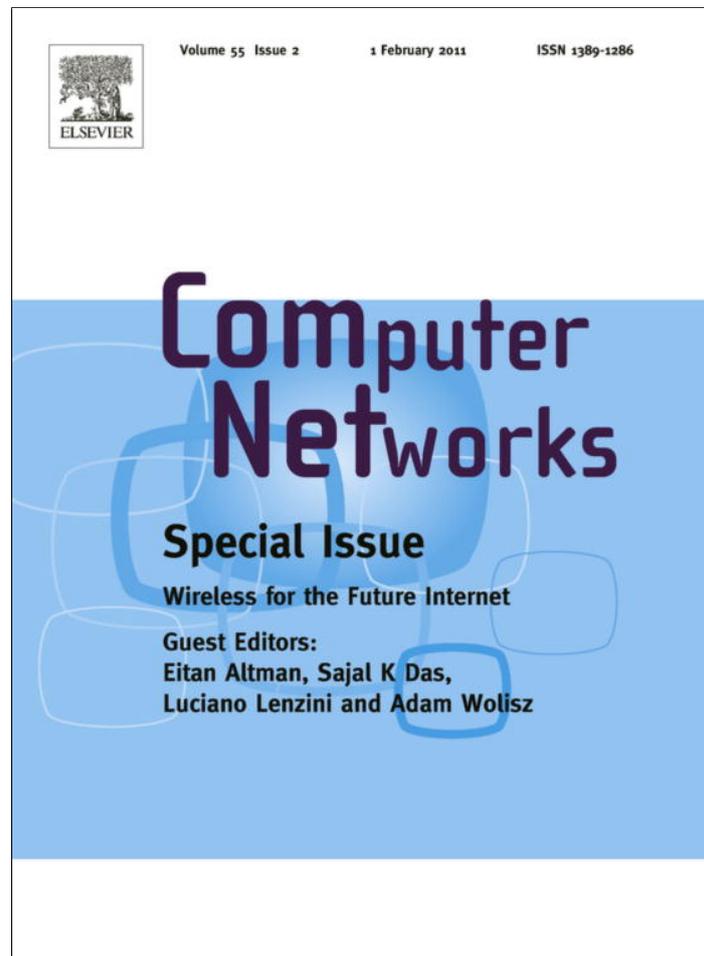


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

Computer Networks

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

Energy efficient wireless Internet access with cooperative cellular networks

Marco Ajmone Marsan^{a,b}, Michela Meo^{a,*}^aElectronics Department, Politecnico di Torino, Italy^bInstitute IMDEA Networks, Madrid, Spain

ARTICLE INFO

Article history:

Available online 27 October 2010

Keywords:

Green networking
Energy efficiency
Cellular systems

ABSTRACT

In this paper we study the energy-aware cooperative management of the cellular access networks of the operators that offer service over the same area. In particular, we evaluate the amount of energy that can be saved by using all networks in high traffic conditions, but progressively switching off networks during the periods when traffic decreases, and eventually becomes so low that the desired quality of service can be obtained with just one network. When a network is switched off, its customers are allowed to roam over those networks that remain powered on. Several alternatives are studied, as regards the traffic profile, the switch-off pattern, the energy cost model, and the roaming policy. Numerical results indicate that a huge amount of energy can be saved with an energy-aware cooperative management of the networks, and suggest that, to reduce energy consumption, and thus the cost to operate the networks, new cooperative attitudes of the operators should be encouraged with appropriate incentives, or even enforced by regulation authorities.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

The number of cellular accesses worldwide has recently surpassed the staggering number of 4 billion (and is expected to reach 4.6 billion by the end of 2009), with an average of more than 6 subscriptions every 10 people, and with a maximum in Italy, where each human being subscribes, on average, over 1.5 wireless network contracts [1,2]. This enormous number of wireless terminals is served by a huge number of base stations, of the order of 4 million worldwide. Each one of those consumes an amount of energy that can be estimated in about 25 MWh per year (the actual power consumption depends on the network technology, and is lower for 3G networks than for 2G networks; typical values range from less than 1 kW to about 3 kW [3]), so that the total energy consumption can be placed around 100 TWh per year. Assuming a unit cost of 0.1 euro (or US dollars) per kWh, this boils down to a cost of *ten billion* euro (or US dollars) per year (2500 euro or US dollars for each

base station). This massive cost represents a significant fraction of the network operational expenditures (OPEX), ranging from slightly less than 20% in mature European markets, to over 30% in emerging economies (like India). Obviously, these energy costs are transferred from network operators to us, the cellular network users.

In Italy, where four cellular access network operators compete, the number of base stations is of the order of 60 thousand, and their total energy consumption is around 2.1 TWh a year, close to 0.7% of the total national electricity consumption, which translates into an energy cost of about 300 million euro per year, or into 1.2 million tons of CO₂ injected into the atmosphere [4]. Quite similar considerations apply to other nations, for example Japan [5].

The energy bill of the wireless Internet access through cellular networks (we do not consider WiFi or WiMax) is bound to increase with time, due to a growing number of deployed base stations (experts predict about 5.5 million base stations in operation by 2013), increasing energy costs, and a 100% yearly growth rate of wireless Internet accesses, in spite of the fact that costs for wireless data are still high. Indeed, while not yet reaching a significant percentage, the wireless Internet access is growing much

* Corresponding author. Tel.: +39 011 5644167; fax: +39 011 5644099.

E-mail addresses: marco.ajmone@polito.it (M. Ajmone Marsan), michela.meo@polito.it, michela@polito.it (M. Meo).

faster than its wired equivalent, so that some forecasts predict that a vast majority of the accesses to the Internet of the Future will be wireless.

As we observed, the energy cost is one of the components of the unit cost that customers pay for network access, so that reducing the energy consumption of the network can translate into a lower bill for the end customers, further stimulating growth of wireless Internet accesses, as well as into a non-marginal reduction of OPEX, both of which can be welcomed in times of recession like those we are living.

This situation has caught the interest of networking researchers, and has stimulated the birth of an innovative research line, often called “Green Networking”. While energy efficiency has traditionally been an issue tackled by hardware designers and equipment manufacturers, the first area of networking that paid attention to energy consumption was represented by sensor networks, where the peculiarity of the network nodes made energy quite a significant element of the network design space. The first work introducing the energy issue in the mainstream networking area appeared at SIGCOMM 2003, and refers to the greening of the Internet [6]. The attention to energy issues in networking has been drastically rising in the last two years, when some specific meetings were organized to discuss the problem. Among the meetings planned for 2010, we can mention e-Energy 2010 [7], GreenMetrics 2010 [8], the First ACM SIGCOMM Workshop on Green Networking [9] and the Third IEEE International Workshop on Green Communications [10]. The European Commission has recently activated a new project, named EARTH, investigating the energy efficiency of mobile communication systems, within its seventh Framework Programme [11], and a European Network of Excellence, named TREND (Towards Real Energy-Efficient Network Design) will be activated soon. As an example of recent works on green networking, see [12–14].

There are two main motivations to focus on energy saving approaches for wireless access networks. The first is an economical motivation. In cellular networks, the mobile operator is a cost aggregator (with respect to energy consumption) to a much higher extent than in wired networks. Indeed, consider the overall energy consumption of a network, including the end users' terminals: in a mobile network, about 80–90% of the overall energy consumption (and thus of the energy cost) is in charge of one entity, the operator; instead, in a wired network, only about 30% of the overall energy consumption is in charge of the operator, the other 70% being distributed among end users, each user being responsible for a rather small amount of energy. Aiming at the objective of reducing consumption, it is easier to achieve significant results, and be effective, acting in a wireless context, and focusing on the access network, that is the main factor responsible of energy consumption (because of the combination of a large number of access devices – base stations – and the high energy consumption of these devices), rather than convincing the end users of a wired network to reduce their already small energy consumption. Moreover, for mobile operators, the energy cost of running their access networks is today of the order of their OPEX; thus, operators are keenly

interested in reducing their energy consumption. The second motivation concerns feasibility. Energy saving schemes, such as our proposals, based on switching on and off devices, that is, based on sleep modes, are particularly effective at the access network, i.e., at the network periphery, where the degree of traffic aggregation is low, so that the difference between peak and off-peak periods is large, and where the network is less vulnerable to possible failures, service discontinuity or degradation, since the number of affected users is limited. However, in wired networks, the resources used for access are not shared, so that access devices cannot be switched off without disconnecting the associated end users. In wireless access networks, the possibility of sharing access resources (i.e. of reusing the local loops) among end users allows a much higher degree of freedom at the access network to provide connectivity, thus allowing switch-off schemes and sleep modes.

