Two Teletraffic-based Schemes for Energy Saving in Cellular Networks with Micro-cells

Ziaul Haq Abbas and Frank Y. Li
Department of Information and Communication Technology, University of Agder (UiA), N-4898 Grimstad, Norway
Email: {ziaul.h.abbas, frank.li}@uia.no

Abstract— **The energy consumption of Base Stations (BSs) is known to constitute a major part of the power consumption in a cellular network. In this paper, we propose a novel approach which may switch a BS off under light traffic conditions in order to conserve the power consumption of such networks. More specifically, when the traffic load in the middle cell of a network with three micro-cells is sufficiently low, the corresponding BS can be switched off and its users will be covered by increasing the transmission power of one sector antenna in each of the two neighboring cells. Two teletraffic-based power saving schemes are proposed in our study. The first scheme analyzes the expected sojourn times of different channel occupancies and switches off the BS *deterministically* when the switching thresholds are met. The second scheme instead switches off the BS *probabilistically* based on a policy designed using a Finite Markov Decision Process (FMDP). Numerical results for the first scheme demonstrate that a reasonable amount of network power can be saved at the cost of slightly higher transmission power. The results for the second scheme indicate that a lower limit on the long-term network transmission power can be obtained using the FMDP-based analysis.**

Index Terms— **Micro-cell, teletraffic, power saving, energy, Markov chain, BS, FMDP, optimization.**

I. Introduction

The ubiquity of the Internet and mobile networks nowadays is causing rapid growth in the number of cellular access networks, especially at the micro- and pico-cell levels. Due to this rapid growth, the energy consumption of such networks is becoming an important issue from both the environmental and the economical points of view. Furthermore, because of the ever increasing data services such as social networking and the huge jumps in the number of mobile phone subscribers, a larger quantity of infrastructure equipment is also required.

On the other hand, the cellular network operators feel an impetus towards reducing the energy consumption of their networks in order to decrease the expense of operating a network. Due to competition in the market, the operators try to reduce their network operation cost while still keeping high Quality of Service (QoS) for the customers. From a system-level point of view, more than two-thirds of the energy in a mobile communication network are consumed by the part responsible for radio access, mainly Base Stations (BSs). Therefore, the operators' focus is shifting from optimizing the network deployment schemes towards developing techniques, such as intelligently switching unnecessary BSs off in order to further reduce the energy consumption of already optimally deployed cellular networks.

Moreover, as the BSs in a telecommunication network consume a major portion of the total network power [1] - [3], the power consumption reduction of BSs may be viewed as a definite direction towards green communication at the system level. Furthermore, in the technological domain, operators have started to incorporate new features in their infrastructure equipment allowing network elements to be remotely controlled, and even switched off under certain circumstances [4] and [5]. Taking such advancements as motivation, we propose in this paper a network energy saving scheme which switches off the middle BS in a linear configuration of a network with three micro-cells when the traffic intensity in the corresponding cell is below a threshold level. Correspondingly, the users of the switched-off BS are covered by increasing the transmission power of a sector antenna from each of the two neighboring cells when their respective traffic intensities are sufficiently low as well.

More specifically, we propose two distinct schemes with different approaches on when and how to switch off the BS. With the first scheme, we analyze the transmission and total power for the network based on the expected sojourn times for different channel occupancies and switch the BS off or on deterministically when the border states are reached. With the second scheme, we intend to minimize the transmission power consumption of the network by switching off the BS with certain probability using a Finite Markov Decision Process (FMDP) based state transition policy.

In the rest of the paper, Sec. II describes related work while Sec. III presents the network scenario along with a few assumptions. Secs. IV and V respectively analyze the two studied schemes. Sec. VI gives an account of BS power consumption. Numerical results for the two studied schemes are presented and discussed respectively in Secs. VII and VIII. Sec. IX gives a brief comparison of the two schemes. Finally,the paper is concluded in Sec. X.

