(\bar{\tau}) \int_r^{+\infty} p \frac{\lambda_b}{k} s^{-\alpha} 2\pi s ds = U(\bar{\tau}) \frac{p\lambda_b 2\pi r^{2-\alpha}}{k(\alpha-2)} \tag{6}$$

where $\frac{\lambda_b}{k} 2\pi s ds$ is the mean number of interfering BSs in an annulus of radii s and $s+ds$, with $s \geq r$. □

Theorem 2 (Mean per-bit delay)**.** The mean per-bit delay $\bar{\tau}$ perceived by a typical best-effort user joining the system when the density of BSs is λ_b and the density of users is λ_u, and transmit power is p can be approximated as the unique solution of the following fixed point problem:

$$\bar{\tau} = \int_0^\infty f(r) \frac{e^{-\lambda_b \pi r^2} \lambda_b 2\pi r}{C(r,\bar{\tau})} \, \mathrm{d}r. \tag{7}$$

with

$$f(r) = \int_0^\infty \int_0^{2\pi} e^{-\lambda_b A(r,x,\theta)} \lambda_u x d\theta dx$$

$A(r,x,\theta)$ 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. $C(r,\bar{\tau})$ is the capacity to a user at a distance r for a reuse factor k, given by $C(r,\bar{\tau}) = (B/k)\log_2\left(1 + \frac{pr^{-\alpha}}{N_0 + \bar{I}(r,k,\bar{\tau})}\right)$, with $\bar{I}(r,k,\bar{\tau})$ given by (4).

For the proof, we refer to appendix A. Note that the expression of $\bar{\tau}$ from Theorem 2 is implicit due to the dependency of the interference on the mean BS utilization, which in turn depends on the mean per-bit delay $\bar{\tau}$ through (4).

IV. Derivation of the energy-optimal network configuration

In this section we present the formulation of the optimization problem, which provides, for a given BS energy model and mean user density, the energy optimal density of BSs and BS transmit power which satisfy the performance constraints. Given the expression of the expected per bit delay, and of the average utilization, the energy optimal BS density derives from solving the following optimization problem:

$$minimize_{\lambda_b,p} \;\; \lambda_b \left[k_1 + U(\lambda_b,\lambda_u,p)\left(k_2 + k_3(p - p_{min})\right)\right] \tag{8}$$

Subject to:

$$U(\lambda_b,\lambda_u,p) \leq 1 \tag{9}$$

$$\tau \leq \tau_0 \tag{10}$$

$$0 \leq \lambda_b \leq \lambda_{b,max} \tag{11}$$

$$p_{min} \leq p \leq p_{max} \tag{12}$$

The maximum BS density $\lambda_{b,max}$ is determined by practical, technical (backhaul technology) and economic considerations. The above problem is non-convex and nonlinear. However, being only a function of the two variables λ_b and p, it can be solved efficiently by exhaustive search.

V. Numerical evaluation

In this section, we apply our approach to investigate numerically the performance of different strategies for energy efficiency in cellular access networks.

A. System setup

Base stations work at a frequency of 1 GHz, and use a bandwidth of 10 MHz. We use a log distance path loss model, with path loss at a reference distance of one meter calculated using Friis equation, and with a path loss exponent $\alpha = 3$. Unless otherwise specified, we consider the case of pure best effort traffic (i.e. $\gamma = 0$), and $\lambda_{b,max} = \infty$. Note however that from the expression of BS utilization in Theorem 3 descends that varying γ has the same effect on the energy-optimal configuration as a change in the values of the coefficients of the BS energy model. Finally, in all scenarios we assume user density to vary within a given range $(\lambda_u^{min}, \lambda_u^{max})$. Unless otherwise stated, we set $\lambda_u^{max} = 0.1$ users per square meter.
We consider two different types of BSs. Namely, macro (with $p_{max} = 10$W), and small cell (with $p_{max} = 0.13$W) ([7], [9], [10]). With reference to Table I, for each BS type we consider two choices of parameters for the energy models. A first one reflects the behavior of the majority of present day installed BSs [7], [9], and it is characterized by a low/moderate load proportionality. A second choice of parameters, labeled load proportional or "LP", reflects the load proportionality achievable through time-domain duty-cycling, i.e. through micro-sleep techniques which deactivate some BS components for short time periods [7], [10]). These techniques, nowadays rarely implemented, make P_{BB} and P_{RF} dependent on load, in a way that increases the overall load dependency of both macro BS and small cell BS. As we can see, such "futuristic" energy models bring the load proportionality (i.e., the percentage of maximum total consumed power which depends on BS load) of both types of BS to about 90%.