In our first works on this topic [15–17] we focused on the energy-aware management of individual cellular access networks, estimating the amount of energy that can be saved by an operator that reduces the number of active cells in its own access network during the periods when they are not necessary, because traffic is low. The periodic reduction of the traffic in some portions of a cellular access network is due to both the typical day-night behavior of users, and the daily swarming of users carrying their mobile terminals from residential areas to office districts and back, resulting in the need for large capacity in both areas at peak usage times, but in reduced requirements during the periods in which the area is lightly populated (day for residential areas and night for office districts). The assumption that some cells in the access network can be switched off when traffic is low implies that radio coverage and service provisioning can be taken care of by the cells that remain active, which requires a, possibly small, increase in the emitted power to increase the size of the cells that remain on, and some adjustment in other network parameters, such as antenna tilting; still, some switch-off patterns may not be feasible, due to specific site positions that require some cells to be always on, to provide full coverage.

Considering that access networks are dimensioned based on the peak hour traffic, so that when traffic decreases due to normal traffic variations, networks are over-dimensioned, suggests a different, possibly more viable, approach, which is based on the fact that metropolitan areas are normally served by several competing operators, which provide coverage, and dimension their networks according to their number of subscribers. When traffic is high, the resources of each operator are exploited at capacity, resulting in the quality of service (QoS) used as a target design objective. When traffic is low, resources become redundant, and at some point just few, or even one, of the existing networks can carry all the traffic in the area with the desired QoS. Thus, if the operators cooperate, they can switch off their networks in turn, and save energy.

In a previous workshop contribution [18] we first investigated this option, assuming that just two operators are present in a metropolitan area, and that they use the same QoS parameter as a design target. In this paper, we extend the approach of [18] by considering the simultaneous

presence of more than two operators in a metropolitan area (as it normally happens), and by assuming that their networks are designed with different QoS targets (again a usual case).

More precisely, we consider a metropolitan area where n cellular network operators provide full coverage with separate infrastructures to populations of users with similar spatial distribution (so that the spatial distributions of traffic are similar across operators). The operators are willing to cooperate in order to save energy, by accepting the competitor's subscribers as roaming customers while their home network is switched off. In other words, in periods of decreasing traffic, operators progressively turn off their access networks, transferring their customers to operators that keep their network on, and that accept the roaming traffic of their competitors. Of course, by doing so, part of the energy required to power the access networks is saved, at the price of transferring some customers from their home network to another operator. This implies some subtle costs deriving, for example, by the information gathered by one operator about the profile of the competitors' customers, which can be highly sensitive; but, as our results will show, this cost is largely compensated by the amount of saved energy, that can be huge, and that translates into significant reductions of the network OPEX.

The technical complexity of the proposed approach is limited, since networks are already designed to carry roaming traffic from other operators' users. However, several operational details must be tackled to make the approach viable in practice, mainly referring to the transients related to the transfer of users from a network that is switched off to a network that remains on. If ongoing sessions (e.g. VoIP calls or video streaming) can be transferred through forced handovers, the impact of the resulting signalling load must be compatible with the signalling channel capacity. This problem can be alleviated by distributing roaming users over several networks and/or over time, as we shall discuss later in this paper.

The switch-off of a network should be made as transparent as possible to end users, and the possible inconveniences due to transients could be compensated by a reduction of tariffs (that should otherwise not change), at least for those users being inconvenienced, transferring to end users a portion of the operators' OPEX savings.

A nice feature of the energy saving approach that we consider is that it is not influenced by user mobility patterns, or cell layout. Indeed, by working at the network level, we must only guarantee that the total traffic in each network, or network portion, (due to its own customers plus the roaming customers) does not exceed the maximum traffic value for which the network was designed. If the network can carry such traffic when it is generated by its own customers, it can also carry the same amount of traffic when it is generated by a mixture of its own and roaming customers. Of course, an underlying assumption of this approach is that the spatial traffic distribution in the networks is not drastically different, so that resources are reusable across networks. If this is not the case, for example because an operator X has most of its customers in Rome, while an operator Y has most of its customers in Milan, when operator Y is switched off, operator X

cannot use its capacity in Rome to serve customers in Milan. However, this does not diminish the power of our approach, since we can cope with such cases by appropriately defining the service area: in the case above, we can define two service areas, one in the city of Rome, and the other in the city of Milan, so that when operator Y is switched off, operator X can use its capacity in Rome to serve roaming customers of Y in Rome, and its capacity in Milan to serve roaming customers of Y in Milan. Of course, this restriction on reuse of resources due to traffic locality leads to reduced gains with respect to the case of total reusability.

We study the proposed energy saving approach with simple analytical models, and we quantify the benefits that can be achieved with different switch-off patterns, under different traffic profiles and cost models. By showing that significant savings can be achieved, our results will hopefully provide a motivation for operators to cooperate, and for regulators to offer adequate incentives to cooperating operators, in order to improve the sustainability of cellular networks, and further stimulate the wireless access to a really ubiquitous Future Internet.

The rest of this paper is organised as follows. In Section 2 we describe the system we consider, introducing the network traffic profiles, the roaming policies, the switch-off patterns, and the energy cost models. In Section 3 we present numerical results for both a sinusoidal traffic profile and a measurement-based traffic profile, always assuming that the QoS target is the same for all operators. The case of different QoS targets is considered in Section 4, where some results for this case are also presented. Finally, Section 5 concludes the paper.

2. The system

We consider an area served by n operators, whose access networks fully cover the service area. The set of networks is denoted by $\mathcal{N} = \{1, 2, \dots, n\}$. Each access network is dimensioned according to the peak traffic demand of the operator's customers, so as to meet a specified QoS constraint. Initially, we assume that all access networks use the same QoS target. Later, in Section 4, we consider the case of different QoS's.

Denote by N_i the number of customers of operator i , and by $f_i(t)$, with $t \in [0, 2T]$ spanning over 24 h ($T = 12$ h), the daily traffic profile of network i . We assume that the daily traffic profile repeats periodically (i.e., we neglect the week-end effect on traffic, which is a conservative assumption as regards energy savings), and that the average per-user traffic in all access networks is the same, so that the overall traffic is proportional to the respective number of users¹:

$$f_i(t) = \alpha_i f(t) \quad (1)$$

with $\alpha_i/\alpha_j = N_i/N_j$, and $f(t)$ a periodic function that describes daily traffic fluctuations. Let f_{max} identify the maximum of

¹ The assumption of equal average per-user traffic in all access networks is actually not necessary for the analysis, and is introduced only to keep notation simple. We could as well assume that the average per-user traffic in network i is τ_i , so that the overall traffic in network i is proportional to $N_i\tau_i$.

function $f(t)$; $\alpha_i f_{max}$ is, thus, the maximum traffic that network i can carry without violating the QoS constraint. With no loss in generality, we assume $\alpha_1 < \alpha_2 < \alpha_3 < \dots < \alpha_n$.

In the next section, we introduce specific traffic profiles for the derivation of numerical results, but it can be noted that quite often real traffic profiles exhibit a periodical structure, with highs around mid-morning and mid-afternoon, and lows at night. In Fig. 1 we report a typical shape of traffic profiles, setting at $t = 0$ the peak traffic value. For the sake of simplicity, we mostly consider traffic profiles that are monotonically decreasing in $[0, T]$, and symmetrical with respect to T , but the extension to the asymmetrical case only requires some heavier notation. In the results section we also discuss one asymmetrical case.

We assume that a subset of the access networks can be switched off when the total traffic reduces to a level such that the networks that remain on can carry the entire traffic of all networks without violating the QoS constraint. Of course, (at least) one network must remain on all day to provide coverage.

We start by assuming that users can roam through any network, and that when some networks are switched off, their customers roam to the networks that remain on, with a probability proportional to the destination network size. This proportionality allows the whole amount of available resources to be exploited at best, reusing the excess capacity of all active networks to accommodate roaming customers. We call Roaming-to-All this roaming scheme. As we mentioned previously, this scheme can alleviate the problems arising in the switch-off transients, by distributing the signalling load over all networks that remain on. In addition, the signalling load burstiness can be reduced by distributing roaming requests over time.

Consider a network switch-off configuration in which the networks in the subset $\mathcal{N}_a \subset \mathcal{N}$ are powered on, while the remaining networks are off. This configuration is possible at time t , if:

$$f(t) \sum_{i \in \mathcal{N}} \alpha_i \leq f_{max} \sum_{i \in \mathcal{N}_a} \alpha_i, \quad (2)$$

where the left side of the expression represents the total traffic to be carried at time t , and the right side is the maximum traffic that the networks in \mathcal{N}_a can carry without violating the QoS constraint. Expression (2) defines the times during a 24 h period in which the configuration is feasible; we call this period the *switch-off zone* of the con-

figuration. In particular, the instant at which the switch-off becomes feasible is T^* , given by:

$$f(T^*) \sum_{i \in \mathcal{N}} \alpha_i = f_{max} \sum_{i \in \mathcal{N}_a} \alpha_i, \quad (3)$$

$$T^* = f^{-1} \left(\frac{f_{max} \sum_{i \in \mathcal{N}_a} \alpha_i}{\sum_{i \in \mathcal{N}} \alpha_i} \right). \quad (4)$$

Notice that the considered switch-off pattern is not feasible without violating the QoS constraint if the term in brackets, argument of $f^{-1}(\cdot)$, is smaller than the minimum value of $f(t)$.