II. Related Work

The interest of the research community in the field of network energy conservation has grown in the past few

Manuscript received October 21, 2011; revised April 15, 2012; accepted June 1, 2012.
Corresponding author: Ziaul Haq Abbas, Email: ziaul.h.abbas@uia.no

doi:10.4304/jcm.7.10.726-739

years. A lot of work has been carried out on the energy consumption reduction of individual network components. However, little attention has been paid on the idea of reducing the network energy consumption by switching part of the network components off while ensuring service availability and continuity. In the following paragraphs, we summarize the related work done by other researchers.

In [6], the authors show that dynamic adjustment of cell size according to the significant traffic variation in a day can help conserve network energy. However, their solution may create blind spots in the geographical coverage of the network. Moreover, the authors of [7] evaluate energy saving that can be achieved using energy-aware cooperative management of two different cellular access networks offering services over the same geographical area. According to their proposal, when traffic intensity is so low that only one of these networks is sufficient to provide desired QoS for the users of both networks, the other network is switched off. However, this approach requires complex potential hardware modification in the involved stations as well as inter-network cooperation algorithms and inter-operator agreement.

Furthermore, in [5], a few concepts to save energy in small-cell wireless communication BSs are presented and one of them suggested to switch off parts of the BS components if traffic load is low. However, no detailed analysis of any of the presented concepts is given in [5]. The possibility to decrease energy consumption of a cellular network is investigated in [8] by reducing the number of active cells during the low-traffic periods according to the time in a day. To do so, the authors carry out an analysis for a 24 hour period by considering low traffic intensity especially in the early morning, late evening, and night times. However, in reality, traffic intensity may become high even during night times, e.g., during festivals, emergency situations etc. How their scheme would deal with such situations is not addressed by these authors. Therefore, a solution which is able to switch off (and on) cells (and/or sectors) according to the variation in traffic intensity, regardless of the time in a day, is needed. This observation indeed partially triggered our work in this paper.

Furthermore, an analytical model to determine the effectiveness of the policies that activate Access Points (APs) in dense Wireless LANs according to user demands is proposed in [9]. However, the authors do not present the optimality of the performance of the proposed policies in terms of energy savings. In [10], the authors propose an approach aiming to realize green radio system by introducing the deployment of Green Base Stations (GBSs) within the already deployed cellular network. GBSs are basically powered by solar and wind energy sources. Coverage optimization for energy saving is presented in their work by optimally increasing or decreasing the footprint of GBSs. However, this approach reduces energy consumption of the network of already existing BSs at the cost of infrastructure expenditure added by the deployment of GBSs. Furthermore, what is the optimal figure of the saved total network energy (considering both types of BSs) is not discussed in [10]. Moreover, the authors in [11] present an insight into optimizing the energy efficiency at the BS in a cell through power control, however, focusing only on the down-link data transfer.

III. NETWORK SCENARIO: POWER SAVING SCHEME AND ASSUMPTIONS

The network scenario studied in this paper is depicted in Fig. 1(a). As shown in the figure, there are three adja-

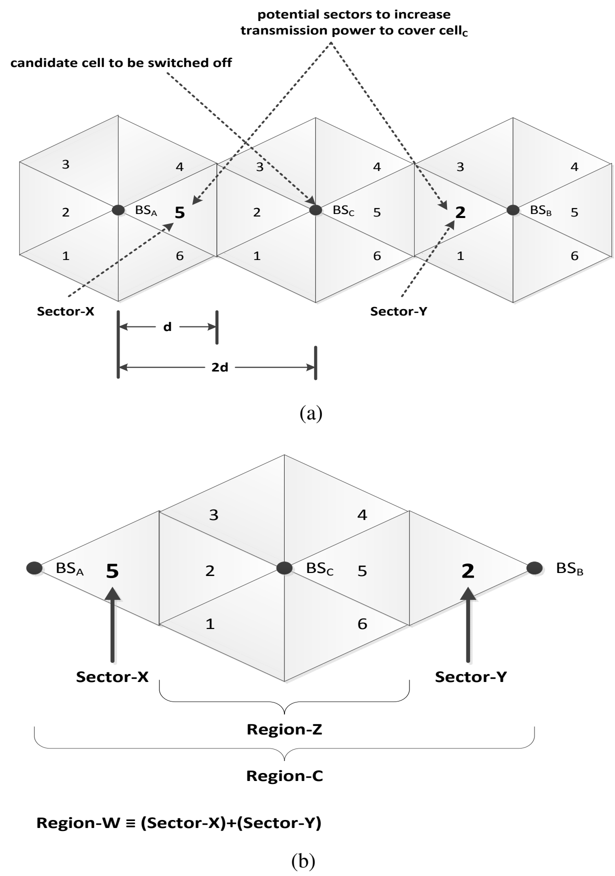


Figure 1. Network scenario (a) Overall network; (b) Definitions of regions.