B. Joint planning and management strategies

In order to assess the maximum energy savings achievable with joint network planning and management strategies, we consider the following parameters:

- λ_b^S is the BS density resulting as solution of the optimization problem in Section IV with $p = p_{max}$ and $\lambda_u = \lambda_{u,max}$; and

Parameter	Macro BS		Small Cell BS		
	Typ	LP	Typ	LP	CRAN
p_{min} [W]	0.1		0.001		
p_{max} [W]	10		0.13		
Max cons. power [W]	1500		20		10.6
Load proportionality (%)	60	90	29	90	90
P_{PA} (%)	60		29		54.7
P_{BB} (%)	13		47		0
k_1 [W]	600	150	14.2	2	1.06
k_2 [W]	0	450	0	12.2	3.54
k_3 [W^{-1}]	90.909		44.961		

TABLE I: Parameters of the energy model of Macro BS and Small Cell BS ([7], [9], [10]).

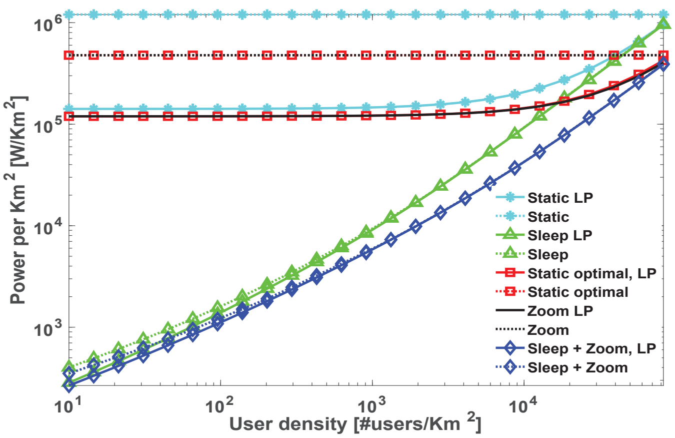


Fig. 1: Power per km^2 consumed with the five joint strategies, as a function of user density, for macro BSs.

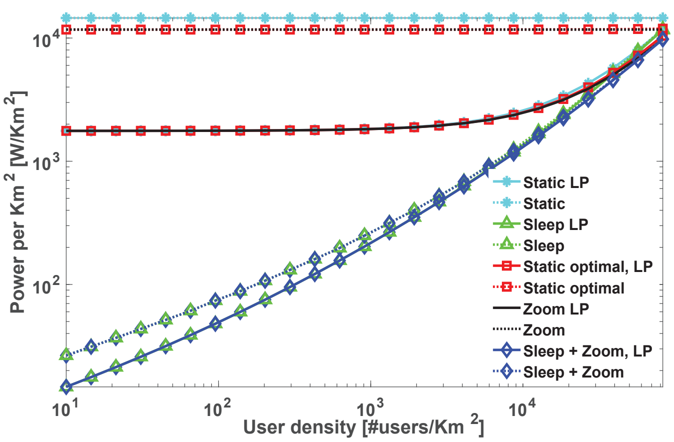


Fig. 2: Power per km^2 consumed with the five joint strategies, as a function of user density, for small cell BSs.

- λ_b^{SO}, p^{SO} are the BS density and transmission power resulting from solving the optimization problem in Section IV with $\lambda_u = \lambda_{u,max}$.

In what follows, and based on the structure of the considered energy model, we consider the following five main joint planning/management strategies:

1) Static: It is a network planning strategy with no dynamic network management. That is, both the density of active BSs

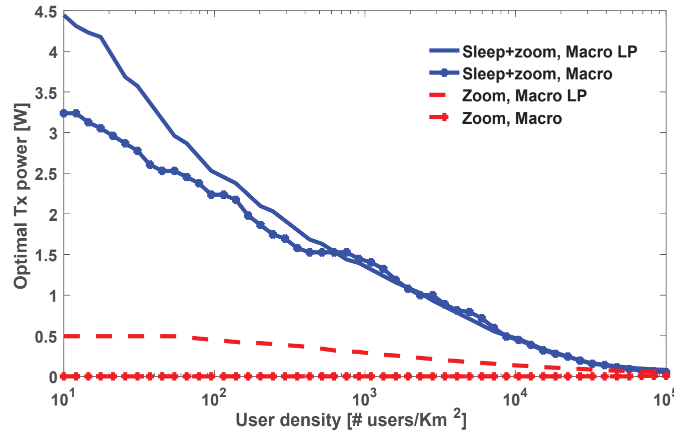


Fig. 3: Optimal transmit power vs user density, for macro BS.

