Computer Networks 78 (2015) 152–163

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Energy-optimal base station density in cellular access networks with sleep modes

Balaji Rengarajan [a], Gianluca Rizzo [b,*], Marco Ajmone Marsan [c]

[a] Accelera, Inc., 3255 Scott Blvd 110, Santa Clara, CA 95054, United States
[b] IIG Institute, HES SO Valais, Technopole 3, 3960 Sierre, Switzerland
[c] Politecnico di Torino and Institute IMDEA Networks, Avda del Mar Mediterraneo 22, Leganes, Spain

ARTICLE INFO

Article history:
Received 18 March 2014
Received in revised form 15 September 2014
Accepted 1 October 2014
Available online 27 November 2014

Keywords:
Green networking
Cellular networks
Sleep modes

ABSTRACT

Sleep modes are widely accepted as an effective technique for energy-efficient networking: by adequately putting to sleep and waking up network resources according to traffic demands, a proportionality between energy consumption and network utilization can be approached, with important reductions in energy consumption. Previous studies have investigated and evaluated sleep modes for wireless access networks, computing variable percentages of energy savings. In this paper we characterize the *maximum* energy saving that can be achieved in a cellular wireless access network under a given performance constraint. In particular, our approach allows the derivation of realistic estimates of the energy-optimal density of base stations corresponding to a given user density, under a fixed performance constraint. Our results allow different sleep mode proposals to be measured against the maximum theoretically achievable improvement. We show, through numerical evaluation, the possible energy savings in today's networks, and we further demonstrate that even with the development of highly energy-efficient hardware, a holistic approach incorporating system level techniques is essential to achieving maximum energy efficiency.

1. Introduction

The ethical imperative to reduce their carbon footprint, combined with the financial realities of increasing energy costs, and the difficulties of network deployment in developing countries with unreliable power grids, has telecommunication network operators keenly interested in energy saving approaches.

In cellular networks, reducing the power consumed by base stations is, by far, the most effective mean to streamline energy consumption. As an example, in the case of UMTS, one typical Node-B consumes around 1500 W, and the multitude of these devices accounts for between 60% and 80% of the network's energy consumption [1,2], often representing the main component of an operator's operational expenditures.

Several international research projects have recently explored the possibilities for reducing energy consumption of base stations [3–5], since the classical assumptions that they can rely on access to a reliable supply of energy with acceptable cost are challenged in the networking context of today. While equipment manufacturers are working to produce more energy-efficient hardware [6], as we show, system-level approaches are called for, to obtain networks with the lowest possible energy consumption. Base stations are deployed according to dimensioning strategies that ensure acceptable user performance at peak (worst-case) traffic loads. However, traffic loads fluctuate

* Corresponding author.
E-mail addresses: balaji.rengarajan@acceleramb.com (B. Rengarajan), gianluca.rizzo@hevs.ch (G. Rizzo), ajmone@polito.it (M. Ajmone Marsan).

http://dx.doi.org/10.1016/j.comnet.2014.10.032
1389-1286/

throughout the day. For example, we expect diurnal patterns in the rate of user requests that mirror human patterns. Additionally, as the users of the network move during the day, they cause fluctuations in the spatial traffic load seen by base stations serving different locations. In [7,8], the possibility of reducing power consumption in cellular networks by reducing the number of active cells in periods of low traffic was considered, but the degradation in performance experienced by users in such a scenario, due to active base stations having to serve larger numbers of users that are located farther away from their serving base station was not explicitly taken into account. However, an important requirement for any energy saving measure, such as the introduction of sleep modes for base stations, is that they must be (almost) transparent to users. This means that the user-perceived performance must be above the target threshold at peak hours, when the load on the network is the highest, and all base stations are active, as well as in non-peak periods, when the load is lower, but the network is operating with reduced resources. In other words, the performance sacrifices that are implied by the introduction of energy-saving measures must be compatible with the target design objectives. Recently, several different approaches have been proposed to turn off base stations to conserve energy and to make the network energy consumption more proportional to utilization. For a very recent survey see [2]. However, to the best of our knowledge, the maximal energy savings that can be achieved under some predefined performance constraint was considered only in [9]. In this paper, we expand on the results in [9], and provide bounds on the minimum density of base stations required to achieve a given performance objective irrespective of the base station topology.

Our objective is to obtain a realistic characterization of the potential energy savings that can be achieved by sleep mode schemes under fixed user performance constraints, and study the impact of base station layout, power consumption model, and user density on the energy-optimal configuration of the access network. The metric we use to capture performance is the *per-bit delay* [10] (whose inverse is the short-term throughput) perceived by a typical user. The network is constrained to maintain, at all times, the average per-bit delay across users below a predetermined threshold. Our contributions are as follows:

- For a given base station layout, we develop a method for estimating the density of base stations that minimizes energy consumption and which is sufficient to serve a given set of active users, with fixed performance guarantees.
- For base stations whose power consumption is independent of load (not unlike current hardware), we derive a layout-independent lower bound on the density of base stations required to support a particular user density and thus an upper bound on energy saving.
- Through numerical evaluation, we compute bounds on the maximum energy saving, and illustrate the impact of various system parameters (user density, base station layout, target per-bit delay, base station energy model). We also assess the impact of user clustering and of correlation between user cluster locations and base station locations. We demonstrate that even with highly energy-efficient hardware, system level techniques are crucial to minimizing energy consumption. We find that the variability in performance across users is sufficiently low, validating the choice of the mean of the per-bit delay as a suitable metric for capturing user performance.

Our results are *bounds* with respect to what can be achieved in real networks, since we assume that *any base station density is achievable*, although this is clearly not possible in practice, since in real networks base stations can be turned off, but their locations cannot be rearranged according to traffic variations. The relevance of our bounds lies in that they indicate what are the theoretical minimum base station densities and energy consumption, allowing the effectiveness of different proposals to be measured against the maximum theoretically achievable improvement. With respect to [9], in this paper we consider a more general user traffic scenario, including both best-effort and constant bit rate services, we study the effect of base station sleep modes on the user terminal battery drain, and we investigate the impact of nonuniform layouts of users and base stations.

The paper is organized as follows. In Section 2, we present our model for the distribution of users and of base stations, and we state the main assumptions underlying our approach. In Section 3, we derive the average and the variance of the per-bit delay. In Section 4, we use the results of the previous sections to compute the energy-optimal base stations density for a given user density, and to estimate the achievable energy savings. Section 5 presents lower bounds on the base station densities required to satisfy the performance constraints. In Section 6, we present numerical results, and we conclude the paper in Section 7.

2. Model and assumptions

We mostly consider the downlink information transfer in a cellular access network, as typically it carries a larger amount of traffic than the uplink, and it has a larger impact on the energy consumption of the mobile network operator. However, we will also later verify the impact of our results on the uplink, by looking at the increase in the average distance between end user terminals and base stations, as well as at the end user terminal power consumption.

Users form a homogeneous planar Poisson point process, Π_u, with intensity λ_u users per square km, while base stations form a planar point process, Π_b, with density λ_b base stations per square km.

While the methodology introduced in this paper is quite general, and can be extended to many different base station configurations, we restrict ourselves to the following models for base station distribution across the service area:

- *Manhattan layout:* base stations lie on the vertices of a square grid, where the side of each square is $l_b = \frac{1}{\sqrt{\lambda_b}}$ km.