cent identical-size hexagonal cells in the network, namely $cell_A$, $cell_C$, and $cell_B$. Every cell has a corresponding BS, i.e., BS_A, BS_C, and BS_B. In the remainder of the paper, the terms $cell_j$ and BS_j would be interchangeably used, where $j = A$, C, and B. Every BS has six sector antennas to cover its corresponding geographical region. For illustration purpose, Sector 5 of $cell_A$ and Sector 2 of $cell_B$ are respectively denoted as *Sector-X* and *Sector-Y*. All the six sectors of $cell_C$ are collectively denoted as *Region-Z*, as also depicted in Fig. 1(b). When a sufficient number of channels are available, Sectors X and Y are the potential sectors to increase the transmission power and cover the Mobile Stations (MSs) of $cell_C$ which will be switched off.

In our scheme, if there are enough channels available in Sectors X and Y and the total traffic intensity in Region-Z is sufficiently low then BS_A can increase the transmission power of the antenna for Sector-X to cover Sectors 1, 2, and 3 of Region-Z. Similarly, BS_B can cover Sectors 4, 5, and 6 of Region-Z by increasing the transmission power of its antenna for Sector-Y. Since all sectors of Region-Z are now supported by the neighboring two cells, $cell_C$ can be switched off. This mechanism can significantly reduce the total network power consumption while still providing coverage to all the users of the network (although only in low/moderate traffic situations). However, as Sectors X and Y have to increase their transmission power while the antennas of $cell_C$ are switched off (i.e., no transmission power consumption for $cell_C$), there would be a tradeoff between the total saved power and the extra consumed transmission power. How to minimize this transmission power consumption is one of the goals of this paper. It is worth mentioning that once a decision to increase the transmission power is taken, both Sectors X and Y increase their transmission power simultaneously and coordinately to fully cover Region-Z. Hence, from here on, the region collectively composed of Sectors X and Y would be referred to as *Region-W*, as depicted in Fig. 1(b). Furthermore, the region comprising collectively of Regions W and Z is referred to as *Region-C*.

Moreover, we define two modes of power consumption for the system. When Region-W increases its transmission power and $cell_C$ is switched off, the system is referred to as working in the Power Save ON (PS-ON) mode. Otherwise, it is referred to as working in the Power Save OFF (PS-OFF) mode.

For the sake of analysis simplicity, we assume that the traffic intensity in Sectors X and Y is the same and that the traffic around the central vertical line in Region-C is symmetric. It is also assumed that the transmission power increments of a sector antenna, and of its corresponding MSs under coverage, remain within the transmission regulation limits. That is, when a sector antenna increases its transmission power to cover the users at a distance of $2d$, the transmission power still remains under the regulation limits. Therefore, the proposed scheme is not suitable for cellular networks with macro-cells. A simplified path-loss model is utilized between the MSs and the BSs. Handshake procedures for cooperation among the BSs are ignored. For the power increasing of the antennas for Region-W, and the corresponding BS switching off, a controller is assumed to be operating on the backend. Furthermore, the switching latency that the BSs need for transmission status change is also ignored.

For calculation convenience, only homogeneous traffic with identical bit rate service is considered. Furthermore, for MSs and BSs, both uplink and downlink power control schemes are assumed to be operating [13]. When the system operates in the PS-ON mode, the interference effects around the vertical line dividing Region-Z into two equal halves (i.e., the line separating Sectors 1, 2, and 3 from Sectors 4, 5, and 6) are ignored. Moreover, the traffic around the mentioned vertical line in Region-C is also assumed symmetric.