In the simple case of $n = 2$, with $\alpha_1 < \alpha_2$, that is considered in Fig. 1, one network is switched off, while the other remains on. Network 1 can be switched off at T_1^* , and network 2 at T_2^* , with:

$$T_1^* = f^{-1} \left(\frac{f_{max} \alpha_2}{\alpha_1 + \alpha_2} \right), \quad (5)$$

$$T_2^* = f^{-1} \left(\frac{f_{max} \alpha_1}{\alpha_1 + \alpha_2} \right). \quad (6)$$

Due to the fact that $f(t)$ is monotonically decreasing in $[0, T]$, we have $T_1^* < T_2^*$, that is, the lower traffic network can be switched off for a longer time, as intuitively expected.

Let us now consider a different roaming scheme. As before, a network is chosen to remain on all day, while the other networks progressively switch off; but, differently from before, all the users of a network that is switched off roam to the only network that during that day never switches off. We call Roaming-to-One this roaming scheme. Consider, as before, a scheme in which the networks in $\mathcal{N}_a \subset \mathcal{N}$ are powered on, and the remaining $\mathcal{N}_o = \mathcal{N} - \mathcal{N}_a$ are off. Assume that network $l \in \mathcal{N}_a$ is the network that remains on all day; l receives all the roaming traffic from the networks that are off. In this case, the traffic constraint that defines if the scheme is possible becomes,

$$f(t) \sum_{i \in \mathcal{N}_o} \alpha_i \leq f_{max} \alpha_l, \quad (7)$$

that is, the scheme is possible if network l , which remains on, can carry, besides its own traffic, all the roaming traffic from the off networks. The switch-off time in this case is:

$$T^* = f^{-1} \left(\frac{f_{max} \alpha_l}{\sum_{i \in \mathcal{N}_o} \alpha_i} \right). \quad (8)$$

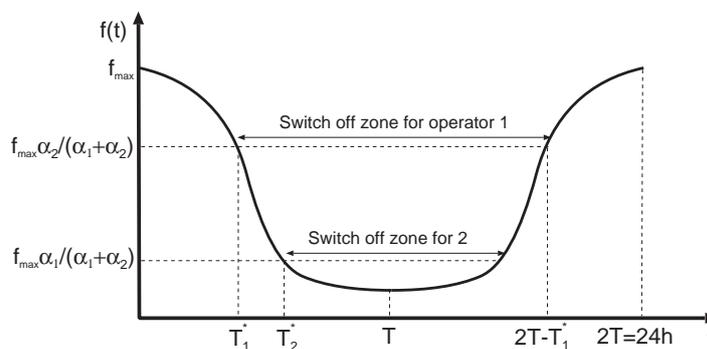


Fig. 1. Typical daily traffic profile $f(t)$, and possible switch-off periods for networks 1 and 2 (only one at a time can be switched off).

Clearly, the savings that can be achieved with this roaming scheme are smaller than with the Roaming-to-All scheme, since not all the capacity of the networks in \mathcal{N}_a is now available for roaming traffic coming from the networks that are off; the only available capacity is the one of network l , which remains on all the time.

Other roaming schemes are possible, and the associated switch-off times can be computed similarly.

2.1. Energy cost

Denote by $C(i)$ the energy cost of network i per unit time. The cost can be expressed as either the consumed energy in kWh, or a corresponding monetary cost. In general, $C(i)$ is given by the sum of two terms: one which is constant with respect to N_i (the number of network i users), and one which depends on N_i . Indeed, the energy cost of the backbone infrastructure and of the access network devices that provide complete radio coverage are more or less independent of the number of subscribers; on the contrary, the number of additional devices needed to provide the necessary capacity in the access network depends on the number of users. For simplicity, we will always consider the two extreme cases in which $C(i)$ is either constant (this case will be termed *same cost*), or directly proportional to N_i (this case will be termed *variable cost*).

For a given switch-off configuration, such as the one previously considered, in which the networks in $\mathcal{N}_a \subset \mathcal{N}$ are powered on, while the remaining networks are off, the daily energy cost can be computed as:

$$C = 2 \left[\sum_{i \in \mathcal{N}} C(i)T^* + \sum_{i \in \mathcal{N}_a} C(i)(T - T^*) \right], \quad (9)$$

since all the networks are on for a time T^* , while only the networks in \mathcal{N}_a are on in the period $T - T^*$; the factor 2 accounts for T being, by definition, half a day.

2.2. Roaming traffic

When network i switches off, its traffic must roam to the networks that are still powered on according to the considered roaming scheme. The switch-off of network i at time T^* generates a daily roaming traffic R_i equal to:

$$R_i = 2 \int_{T^*}^T \alpha_i f(t) dt, \quad (10)$$

where the factor 2 again comes from the fact that we consider 12-hour intervals $[0, T]$. This roaming traffic is directed to the active networks in \mathcal{N}_a according to the roaming scheme. For the Roaming-to-All scheme, the daily traffic roaming from network i to network $j \in \mathcal{N}_a$ is,

$$R_{i,j} = \frac{\alpha_j}{\sum_{k \in \mathcal{N}_a} \alpha_k} R_i. \quad (11)$$

For the Roaming-to-One scheme, the daily traffic roaming from i to j is,

$$R_{i,j} = \begin{cases} 0, & \text{for } j \neq l, \\ R_i, & \text{for } j = l, \end{cases} \quad (12)$$

(remember that l is the network that remains always on).

2.3. Switch-off patterns

We focus now on different switch-off patterns, indicating with this term the sequence according to which the networks are switched off, together with their switch-off instants. We assume that all but one networks switch off, in each 24 h period, provided enough capacity is available.

A switch-off pattern P is thus defined by

- $\{x_i, i = 1, \dots, n-1\}$ with $x_i \in \{1, 2, \dots, n\}$ – the sequence that specifies the order in which networks switch off; e.g., $x_i = k$ means that the i th network to switch off is network k . The network x_n never switches off in pattern P .
- $\{T_i, i = 1, \dots, n-1\}$ with $T_i \in [0, T]$ – the sequence of switch-off instants, i.e., network x_i switches off at time T_i .

The energy cost of pattern P , C_P , can be computed as:

$$C_P/2 = T_1 \sum_{i=1}^n C(x_i) + (T_2 - T_1) \sum_{i=2}^n C(x_i) + \dots + (T - T_{n-1})C(x_n), \quad (13)$$

$$= T_1 C(x_1) + T_2 C(x_2) + \dots + T C(x_n), \quad (14)$$

$$= \sum_{i=1}^{n-1} T_i C(x_i) + T C(x_n), \quad (15)$$

$$= \sum_{i=1}^{n-1} f^{-1} \left(\frac{f_{\max} \sum_{j=i+1}^n \alpha_{x_j}}{\sum_{j=1}^n \alpha_j} \right) C(x_i) + T C(x_n). \quad (16)$$

The optimal switch-off pattern is the one minimizing C_P .

In the *same cost* case, calling C the daily energy cost to power one network, according to (16), the cost of the switch-off pattern becomes:

$$C_P = 2C \sum_{i=1}^{n-1} f^{-1} \left(\frac{f_{\max} \sum_{j=i+1}^n \alpha_{x_j}}{\sum_{j=1}^n \alpha_j} \right) + 2CT. \quad (17)$$

Under this cost model, the optimal scheme corresponds to switching off the networks in order of increasing number of users; i.e., network i is switched off before network j if $\alpha_i < \alpha_j$. This implies that $x_j = j$ so that $\alpha_{x_j} = \alpha_j$.

To prove optimality, we need to show that setting $\alpha_{x_j} = \alpha_j$ minimizes the cost in (17). To do this, we neglect the additive and multiplicative constants, and focus on the minimization of the summation, that we rewrite as:

$$\sum_{i=1}^{n-1} f^{-1} \left(K \sum_{j=i+1}^n \alpha_{x_j} \right), \quad (18)$$

where K is a constant, equal to:

$$\frac{f_{\max}}{\sum_{j=1}^n \alpha_j}.$$

Setting $\alpha_{x_j} = \alpha_j$, we can rewrite (18) as:

$$\begin{aligned} \sum_{i=1}^{n-1} f^{-1} \left(K \sum_{j=i+1}^n \alpha_j \right) &= f^{-1} (K[\alpha_2 + \alpha_3 + \alpha_4 + \dots + \alpha_n]) \\ &\quad + f^{-1} (K[\alpha_3 + \alpha_4 + \dots + \alpha_n]) \\ &\quad + f^{-1} (K[\alpha_4 + \dots + \alpha_n]) + \dots \\ &\quad + f^{-1} (K[\alpha_{n-1} + \alpha_n]) + f^{-1} (K\alpha_n). \end{aligned} \quad (19)$$

If we alter the switch-off pattern, for example by switching off network 3 before network 2, the cost becomes:

$$f^{-1}(K[\alpha_2 + \alpha_3 + \alpha_4 + \dots + \alpha_n]) + f^{-1}(K[\alpha_2 + \alpha_4 + \dots + \alpha_n]) + f^{-1}(K[\alpha_4 + \dots + \alpha_n]) + \dots + f^{-1}(K[\alpha_{n-1} + \alpha_n]) + f^{-1}(K\alpha_n), \quad (20)$$

so that only the second term changes. Due to the fact that $\alpha_2 < \alpha_3$, the argument in the second term decreases, and, because of the monotonically decreasing behavior of the traffic profile, the term increases.