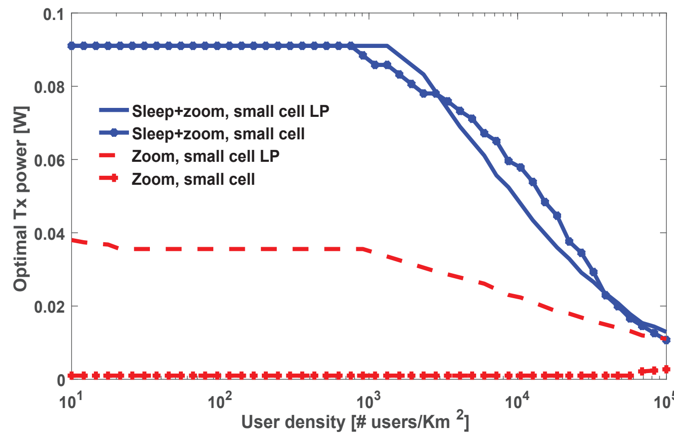


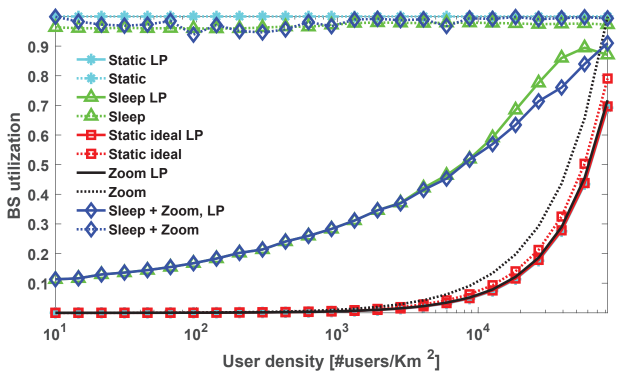
Fig. 4: Optimal transmit power vs user density, for small cell BS.

and the transmit power do not change while the network is operating. We assume transmit power to be equal to max transmit power, as it is typical in present day access networks [1]. As for the density of installed BS, we assume it is equal to λ_b^S.

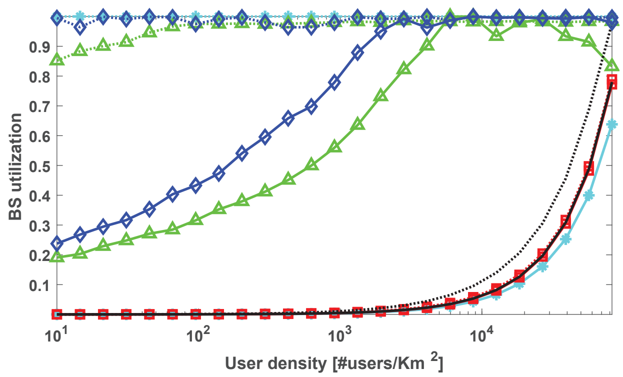
2) Sleep: In this joint strategy, the density of installed BS is λ_b^S. During network operation, TX power is kept constant at p_{max}, while the density of active BSs, for a given mean value of user density is derived from solving the problem in Section IV with $p = p_{max}$, at that value of user density.

3) Zoom: In this strategy, the density of active BSs does not vary over time, while their transmit power is varied in order to adapt energy consumption to load. We assume that in the planning phase, the derivation of the optimal density of installed BSs assumes that transmit power can be tuned during network operation. Hence, the density of installed BS is λ_b^{SO}. For any value of user density, the corresponding optimal transmit power is derived from solving the problem in Section IV over transmit power only.

4) Sleep+Zoom: In this case, the density of installed BSs is λ_b^{SO}. During network operation, for a given mean user density, both the density of active BSs and transmit power are set to be equal to the solution of the problem in Section IV for that user density.