IV. Deterministic Scheme Based on Sojourn Time Analysis

This section presents the power consumption analysis of our first scheme, i.e., sojourn time-based deterministic scheme.

Each of the regions W, Z, and C are individually modeled using separate Markov Chains (MCs). Each state in these MCs represents the collective number of occupied channels in the corresponding region. Thereafter, we obtain the blocking probabilities for each case from the MCs.

Based on our analysis, a power saving policy is developed with a few conditions being imposed on the individual thresholds of the number of channels occupied in each sector/region in order to avoid frequent mode switching between the PS-ON mode and the PS-OFF mode.

A. Markov Chains for the Regions

Each of the Regions W, Z, and C is modeled as a Birth-Death (BD) process, with Poisson arrival rate and exponential service times, using MCs. As an example, Fig. 2 shows the process for Region-W which has $2r$ channels in total. In this figure, the arrival rate is denoted by λ_W and the departure rate by μ_W. The initial state of the process is 0 and the final state is $2r$, where r is also the maximum number of channels available in a sector.

For the BD process of Region-Z (whose MC is not shown as a figure due to its minor pictorial variation from Fig. 2), the arrival and departure rates are represented as λ_Z and μ_Z, respectively. Since six sectors are considered collectively as Region-Z, there are a total number of $6r+1$ possible states for the MC of Region-Z. The lower and upper threshold values of channel occupancy in Regions W and Z, denoted respectively by t_l, t_u and t'_l, t'_u, would be discussed in Subsection IV-B.

However, as Region-C which jointly represents the area covered by Regions W and Z is obtained by increasing the transmission power of Region-W, there are still only $2r$ channels available in Region-C. The MC for Region-C is illustrated in Fig. 3.

Each state of the MC for Region-C represents the sum of the channels occupied in Region-W and Region-Z. Hence the arrival and departure rates are respectively denoted by $\lambda_C \equiv \lambda_W + \lambda_Z$ (because independent arrivals are considered in Regions W and Z) and $\mu_C \equiv \mu_W$ (because, like Region-W, only $2r$ channels are available to Region-C, as it is only a geographical (and not channel wise) extension of Region-W). The threshold values in Region-W and Region-Z are used by Region-C to decide whether the system operates in the PS-ON or the PS-OFF mode.

The blocking probabilities for each sector can be obtained according to the Erlang-B formula. For example, for Region-W

$$p_{2r} = \frac{(\frac{\lambda_W}{\mu_W})^{2r}/(2r)!}{\sum_{j=0}^{2r}(\frac{\lambda_W}{\mu_W})^j/j!} \quad , \qquad j = 0, 1, 2, \ldots, 2r. \qquad (1)$$

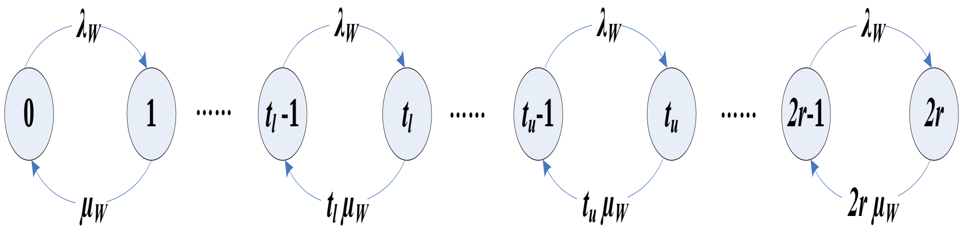


Figure 2. BD process for Region-W.

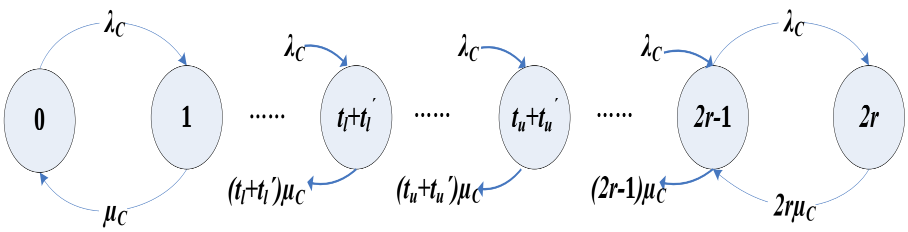


Figure 3. BD process for Region-C.