This is sufficient to prove that an alteration of the sequence of switch-offs consisting in a swap of two consecutive networks that violates the order of increasing number of users, produces a higher cost. This also means that, given a sequence, if a smaller network is switched off right after a larger network, swapping the two switch-offs produces a lower cost. A repeated iteration of swaps can thus be used (much like in the bubble sort algorithm) to obtain a complete ordering, with minimum cost.

From the computed energy cost, it is possible to compute the fraction of total energy saving for switch-off pattern P ,

$$G_p = 1 - \frac{C_p}{\sum_{i \in \mathcal{V}} C(i)2T} \quad (21)$$

and the fraction of energy saving of operator k that switches off as the i th one, i.e., such that $x_i = k$,

$$G_p(k) = 1 - \frac{2T_i}{2T} = 1 - \frac{f^{-1}\left(\frac{f_{max} \sum_{j=i+1}^n \alpha_{x_j}}{\sum_{j=1}^i \alpha_j}\right)}{T}. \quad (22)$$

3. Numerical results

In this section we discuss numerical results, starting from the case of sinusoidal traffic profiles. We then consider a traffic profile derived from measured traffic data, and we finally address the case of different QoS targets, in the case of just two operators, for the sake of simplicity.

3.1. Sinusoidal-like traffic profiles

In the derivation of numerical results we start by considering a simple daily traffic profile that follows a sinusoidal-like behavior expressed by:

$$f(t) = \frac{1}{2^b} [1 + \sin(\pi t/12 + \pi/2)]^b f_{max} \quad (23)$$

with $b = 1, 3$. The traffic profiles in the 12-hour period following the peak are shown in Fig. 2, assuming $f_{max} = 1$. It must be noted that with $b = 3$ the curve has steeper slope, and the average traffic is lower.

We consider an area served by $n = 4$ operators (as it happens in Italy) with parameters α_i given by:

$$\alpha_i = (1 - a) + \frac{ia}{4}, \quad (24)$$

where a is a parameter that we call network unbalance. When $a = 0$, all the networks carry the same amount of traffic; when $a = 1$, the network loads are quite different,

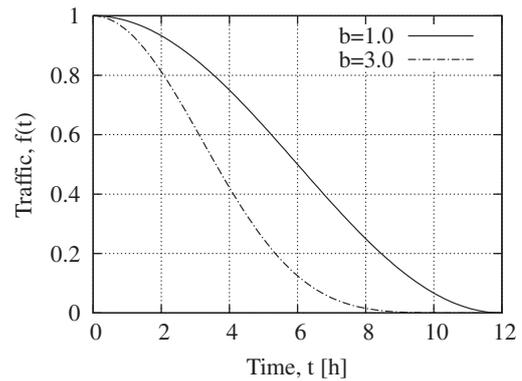


Fig. 2. Sinusoidal-like traffic profiles.

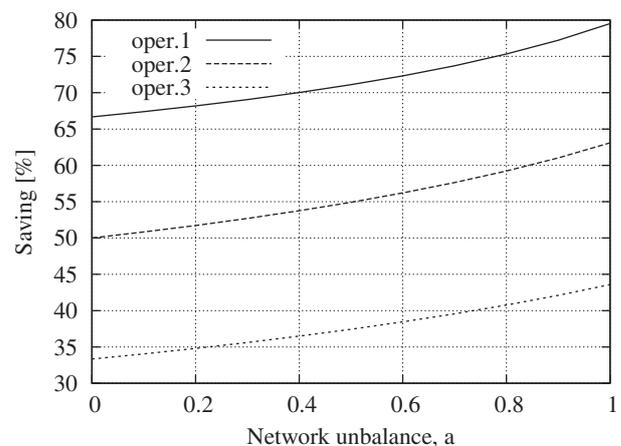


Fig. 3. Energy savings of individual operators for traffic shape parameter $b = 1$. Roaming-to-All scheme.

since the first network carries just one fourth of the traffic of the last one. To compare the unbalanced case with the same amount of total carried traffic, we also set f_{max} so that, in any configuration, the maximum amount of carried traffic sums to 4.

$$f_{max} = \frac{4}{\sum_{i=1}^4 \alpha_i} = \frac{4}{4 - 3a/2}. \quad (25)$$

We first consider a switch-off pattern such that the highest load access network (network number 4) is never switched off, while the other networks are switched off in increasing order of traffic, i.e., network 1, which is the one with the lowest traffic, is switched off first, network 2 is switched off second, and then network 3 follows. This switch-off pattern is optimal (as we proved) in the same cost case. Figs. 3 and 4 show the percentage energy savings that can be achieved by each operator for the traffic profiles with $b = 1$ and $b = 3$, respectively, in the case of the Roaming-to-All scheme. These percentage energy savings coincide with the percentage of time the networks are off, and they do not depend on the cost model, since we consider one network at a time.

The case $b = 3$ yields larger savings, because in this case the traffic profile decreases very rapidly, and the total amount of traffic is smaller. The sharp decrease allows network to be switched off early, shortly after the peak traffic,

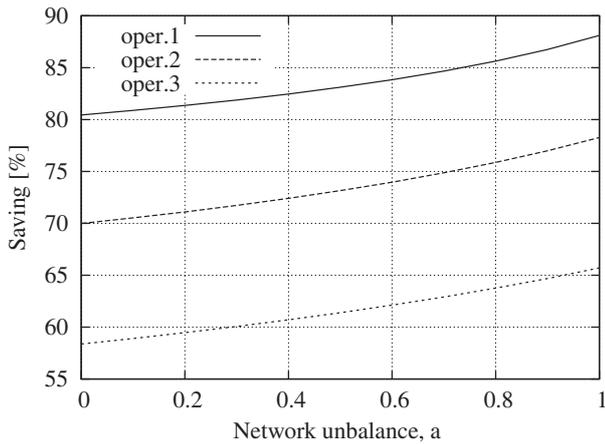


Fig. 4. Energy savings of individual operators for traffic shape parameter $b = 3$. Roaming-to-All scheme.

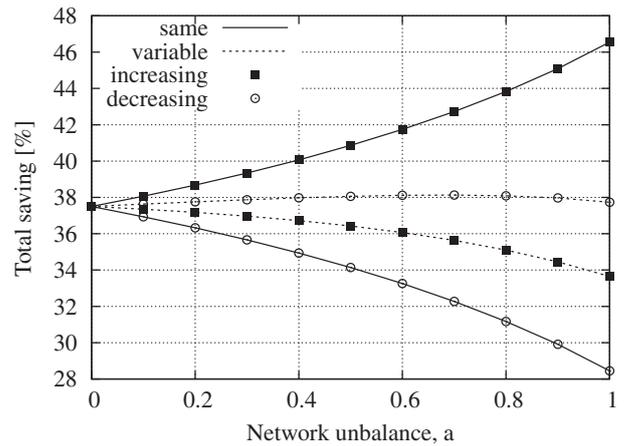


Fig. 6. Total saving under same and variable cost models for traffic shape parameter $b = 1$ with increasing and decreasing switch-off patterns. Roaming-to-All scheme.

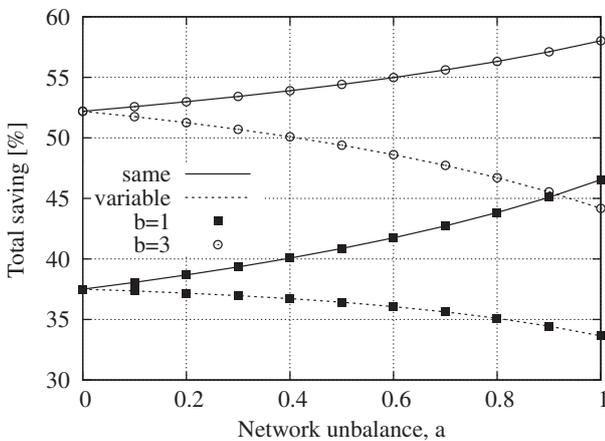


Fig. 5. Total saving under same and variable cost models for traffic shape parameter $b = 1,3$. Roaming-to-All scheme.