(a) Small Cell BS



(b) Macro BS

Fig. 5: Optimal BS utilization vs user density.

5) Static Optimal: This is also a network planning strategy with no dynamic network management, like 1) above. The density of installed BSs is set to λ_b^{SO}. The value of transmit power, which is kept constant during network operation, is derived by solving the problem in Section IV only over transmit power, assuming $\lambda_b = \lambda_b^{SO}$, over all the interval $(\lambda_u^{min}, \lambda_u^{max})$, and taking the maximum over the interval. The underlying idea is to avoid dynamic network management, taking for both transmit power and for the density of installed BSs, the maximum values assumed by these quantities in the sleep+zoom strategy, for a given interval of user densities. This brings to a conservative configuration for the network, though less conservative than the purely static strategy.

Note that, being defined on a homogeneous spatial distribution of users and of BSs, and assuming BS density can be varied with continuity (with no switching costs and delays), these strategies are not directly applicable to a realistic scenario. They have been considered because they give an indication of the maximum energy savings achievable through a given network planning/management approach.

C. Assessment of the joint strategies

In a first set of evaluations, we have assessed the performance of the five strategies over a range of user densities of $(10, 10^5)$ users per km^2, as the typical values commonly considered in the planning phase usually fall within such an interval. In Fig. 1 and Fig. 2 we have plotted the power consumed per km^2 as a function of user density, for the five strategies. First, it is very important to note that the adoption of a small cell BS layout implies a reduction of the energy consumption of 1-2 orders of magnitude, in spite of the increase in the number of necessary BSs of almost one order of magnitude, as we will see later. Second, as expected, when micro-sleep modes at the device level are in use at BSs, the static strategy is already able to save energy with respect to present day BSs models. However, the explosion of traffic and the consequent network densification foreseen in future 5G networks will push for more energy-efficient solutions. Moreover, we can see that sleep modes, especially when combined with zooming, still enable the largest energy savings, especially at low user densities. Tuning the transmit power, whether combined with sleeping or not, saves energy with respect to the static strategy even at peak user density, as it allows to decrease interference levels and the power inefficiencies it induces.

As for the static optimal strategies, we can see that it performs almost identically as the zoom strategy, for all user densities and across all BS power models. Both these strategies are overperformed by sleep and sleep+zoom at very low user densities. However, for user densities close to the peak, the static optimal strategy is more energy efficient than both the pure static and the sleep strategies.

The configuration induced by each strategy can be better understood by Figs. 3 and 4, which show the optimal transmit power as a function of user density, and Figure 5, for the optimal BS utilization. The energy savings (of around 60% for macro BS) which zoom, zoom+sleep and static optimal strategies exhibit at peak user density are related to the fraction of total power consumed by a BS which is due to the power amplifier. Indeed, at high user densities, zooming acts by minimizing this contribution by reducing transmit power, while keeping a high BS utilization. When user density decreases, zooming acts by decreasing utilization while at the same time increasing transmit power.

Fig. 6 shows the optimal BS density for the given strategies, as a function of user density. As we can expect, the use of small cells implies much higher BS densities, with differences of up to one order of magnitude. In addition, we see that,

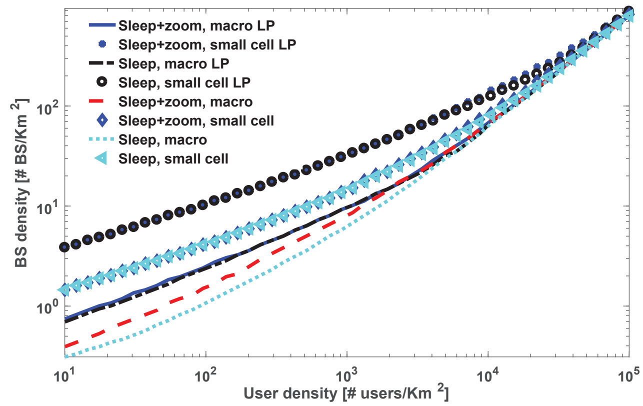


Fig. 6: Optimal BS density vs user density.

when zooming (with or without sleeping) is in place, or when the static ideal strategy is adopted, even the optimal density of installed BSs is decreased with respect to the static (or zooming) strategies. This happens because at high user densities the system is dominated by interference, in a way that increasing the transmitted power of all BSs has no impact on SINR. Zooming therefore acts by reducing the energy inefficiency due to interference, by decreasing both the density of installed BSs and the transmit power at peak user density. Overall, strategies which put BSs to sleep save energy when the user density is low (e.g. during night), while strategies which optimize over transmit power have the largest amount of energy saved during traffic peaks, i.e. when the power consumed by the network is at its highest. Bottom line, these trends suggest that among the most important factors in determining the best strategy in a given setting are the peak user density, the peak/valley ratio of the temporal traffic profile, and the amount of time in which the traffic is at high and low values.