Here, p_{2r} is the probability that all $2r$ channels are occupied.

The blocking probability can be taken as a Grade of Service (GoS) measure in a system. In a loss system, GoS is used to indicate the proportion of calls that are lost due to congestion. Thus for a particular required GoS, the traffic intensity offered by each user (i.e., λ/μ) and the number of channels in use, we can find the total offered traffic intensity through the Erlang table. This leads to the number of users that can be supported in a region [14]. For decision making about mode switching, the blocking probability can then be utilized as a criterion, i.e., the threshold values of channel occupancy can be adjusted according to the blocking probability requirements.

B. Power Saving Policy

Our power-saving policy is primarily based on the threshold channel occupancy values of Region-W and Region-Z, considered collectively in the MC for Region-C.

Figure 4 illustrates the policy. As shown in the figure, when the MC is in any of the states from 0 to t_l+t_l', the network operates in the PS-ON mode. Similarly, when the MC is in any of the states from t_u+t_u' to $2r$, the network operates in the PS-OFF mode. However, when the system reaches state t_l+t_l' or state t_u+t_u', mode switching may happen.

To prevent the system from frequent (and undesirable) switching between the PS-ON and the PS-OFF modes, there is a hysteresis region defined between states t_l+t_l' and t_u+t_u'. This means that if the system reaches state t_u+t_u' from any of the states between $t_u+t_u'+1$ and $2r$ (i.e., the number of occupied channels is decreasing and the system is already in the PS-OFF mode), it does not simply trigger the PS-OFF mode until it reaches state t_l+t_l'. On the other hand, if the system is in the PS-ON mode and the number of occupied channels is increasing, it will remain in the same mode till it reaches the other threshold t_u+t_u'. Thus, states t_l+t_l' and t_u+t_u' are triggering states. There is no decision-making within the hysteresis region, i.e., between t_l+t_l' and t_u+t_u'.

There are a few necessary conditions/requirements for our model to be effective, as given below:

1) thresholds t_l and t_l' are respectively less than t_u and t_u';
2) $t_u + t_u' \leq c_1 \times 2r$ and $t_l + t_l' \geq c_2 \times 2r$. Here, c_1 and c_2 are scalers chosen according to the requirement/priority given to the PS-OFF and the PS-ON modes, respectively, and $0 < c_2 < c_1 < 1$. This guarantees that the upper threshold in Region-C is lower than $2r$ and the lower threshold is greater than 0, and there is a non-zero hysteresis region;
3) for the MCs of all sectors to be stable, $\lambda_i < \mu_i$, $i \in \{W, Z, C\}$.

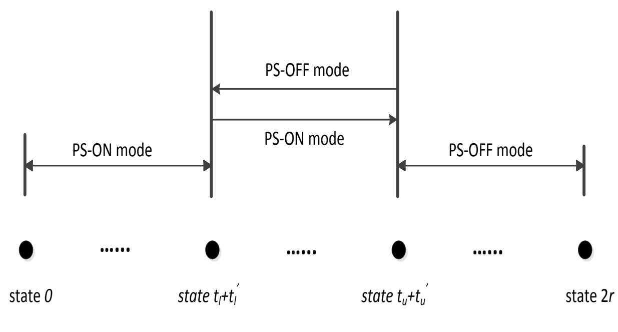


Figure 4. Power saving policy for the network.

C. The Matrix S

To evaluate the possible benefit of energy saving, we need to calculate the expected amount of time the system stays in the PS-ON and the PS-OFF modes. For this purpose, from the MC of Region-C, the sum of the expected times that the process remains within lower threshold is calculated. Thus, we calculate the total expected time of the process in state $t_l + t'_l$ given that it starts from any of the states $0, 1, \ldots, t_l + t'_l$ (i.e., a pure PS-ON mode). Furthermore, according to the above described power saving policy, we also need to obtain the total expected time that the process remains in the PS-ON mode (the middle part of Fig. 4) till it reaches state $t_u + t'_u$, i.e., the triggering state for the PS-OFF mode. Summing these two total times (i.e., the time in the pure PS-ON mode and the time in the hysteresis region while moving towards the upper threshold) gives the total time for the PS-ON mode, T_{PS-ON}.