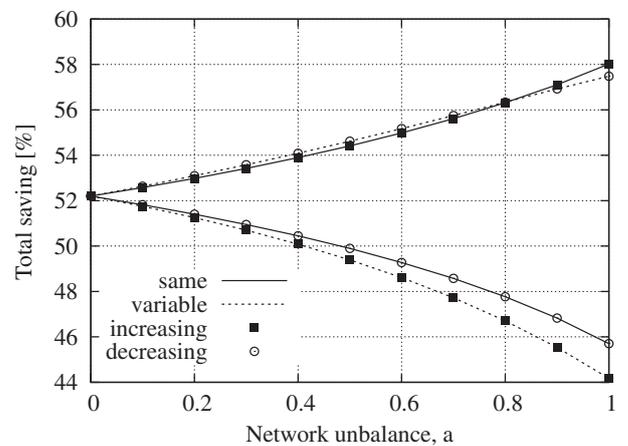


Fig. 7. Total saving under same and variable cost models for traffic shape parameter $b = 3$ with increasing and decreasing switch-off patterns. Roaming-to-All scheme.

and this leads to large savings. Being switched off first, operator 1, which carries the least traffic, saves the largest amount of energy. The more unbalanced the network traffic loads are, the better it is from the point of view of the saved energy. Indeed, the highest traffic network, which is not switched off, is the one with the most capacity, and can easily carry the traffic of the other networks, specially if it is low. It is worth noting that for high traffic unbalance the energy savings for low traffic networks can be huge, of the order of 80–90%.

Fig. 5 reports the total percentage energy saving, considering all networks, as in (21). Solid lines refer to the case in which all networks have the same cost model, while dashed lines refer to the variable cost model. Both the traffic profiles with $b = 1$ and $b = 3$ are considered. While under the same cost model the total saving increases with the network unbalance (coherently with the previous figures), under the variable cost model the total percentage energy saving decreases with the network unbalance. This is due to the fact that the longer time spent off by small networks does not compensate the higher cost associated with the power consumed by larger networks. Again, the case

$b = 3$ leads to higher savings, because of the steeper slope of the curve and the smaller daily traffic volume.

To discuss the impact of switch-off patterns, we compare the scheme where networks are switched off in order of increasing number of users, which is optimal for the same cost case, and a scheme in which networks are switched off in the reverse order, i.e., starting from the highest traffic network to the lowest traffic network, which is never switched off (this case is called decreasing). Figs. 6 and 7 show the total percentage energy saving for traffic profiles with $b = 1$ and $b = 3$, respectively. Let us first focus on the same cost model case (solid lines). The gap between the two switch-off patterns is remarkable, as expected, the increasing pattern being optimal. Interestingly, under the variable cost models, the decreasing pattern is instead more convenient. With a sinusoidal traffic profile with parameter $b = 1$, the difference between the two alternatives is much larger in the same cost case than in the variable cost case. On the contrary, with traffic shape parameter $b = 3$, the difference is very similar, and, quite surprisingly, the total percentage energy saving obtained

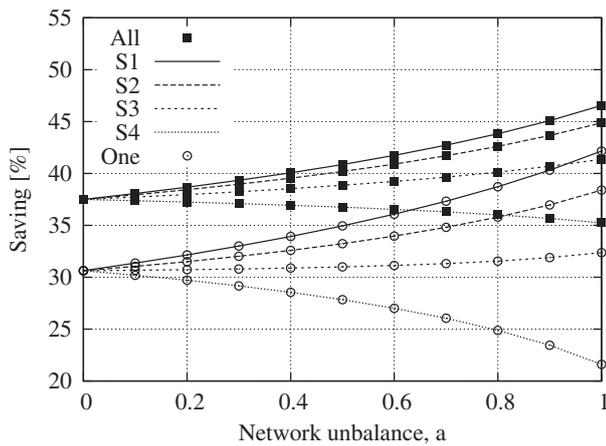


Fig. 8. Saving for different switch-off patterns under the same cost model for traffic shape parameter $b = 1$ with different roaming schemes.

for the increasing (optimal) pattern in the same cost case is almost identical to the saving obtained with the decreasing pattern in the variable cost case. These surprising results indicate that a careful evaluation of the possible approaches is necessary, in order to select the most effective options.

We now consider just the same cost model to compare the Roaming-to-One and Roaming-to-All schemes. Since a cooperation agreement among operators might reasonably impose that all the networks periodically remain on for the whole day (so that every operator has the opportunity to periodically carry other operators' roaming traffic), we consider four switch-off patterns obtained in the following way. We first choose the network that remains on the whole day, and let the other 3 networks switch off in order

of increasing size, so as to guarantee that the energy saving is maximized, given the choice of the network that remains always on. The resulting switch-off patterns are the following:

Pattern	Switch-off order	Always on
S1	1,2,3	4
S2	1,2,4	3
S3	1,3,4	2
S4	2,3,4	1

The savings that can be achieved under these switch-off patterns are reported in Fig. 8 for the traffic profile with $b = 1$; results are shown for both the Roaming-to-One (labeled 'One' in the figure) and Roaming-to-All (labeled 'All') approaches.

When all networks have the same size (i.e., for network unbalance $a = 0$) all switch-off patterns are (obviously) equivalent, and achieve the same energy saving. In unbalanced conditions, the energy savings increase with the size of the network that remains always on, S1 being, as previously discussed, the optimal switch-off pattern. The advantage in terms of energy saving of using the Roaming-to-All scheme is quite remarkable, as expected, since Roaming-to-All allows the capacity of all the networks that are powered on to be fully exploited, while Roaming-to-One must rely only on the capacity of the only network that remains on all day.

The switch-off instants for each operator, under all the switch-off patterns S_i are reported in Fig. 9; solid lines refer to the Roaming-to-All scheme and dashed lines to the Roaming-to-One scheme. Clearly, the lower the switch-off instant

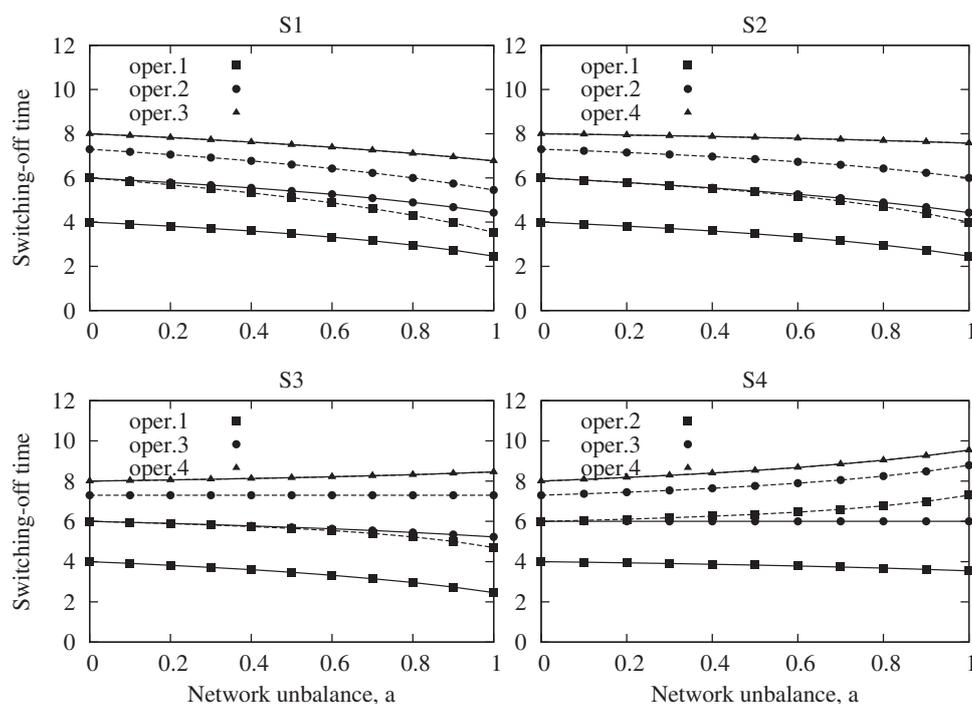


Fig. 9. Switching-off time (in hours) for each operator under the different switch-off patterns for the Roaming-to-All scheme (solid lines) and Roaming-to-One scheme (dashed lines); traffic shape parameter $b = 1$.