D. Shanghai setup

As we have seen, the relative performance of the five strategies depends heavily on how traffic varies during the day. Hence, for a more realistic assessment, we have characterized the increase in energy efficiency enabled by the considered strategies with respect to the static strategy, in a setup in which daily patterns of user density are taken from a realistic scenario. [11] identifies five different types of cellular traffic profiles, each related to the main activities of the areas which generate it (office, residential, entertainment, transport, and mixed). We have assumed for each profile the same peak user density of one user per $10m^2$, and a mean per-bit delay of $100kb/s$. Fig. 7 shows the temporal traffic patterns, as well as the energy savings over 24 hours for a typical working day, with respect to the static policy, for the 4 policies and the 5 traffic patterns.

Moreover, in order to take into account the impact of cloud RAN approaches on our strategies, we have considered a cloud RAN small cell power model, derived from the small cell LP model, but in which the baseband processing is performed on a dedicated device instead of the BS (see Table I). We can first observe that, as expected, the sleep+zoom policy achieves very large savings (between 70% and 85%), which largely outperform the sleep strategy. Second, the zoom strategy, which does not involve BS duty cycling, has a better performance than the sleep strategy, in almost all settings, and all the BS types and configurations. Hence, it is possible to adopt strategies which avoid putting BS to sleep (with all associated issues relative to interference management, and to increase of fault rates, bringing to a decrease of service availability due to a shortening of BS lifetime [6]) without compromising on energy efficiency. Third, the static optimal strategy, which is based only on a network planning which is optimized according to our approach, has savings comparable to the zoom and sleep strategies, while avoiding all the complexities associated with dynamically managing network resources.

E. Model Validation

Finally, in order to assess the impact of the approximations introduced in the derivation of Theorem 2 on the accuracy of the model, we have simulated the system for a range of values of user density between 10 and 10^5 users per km^2, and for base station densities between 0.1 and 10^3 BS per km^2. Over these ranges of values, we have compared the measured per-bit delay with the value derived from Theorem 2. The simulation values have been measured with a 97% confidence interval of 3% of the mean value. As a result, the mean difference between simulation and analysis has always been within 5%, and never larger than 8%.

VI. Conclusions

In this paper we have presented an analytic framework for the assessment of the potential for energy efficiency of several joint network planning and dynamic management strategies. By including in our model for user perceived performance the effects of interference, our approach allows to realistically assess the potential savings of these strategies, and their relative performance as a function of the temporal traffic profile. Far from defining practical schemes for planning and management, our approach gives an indication as to which type of strategies has the largest potential for energy savings in a given urban scenario, and for a given temporal load profile. In this sense, our work is a first step towards the choice of practical strategies in a given scenario, as well as a reference for the evaluation of their performance.

References

[1] L. Budzisz *et al.*, "Dynamic Resource Provisioning for Energy Efficiency in Wireless Access Networks: A Survey and an Outlook," *Communications Surveys Tutorials, IEEE*, vol. PP, no. 99, June 2014.

[2] D. López-Pérez *et al.*, "Towards 1 Gbps/UE in cellular systems: Understanding ultra-dense small cell deployments," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2078–2101, 2015.

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

[4] B. Rengarajan, G. Rizzo, and M. A. Marsan, "Energy-optimal base station density in cellular access networks with sleep modes," *Computer Networks*, vol. 78, pp. 152–163, 2015.

[5] X. Weng, D. Cao, and Z. Niu, "Energy-Efficient Cellular Network Planning under Insufficient Cell Zooming," in *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*, May 2011, pp. 1–5.

[6] L. Chiaraviglio, M. Listanti, and E. Manzia, "Life is Short: The Impact of Power States on Base Station Lifetime," *Energies*, vol. 8, no. 12, pp. 14 407–14 426, 2015.