Similarly, the total expected time the system spends in the PS-OFF mode, T_{PS-OFF}, can be calculated. The expected time within the hysteresis region is shared by both the PS-ON and the PS-OFF modes.

The above mentioned expected times can be obtained from a matrix **S** whose entries represent the expected sojourn times in different transient states, as described in the following.

From the BD process of Region-C, the rate transition matrix **Q** can be obtained. Using **Q**, the probability transition matrix **F** is computed[1] through the uniformization method [15] , i.e., each element q_{ij} of **Q** is transformed to f_{ij} of **F** as

$$f_{ij} = \begin{cases} q_{ij}/\theta \quad if \quad i \neq j, \\ 1 + q_{ii}/\theta \quad if \quad i = j, \end{cases} \tag{2}$$

with θ chosen such that $\theta \geq \max |q_{ii}|$. Please note that q_{ii} are the entries on the main diagonal of **Q** for the BD process and thus are negative numbers.

Let s_{ij} denote the expected number of time periods that the MC (for Region-C) is in state j given that it starts in state i. Let **S** denote the matrix of values s_{ij}, $i, j = 0, 1, \ldots, 2r$, i.e., **S** =

$$\begin{vmatrix} s_{0,0} & \cdots & s_{o,t_l+t'_l} & \cdots & s_{o,t_u+t'_u} & \cdots & s_{0,r} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ s_{t_l+t'_l,0} & \cdots & s_{t_l+t'_l,t_l+t'_l} & \cdots & s_{t_l+t'_l,t_u+t'_u} & \cdots & s_{t_l+t'_l,r} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ s_{t_u+t'_u,0} & \cdots & s_{t_u+t'_u,t_l+t'_l} & \cdots & s_{t_u+t'_u,t_u+t'_u} & \cdots & s_{t_u+t'_u,r} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ s_{r,0} & \cdots & s_{r,t_l+t'_l} & \cdots & s_{r,t_u+t'_u} & \cdots & s_{r,r} \end{vmatrix}$$

The matrix **S** is easily obtained from **F** as [16]

$$\mathbf{S} = (\mathbf{I} - \mathbf{F})^{-1}, \tag{3}$$

where **I** is an identity matrix.

Now, using **S**, the mean times spent in the PS-ON and the PS-OFF modes (as mentioned above) can be calculated by summing the relevant elements of **S**.

Note that in our numerical results, to be presented later in Secs. VII and VIII, T_{PS-ON} and T_{PS-OFF} are normalized with the total time, T, of **S**, and thus are *unit-less numbers*. Moreover, this implies that the energy consumption of these modes is indeed reflected even though the figures are labeled as power.

By using T_{PS-ON} and Eqs. (12) and (13) of Sec. VI, the total transmission and network power consumption in the PS-ON mode can be found. Similarly, T_{PS-OFF} can be used to calculate the power consumption values in the PS-OFF mode.

V. Probabilistic Scheme Based on FMDP Analysis

In the sojourn time-based scheme presented above, once the system reaches a triggering state, the mode is deterministically changed. In this section, we present the transmission power consumption analysis for our second scheme, i.e., FMDP-based scheme, in which the switching decision is made probabilistically instead. That is, the decisions to switch to the PS-ON or the PS-OFF mode are taken with certain probability (to be explained in the following).

Initially, we briefly present the necessary traffic analysis utilized for this scheme, followed by an example of how such analysis can be helpful in the policy development for FMDP. A state-transition matrix **P** is introduced, followed by the FMDP and the policy for minimizing the network transmission power consumption. Later on, a cost function, which is the long-term expected transmission power consumption, is defined and minimized using linear programming.