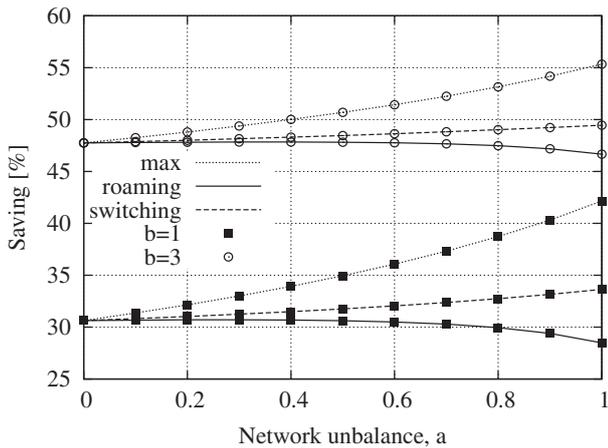


Fig. 10. Comparison between energy savings for the same roaming and same switch-off patterns, under the same cost model for traffic shape parameter $b = 1,3$ and Roaming-to-One.

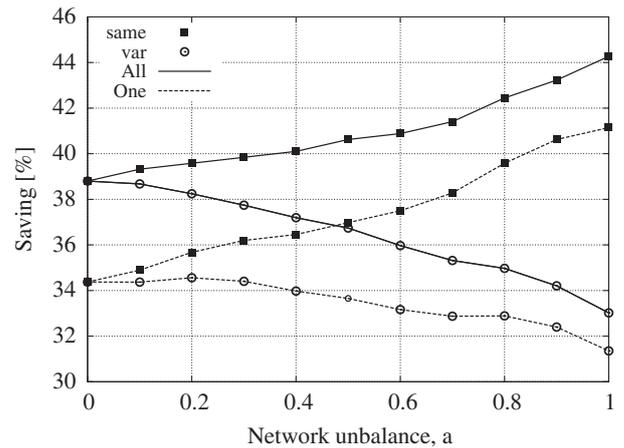


Fig. 12. Real traffic profile: total saving under same and variable cost models with increasing switch-off pattern.

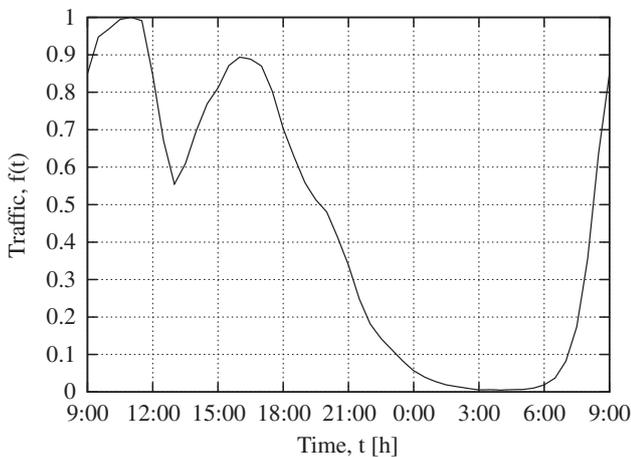


Fig. 11. The real normalized traffic profile.

is, the longer the off time is, and the higher the energy saving is. As expected, the Roaming-to-One scheme tends to have higher switch-off instants, the worst case being S4, in which the smallest network, i.e., network 1, remains on all the time, and is in charge of all the roaming traffic.

The fact that switch-off times tend to decrease for growing unbalance when large networks remain on (e.g., in switch-off patterns S1 and S2), while they tend to increase when smaller networks remain on (S3 and S4) is what is intuitively expected. The fact that variations with unbalance are not very high is quite encouraging from the point of view of the feasibility of the approach.

Notice that all the switch-off configurations are possible in the considered scenario, due to the fact that the traffic profile goes to zero during the night. For some traffic profiles with non negligible traffic during the night, some configurations might result not to be feasible.

The switch-off patterns considered above can be combined to form more general schemes. Let p_i be the fraction of time that pattern S_i is adopted with $i = 1, \dots, 4$ and $\sum_{i=1}^4 p_i = 1$. The total energy saving in this case can be derived by a weighted sum of the savings of the patterns S_i ,

$$G = \sum_{i=1}^4 p_i G_{S_i}. \quad (26)$$

When all the p_i 's are the same, we obtain a switch-off balance among operators, such that each operator has the same switch-off frequency. By combining the proposed schemes with properly selected values of the p_i 's, it is possible to achieve other balance objectives. As an example, we consider a *roaming balance* in which each operator has zero roaming cost, i.e., its outgoing roaming traffic is equal to the roaming traffic it has to carry when the other networks are off. Fig. 10 compares the savings achieved under the switch-off balance and roaming balance; as a reference, the case leading to the maximum energy saving is also reported. The constant cost model and the Roaming-to-One scheme are adopted; both the $b = 1$ and $b = 3$ cases are reported. While not being optimal, all the proposed switch-off patterns provide large savings with only a limited reduction with respect to the maximum. Again, this is quite a promising characteristic of the proposed approach.

3.2. Measurement-based traffic profile

As a second scenario, we now consider a more realistic traffic profile, derived from true traffic data collected in the network of an Italian operator. Fig. 11 reports the measured daily traffic profile of the operator, normalized so that $f_{max} = 1$. Observe that, differently from the previous case, this traffic profile is asymmetric with respect to T .

Fig. 12 reports results for the switch-off pattern where networks are switched off in order of increasing size, for the same and variable cost models, with the Roaming-to-One and Roaming-to-All schemes. The case of the switch-off pattern where networks are switched off in order of decreasing size is considered in Fig. 13. For the value $a = 0$, since the four networks are equal, savings are the same, regardless of the cost model. The increasing size switch-off pattern is better for same cost at increasing unbalance. It is interesting to see that also in this case the decreasing size switch-off pattern yields larger savings for variable cost. Observe also that the amount of energy

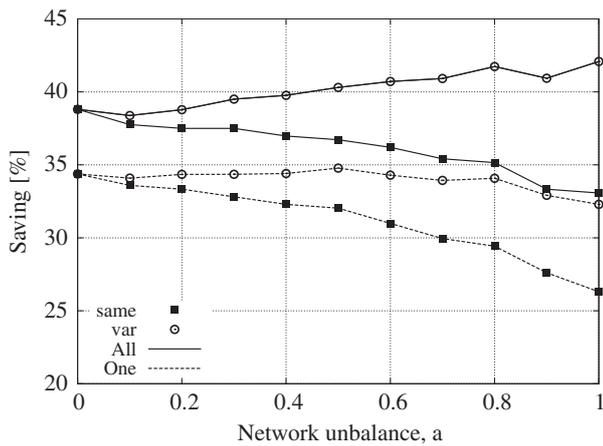


Fig. 13. Real traffic profile: total saving under same and variable cost models with decreasing switch-off pattern.

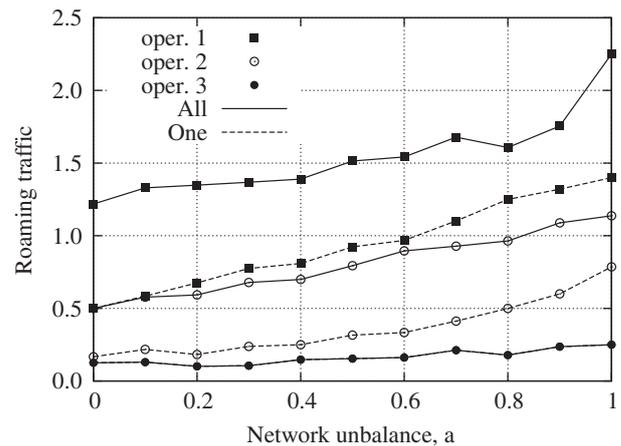


Fig. 15. Real traffic profile: outgoing roaming traffic with increasing switch-off pattern, Roaming-to-All and Roaming-to-One schemes.

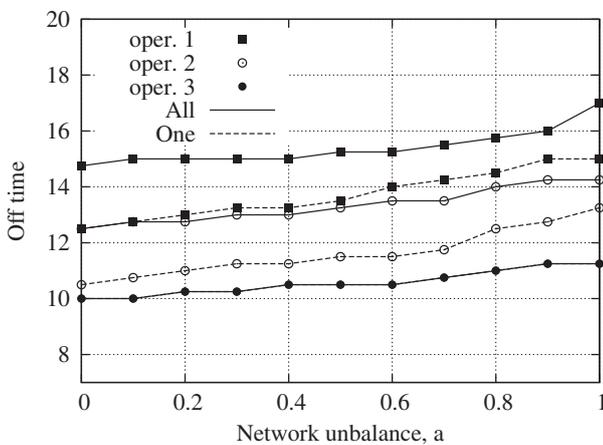


Fig. 14. Real traffic profile: total time (in hours) spent off with increasing switch-off pattern, Roaming-to-All and Roaming-to-One schemes.