[7] C. Desset *et al.*, "Flexible power modeling of LTE base stations," in *WCNC*. IEEE, 2012, pp. 2858–2862.

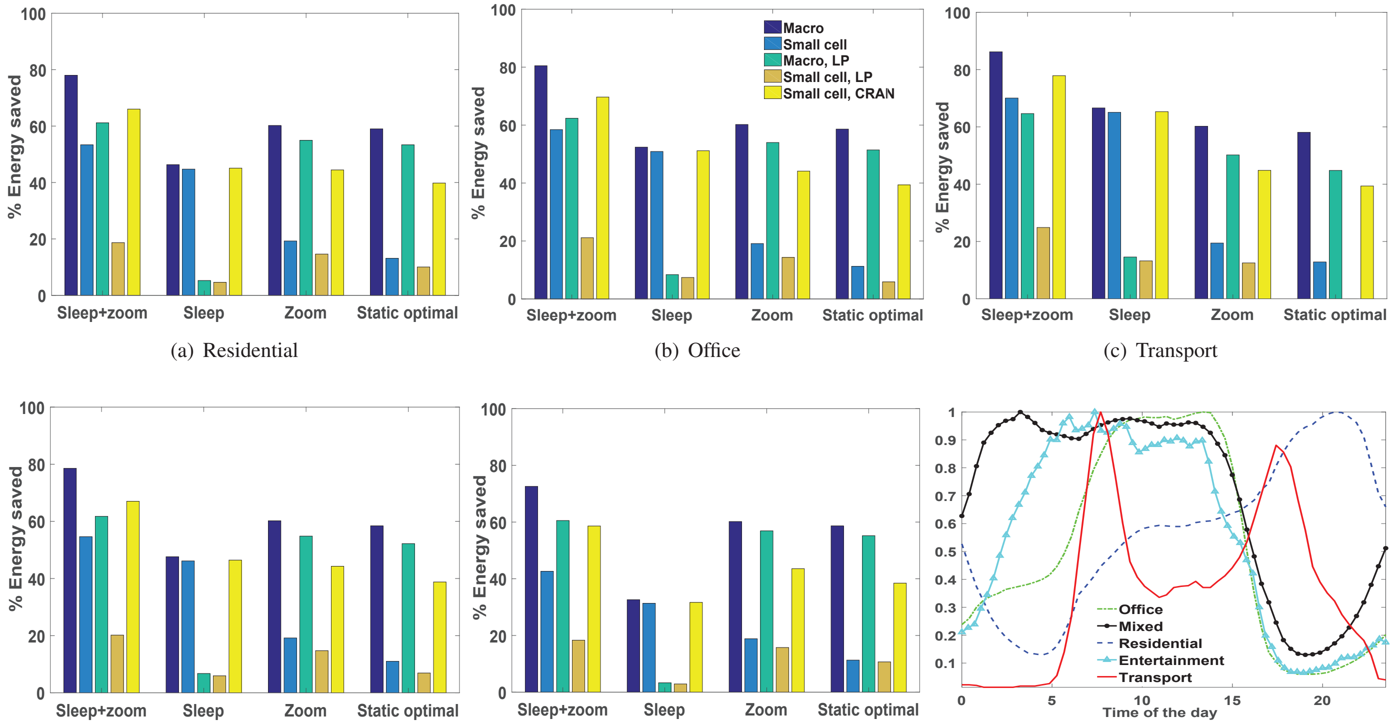
[8] F. Baccelli, B. Błaszczyszyn *et al.*, "Stochastic geometry and wireless networks: Volume II Applications," *Foundations and Trends in Networking*, vol. 4, no. 1–2, pp. 1–312, 2010.

[9] M. Imran *et al.*, "INFSO-ICT-247733 EARTH Deliverable D2.3: Energy Efficiency Analysis of the Reference Systems, Areas of Improvements and Target Breakdown," Tech. Rep., 2012.

[10] B. Debaillie, C. Desset, and F. Louagie, "A flexible and future-proof power model for cellular base stations," in *Vehicular Technology Conference (VTC Spring), 81st*. IEEE, 2015, pp. 1–7.

[11] F. Xu, Y. Li, H. Wang, P. Zhang, and D. Jin, "Understanding mobile traffic patterns of large scale cellular towers in urban environment," *IEEE/ACM Transactions on Networking (TON)*, vol. 25, no. 2, pp. 1147–1161, 2017.