A. Traffic Analysis

As Regions W and Z are separately modeled as BD processes (described in the previous section), we can obtain the corresponding steady-state probabilities of channel occupancy. Let p_n and p_m be the probabilities that n and m channels are respectively occupied in Regions W and Z. Using Erlang's first formula [17], we have

$$p_n = \frac{A_W^n/n!}{\sum_{k=0}^{2r} A_W^k/k!} \quad , \qquad n = 0, 1, 2, \ldots, 2r, \tag{4}$$

$$p_m = \frac{A_Z^m/m!}{\sum_{k=0}^{6r} A_Z^k/k!} \quad , \qquad m = 0, 1, 2, \ldots, 6r, \tag{5}$$

where $A_W = \lambda_W/\mu_W$ and $A_Z = \lambda_Z/\mu_Z$. Here, λ_W/μ_W and λ_Z/μ_Z are the respective average traffic injected by each user in the corresponding region.

As the probabilities in Eqs. (4) and (5) are obtained from two independent regions (i.e., W and Z), the total

[1]In **F**, a few of the row sums are less than 1 while others are 1. This is because that the elements of **F** are the probabilities from one *transient* state to another *transient* state.

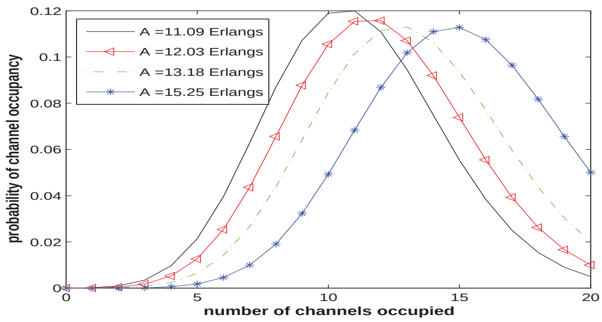


Figure 5. Effects of traffic load, A, on the probability of channel occupancy.

probability, p_R, that n channels are occupied in Region-W and m channels are occupied in Region-Z becomes $p_R = p_n \cdot p_m$.

The corresponding blocking probabilities for the regions can be obtained according to the Erlang-B formula [17], as already described in Sec. IV-A.

For a given region, Fig. 5 illustrates the behavior of the probability of channel occupancy for different values of A. As illustrated in the figure, the probability of higher number of channel occupancy is a little skewed towards right for increasing levels of A. This gives us an indication that for a given number of total available channels, with higher A, there is a higher tendency of channel occupancy in the upper half of the total channel range. Thus, there is a relatively higher chance of the system moving to a higher channel occupancy level from the lower levels. Furthermore, the channel occupancy probability drops, however less sharply, with increasing value of A near the maximum channel occupancy level. This means, e.g., for $A = 15.25$ Erlangs in the figure, the highest probability of channel occupancy is for 15 channels, and not for 10 channels (which is the middle value of channel occupancy) or for 20 channels (which is the maximum value of channel occupancy). Thus, once the transmission power of Region-W is increased (based on the channel occupancy information) to cover $cell_C$ (an action in the action space; to be explained in the next section), the probability of reducing the transmission power (another action in the action space) should be relatively low. Such information can be useful for the policy development, e.g., for assigning the probabilities of the system going from one channel occupancy interval to another in the state-transition matrix **P**.

B. State-transition Matrix *P*

In such type of optimization problems, the state-transition matrix **P** may be obtained after gathering historical data based on the inspection results [18]. In this paper, we obtain **P** by sampling the arrival and departure statistics evolved through time for a given network. Thus, we can assign reasonably appropriate values to the elements of **P** by observing the trend of channel occupancy probabilities for different number of occupied channels.

Each state of the MC for **P** represents a bounded range for the total number of channels occupied collectively in Regions W and Z. We define two low and two high values of threshold channel occupancy in these regions respectively. Let t_{Wl_1} and t_{Wl_2} denote these lower threshold values and t_{Wh_1} and t_{Wh_2} denote the higher threshold values in Region-W, and similarly t_{Zl_1}, t_{Zl_2}, t_{Zh_1}, and t_{Zh_2} for Region-Z. We can granulate[2] the states ($\mathbf{s} = \{s_0, s_1, ..., s_5\}$) used in **P** as follows, where *occupied channels* means the total number of channels occupied collectively in Regions W and Z:

- s_0: 0 occupied channels,
- s_1: $0 <$ occupied channels $\leq t_{Wl_1} + t_{Zl_1}$,
- s_2: $t_{Wl_1} + t_{Zl_1} <$ occupied channels $\leq t_{Wl_2} + t_{Zl_2}$,
- s_3: $t_{Wl_2} + t_{Zl_2} <$ occupied channels $\leq t_{Wh_1} + t_{Zh_1}$,
- s_4: $t_{Wh_1} + t_{Zh_1} <$ occupied channels $\leq t_{Wh_2} + t_{Zh_2}$,
- s_5: $t_{Wh_2} + t_{Zh_2} <$ occupied channels $\leq 2r$.