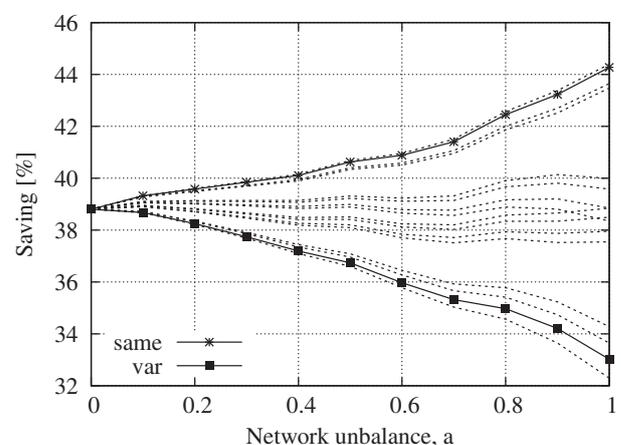


Fig. 16. Real traffic profile: total saving under mixed cost scenarios, increasing switch-off pattern, Roaming-to-All scheme.

saving for this realistic scenario is remarkably always above 25%, and can be more than 40%.

To investigate further the increasing size switch-off pattern, Fig. 14 shows the total time (in hours) spent off by all the operators (except for the fourth one, that remains on all the time) in the same cost model. As expected, networks of operators 1 and 2 can be switched off for a longer period of time in the Roaming-to-All scheme, since the unused bandwidth of all active networks can be used for roaming customers. Fig. 15 shows the total amount of outgoing roaming traffic for the three networks that switch off. The different effectiveness of the Roaming-to-One and Roaming-to-All schemes is reflected here also by the different amount of roaming traffic: by switching off later, under the Roaming-to-One scheme, each operator roams a smaller amount of traffic.

The curves of operator 3 overlap in both figures, because operator 3 is the last to switch off, and in both cases it must transfer its customers to operator 4, which after the switchoff will be in charge of all customers.

Finally, we consider the case in which the operators have different cost models, meaning that while some operators have the same cost to run their networks,

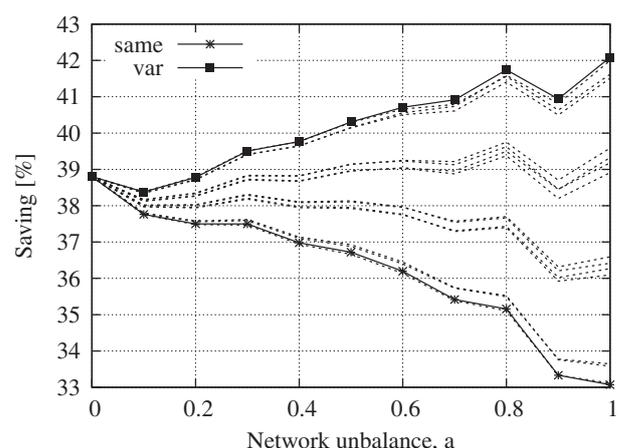


Fig. 17. Real traffic profile: total saving under mixed cost scenarios, decreasing switch-off pattern, Roaming-to-All scheme.

independently of their number of subscribers, other operators have a cost which is proportional to the number of their subscribers. We assume that in the two cases considered before (as well as in the new cases), where either all

the networks have the same cost model, or all the networks have the variable cost model, the total cost for running the four networks is the same.

Figs. 16 and 17 report the energy saving that can be achieved in the 16 possible mixes of cost models, for the increasing and decreasing switch-off patterns, respectively. The Roaming-to-All scheme is considered. The cases of all networks with same or variable cost models are emphasized with markers. These results suggest that the saving that can be achieved is significant, even under complex cost structures that can be somehow derived from the simple *same* and *variable* cost models previously described.

4. Different QoS targets

Consider now the case of access networks which are dimensioned according to their peak traffic demands, so as to meet different QoS constraints. For the sake of simplicity, we will discuss the case of only two access networks, but the extension to a larger number of networks is trivial.

Let Q_1 and Q_2 denote the QoS targets in the two access networks (respectively, network 1 and network 2). For example, Q_i could be blocking probabilities, in the case of circuit switched networks, or average delays, in the case of packet switched networks. The QoS targets depend on the network characteristics (bandwidth B_i , number of circuits, ...), on the service characteristics (average μ^{-1} and variance of the service duration, ...), and on the service request arrival rate.

When a network is switched off, and its traffic roams to the other, it becomes impossible (or unreasonably complex) to differentiate the QoS offered to the subscribers of the two operators, so we assume that the better QoS is guaranteed to all users.

Consider the case with $\alpha_2 > \alpha_1$ (the second network is the larger one) and $Q_2 < Q_1$, i.e., the QoS constraint on network 2 is tighter. If network 1 is switched off at time T_1^* , and its customers roam to network 2, since the QoS constraint Q_2 is more stringent than Q_1 , the only constraint is that network 2 can carry the traffic of network 1:

$$f(T_1^*)(\alpha_1 + \alpha_2) = f_{max}\alpha_2. \quad (27)$$

Instead, if network 2 is switched off at time T_2^* , and its customers roam to network 1, it is necessary to guarantee that network 1 can carry its own traffic, as well as the traffic of network 2, with the tighter QoS Q_2 rather than Q_1 .

Assuming for simplicity that the QoS is expressed as the average delay in a M/M/1 queue, so that:

$$Q_1 = \frac{1}{\mu B_1 - \alpha_1 f_{max}}, \quad (28)$$

$$Q_2 = \frac{1}{\mu B_2 - \alpha_2 f_{max}}, \quad (29)$$

it is necessary to guarantee that:

$$f(T_1^*)(\alpha_1 + \alpha_2) = f_{max}\alpha_1\beta, \quad (30)$$

where $\beta < 1$ is defined by:

$$Q_2 = \frac{1}{\mu B_1 - \beta\alpha_1 f_{max}}, \quad (31)$$

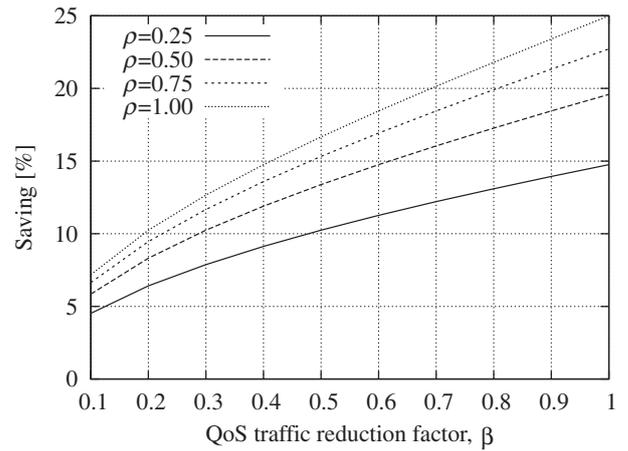


Fig. 18. Two networks with different QoS level. Total energy saving when switching off the smaller network, for variable QoS traffic reduction factor β . Sinusoidal traffic profile.

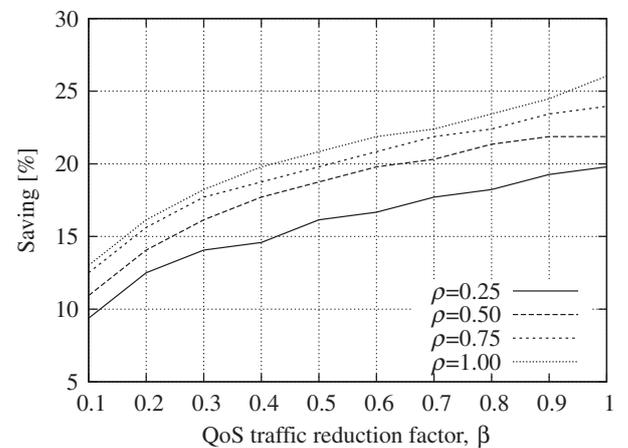


Fig. 19. Two networks with different QoS level and real traffic profile. Total energy saving when switching off the smaller network, for variable QoS traffic reduction factor β . Sinusoidal traffic profile.

so that:

$$\beta = \frac{\mu(B_1 - B_2) + \alpha_2 f_{max}}{\alpha_1 f_{max}}. \quad (32)$$

Eq. (30) suggests that a tighter QoS constraint translates into a more stringent constraint on the roaming traffic that a network can carry, and this also means a reduction of the time for which a network can be switched off.

As an example, consider the case of two networks with sinusoidal traffic profile with $b = 1$, same cost model, and in which the ratio of the network sizes is given by ρ ,

$$\rho = \frac{\alpha_1}{\alpha_2}. \quad (33)$$

Since we assume, as usual, that the second network has a larger number of users, $\rho < 1$.