[12] D. Blackwell, "Discounted dynamic programming," *The Annals of Mathematical Statistics*, vol. 36, no. 1, pp. 226–235, 1965.



(a) Residential (b) Office (c) Transport

(d) Entertainment (e) Mixed (f) Normalized traffic patterns

Fig. 7: Average energy saved over a typical working day by the considered strategies with respect to the static strategy, for five different temporal traffic profiles in the city of Shanghai [11]

Appendix

A. Proof of Theorem 2

Here we sketch the main steps of the proof. The derivation of the expression of the per-bit delay goes along the same lines as the proof of Theorem 3.1 in [4]. The main variants are the use of the palm expectation of $I(r,k)$, $\bar{I}(r,k,\bar{\tau})$, instead of $I(r,k)$, in order to make the derivation analytically tractable.

Lemma 2. For $p \le 20W$, and $\lambda_b \le 10^4\ BS/Km^2$, the fixed point in Theorem 2 admits a unique solution.

Proof. In order to prove that (7) has a unique fixed point, we have to prove that the operator $T(\bar{\tau})$, whose expression is given by the right member of (7), is a contraction. To do so, we verify that Blackwell's sufficient conditions for a contraction [12] hold for T. The monotonicity of T is straightforward, as with increasing $\bar{\tau}$ increases the mean BS utilization, and hence their interference. And this translates into an increase in per-bit delays computed in (7).

For the discounting property, we have to prove that $\exists \beta \in (0,1)$ such that $T(\bar{\tau}+a) \le T(\bar{\tau}) + \beta a$, $\forall a \ge 0$ and for all system parameter values for which $\bar{\tau}$ is defined. We have $U(\bar{\tau}+a) = U(\bar{\tau}) + Ka$, with $K = \left[1+\gamma\left(R_0 \frac{\mu_H^{-1}}{\mu_H^{-1}+\mu_W^{-1}}\bar{\tau}^0 - 1\right)\right]/\bar{\tau}^0$. Substituting into the capacity formula, and as $Ka \ge 0$, we have

$$(B/k)\log_2\left(1+\frac{pr^{-\alpha}}{N_0+I(r,k)+Kapr^{-\alpha}}\right) \ge$$

$$\ge (B/k)\left[\log_2\left(1+\frac{pr^{-\alpha}}{N_0+I(r,k)}\right) - \log_2\left(Kapr^{-\alpha}\right)\right]$$

Then we can write

$$T(\bar{\tau}+a) \le \int_0^\infty \left(\int_0^\infty \int_0^{2\pi} e^{-\lambda_b A(r,x,\theta)} \lambda_u x \, \mathrm{d}\theta \, \mathrm{d}x\right) \cdot$$

$$\cdot \frac{e^{-\lambda_b \pi r^2}\lambda_b 2\pi r}{C(r,\bar{\tau}) - (B/k)\log_2\left(Kapr^{-\alpha}\right)} \, \mathrm{d}r$$

We apply the Taylor series expansion of $\frac{1}{c-x}$ for $x \to 0$ to the fraction at the integrand, and we have, for $a \to 0$

$$\frac{1}{C(r,\bar{\tau}) - (B/k)\log_2\left(Kapr^{-\alpha}\right)} \ge \frac{1}{C(r,\bar{\tau})} + \frac{B\log_2\left(Kapr^{-\alpha}\right)}{kC^2(r,\bar{\tau})}$$

Now we have

$$(B/k)\frac{\log_2\left(Kapr^{-\alpha}\right)}{C(r,\bar{\tau})} \le Ka\frac{\log_2\left(pr^{-\alpha}\right)}{log_2\left(1+\frac{pr^{-\alpha}}{N_0+I(r,k)}\right)} \quad (13)$$

By substituting in the expression (6) for the interfering power, for $p \le 20W$, and $\lambda_b \le 10^4\ BS/km^2$, the right member in (13) is upper bounded by $K\beta a$, with $0 \le \beta < 1$. Hence $T(\bar{\tau}+a) \le T(\bar{\tau})(1+K\beta a)$, and as the utilization $U \le 1$, $T(\bar{\tau}+a) \le T(\bar{\tau}) + \beta a$, which proves the discounting property. Being $T(\bar{\tau})$ a contraction, by the Banach fixed point theorem, the fixed point problem in (7) admits a unique solution. □