Fig. 18 shows the energy saving that can be achieved when switching off the larger network (network 2, whose QoS constraint is more stringent), given different values of β as in (32). Observe that tight values of the QoS index of network 2 correspond to small values of β , and to small

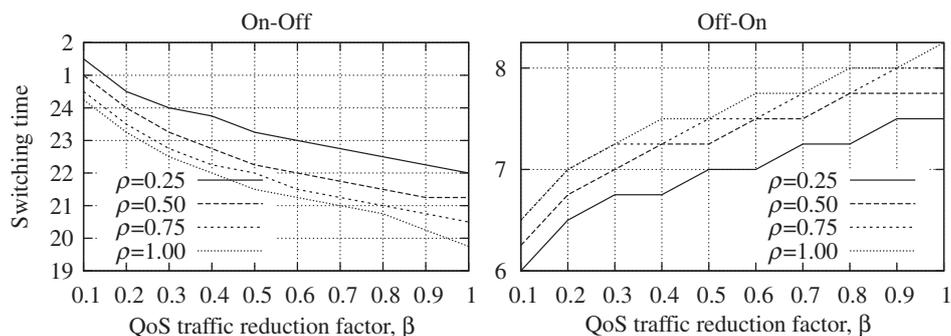


Fig. 20. Two networks with different QoS level and real traffic profile. Switching time from On to Off (left) and from Off to On (right) when switching off the smaller network, for variable QoS traffic reduction factor β .

values of the possible saving. Indeed, we consider the most critical case, where the larger network (with more stringent QoS) is switched off, and its traffic must roam to a smaller network, which in addition to carrying the roaming traffic, must also tighten its QoS constraint. Again, the larger the difference in network size (small values of ρ), the smaller the possible energy savings.

Similar conclusions can be drawn from the results reported in Fig. 19 for the real traffic profile; the considered scenarios are similar to the previous case. Observe that savings are of the same order as in the previous case: the still large values of the possible savings for very small values of β is due to the fact that the traffic becomes negligible during night and even extremely tight QoS constraints can be met during night hours.

The switch-off instants in the evening, and the switch-on instants in the morning, are reported in Fig. 20 for the latter scenario. It can be observed that small values of the size of network 1 (the one that remains on), or of the traffic reduction parameter (β) mean small energy savings because the large network 2 can switch off late in the evening and has to switch on again early in the morning.

5. Conclusions

Our study started with the observation that metropolitan areas are normally served by a few competing cellular network operators, which provide 24/7 full coverage, each dimensioning its network according to peak traffic, but providing redundant resources when traffic is low. It becomes thus possible to save energy by switching off some networks when their resources are not necessary, provided that operators are willing to cooperate, by accepting the competitor's subscribers as roaming customers while their home network is switched off, i.e., by temporarily becoming virtual operators.

We showed how to evaluate the feasibility and effectiveness of a switch-off scheme while preserving the desired QoS level. Our results, generated with different traffic patterns, showed that the amount of energy that is possible to save with this approach can be substantial (of the order of 20%, and above) and that many parameters influence the actual amount of energy that can be saved, including the roaming scheme and the order in which networks are switched off.

While in this paper we only considered the energy consumed by BSs, to feed the downlink channels, some comments on the uplink channel energy consumption are in order, because we must avoid shifting energy consumption from BSs to user terminals, specially because user terminals emit power very close to human bodies. Consider just two networks. Assuming that the density of BSs is the same in both networks, roaming from a network to another does not modify the average required uplink transmission power. The situation might be different, if we consider networks with different BS density, and it might favor some users with respect to others. Still, depending on the roaming direction, some users will gain, and some will lose, with a net average balance. This means that our proposed switch-off schemes do not imply additional energy consumption at the user terminals.

Our results suggest that the coexistence of operators in a competitive market is not energy efficient, unless a cooperative approach is used. Conversely, our results indicate that an energy-efficient approach to cellular networking is coherent with the present trend, which consists in several virtual operators sharing one physical infrastructure; the design of the Future Wireless Internet should include and encourage with proper incentives similar approaches, based on cooperations among operators.

References

- [1] ITU, Measuring the Information Society – The ICT Development Index, 2009.
- [2] EC, Progress Report on the Single European Electronic Communications Market 2008 (14th Report), March 24, 2009.
- [3] Solution-The green CDMA base station, <<http://www.huawei.com/publications/view.do?id=5715&cid=10549&pid=61>>.
- [4] S. Pileri, Energy and communication: engine of the human progress, <http://www.ega.it/intelec2007/img/Stefano_Pileri.pdf>, Intelec Opening Keynote Speech, 2007.
- [5] M. Etoh et al., Energy consumption issues on mobile network systems, in: IEEE SAINT 2008, Turku, Finland, July 2008.
- [6] M. Gupta, S. Singh, Greening of the Internet, in: ACM SIGCOMM 2003, Karlsruhe, Germany, August 2003.
- [7] <<http://www.e-energy.uni-passau.de/>>.
- [8] <<http://www.sigmetrics.org/sigmetrics2010/workshops.shtml>>.
- [9] <<http://conferences.sigcomm.org/sigcomm/2010/gncfp.php>>.
- [10] <<http://www.green-communications.net/globecom2010>>.
- [11] <<http://www.ict-earth.eu/>>.
- [12] M. Pickavet et al., Worldwide energy needs for ICT: the rise of power-aware networking, in: IEEE ANTS Conference, Bombay, India, December 2008.
- [13] M. Liebsch, X. Perez-Costa, Utilization of the IEEE 802.11 power save mode with IP paging, in: ICC 2005, Seoul, Korea, May 2005.

- [14] A. Jardosh, K. Papagiannaki, E. Belding, K. Almeroth, G. Iannaccone, B. Vinnakota, Green WLANs: on-demand WLAN infrastructure, Mobile Networks and Applications, 2009 (Special issue on Recent Advances in WLANs).
- [15] L. Chiaraviglio, D. Ciullo, M. Meo, M. Ajmone Marsan, Energy-aware UMTS access networks, in: W-GREEN 2008, Lapland, September 2008.
- [16] L. Chiaraviglio, D. Ciullo, M. Meo, M. Ajmone Marsan, Energy-efficient management of UMTS access networks, in: 21st International Teletraffic Congress (ITC 21), Paris, France, September 2009.
- [17] M. Ajmone Marsan, L. Chiaraviglio, D. Ciullo, M. Meo, Optimal energy savings in cellular access networks, in: GreenComm 2009 Workshop, Dresden, Germany, June 2009.
- [18] M. Ajmone Marsan, M. Meo, Energy efficient management of two cellular access networks, in: GreenMetrics 2009 Workshop, Seattle, WA, USA, June 2009.



Marco Ajmone Marsan is a Full Professor at the Electronics Department of Politecnico di Torino, in Italy, and a part-time Chief Researcher at IMDEA Networks in Madrid, Spain. He holds degrees in Electronic Engineering from the Politecnico di Torino and the University of California, Los Angeles. He was at Politecnico di Torino's Electronics Department from November 1975 to October 1987 - first as a researcher and then as an Associate Professor. He was a Full Professor at the University of Milan's Computer Science

Department from November 1987 to October 1990. From September 2002 to March 2009 he was the Director of the Institute for Electronics, Information and Telecommunications Engineering of the National Research Council. During the summers of 1980 and 1981, he was with the Research in Distributed Processing Group, Computer Science Department, UCLA. During the summer of 1998 he was an Erskine Fellow at the Computer Science Department of the University of Canterbury in New Zealand.

He has co-authored over 300 journal and conference papers in Communications and Computer Science, as well as the two books "Performance Models of Multiprocessor Systems," published by the MIT Press, and "Modelling with Generalized Stochastic Petri Nets," published by John Wiley.

In 1982, he received the best paper award at the Third International Conference on Distributed Computing Systems in Miami, Florida. In 2002, he was awarded a honorary doctoral degree in Telecommunications Networks from the Budapest University of Technology and Economics. He is the chair of the Italian Group of Telecommunications Professors, and the Italian Delegate in the ICT Committee of the Seventh Framework Programme of the EC.

Marco Ajmone Marsan is a Fellow of IEEE, and a corresponding member of the Academy of Sciences of Torino. He is a member of the steering committee of the IEEE/ACM Transactions on Networking. He participates in editorial boards of international journals, including the Computer Networks Journal by Elsevier. He is listed by ISI among the highly cited researchers in Computer Science.



Michela Meo received the Laurea degree in Electronics Engineering in 1993, and the Ph.D. degree in Electronic and Telecommunications Engineering in 1997, both from the Politecnico di Torino, Italy. Since November 1999, she is an Associate Professor at Politecnico di Torino. She co-authored more than 120 papers, about 40 of which are in international journals. She edited six special issues of international journals, including ACM Monet, Performance Evaluation, Journal and Computer Networks. She was program co-chair of

two editions of ACM MSWiM, general chair of another edition of ACM MSWiM, program co-chair of IEEE QoS-IP, IEEE MoVeNet 2007, IEEE ISCC 2009 and she was in the program committee of about 50 international conferences, including Sigmetrics, Infocom, ICC and Globecom. Her research interests are in the field of performance evaluation and modeling, traffic classification and characterization, P2P, green networking.