Each element of **P** represents a probability of transition between two states. For example, an element $p_{s_0 s_1}$ of **P** means a probability of transition from s_0 to s_1, i.e., the probability that fewer than or equal to $t_{Wl_1} + t_{Zl_1}$ number of channels are occupied in the next step given that currently no channels are occupied.

Now, for illustration clarity, we can write **P** in a two-dimensional elemental form as p_{ij}, $i, j = 0, 1, ..., 5$. Here, i and j represent the row and column indices and respectively indicate the present and next state. Moreover, $\sum_j p_{ij} = 1$. The values to the elements of **P** can be assigned using the information obtained from Eqs. (4) and (5), and Fig. 5.

In order for our model to be effective, there are a few necessary conditions to be met by the above mentioned thresholds, described as follows:

1) $0 < t_{Wl_1} + t_{Zl_1} < t_{Wl_2} + t_{Zl_2} \leq r$. This guarantees that the lower thresholds remain below 50% of the total allowed channel occupancy in Region-W, ensuring a fairer distribution of channels among the above mentioned six possible states of the MC.
2) $r < t_{Wh_1} + t_{Zh_1} < t_{Wh_2} + t_{Zh_2} \leq 2r$. This guarantees that the higher thresholds in Regions W and Z remain within the supportable channel limits of Region-W. Hence, Region-W is given priority over Region-Z because it is Region-W that needs to support the users of Region-Z if BS_C is switched off.

Furthermore, the steady-state values of the elements of **P**, i.e., π_i, $i = 0, 1, ..., 5$, can be easily obtained from **P** [16]. These are the limiting values of the probability of the system being in each of the states, irrespective of the initial state. For example, if p_{12} represents the transition probability to state s_2 from s_1 then π_2 is the long-term probability of the system being in state s_2. The values of π_i are used to obtain the steady-state values of the unconditional probability y_{ik} (see Eqs. (7) and (8)) from the conditional probability D_{ik} (see Eq. (6)) of action i

[2]Here, we have used six states, i.e., $s_0, ..., s_5$. Finer granularity would lead to more accuracy but at a cost of more computational complexity.

in state k. Hence, y_{ik} are our decision variables to be optimized in the cost function for the policy (explained in Subsection V-D).

Having developed **P** and associated necessary conditions, we now move on to develop the policy to be incorporated into the FMDP. Later, we will furnish a linear program for the optimization purpose.

C. FMDP and the Policy

An FMDP is a reinforcement learning technique that satisfies the Markov property [18]. It is defined by its state and action spaces[3] and by one-step dynamic of the environment. The state space is composed of all the states that the process can assume (i.e., for example, $s_0, s_1, ..., s_5$) while the action space consists of all the possible actions that can be taken from these states. Given any state and action, there is a probability of transition to another state. Thus, a state-transition in our FMDP means that the process moves with certain probability to another state based on the current state-action pair. Such probabilities are conditional and are generally named as transition probabilities. Furthermore, given any current state and action, together with any next state, there is an expected value of the associated cost (due to the action taken). This cost, e.g., in our problem, is the transmission power consumption (in Watts) when another action is taken based on the current state and action. This action may be to switch to the PS-ON or the PS-OFF mode, thereby, increasing or decreasing the transmission power consumption. As the actions are taken with certain probabilities, our aim in this scheme is to minimize this *expected* cost. Therefore, we develop a policy for this purpose.

The policy is meant to determine a probability distribution for the actions k $(k = 1, ..., K)$, to be taken when the system is in state i $(i = 0, 1, ..., M)$. Thus, there are a total number of K possible actions and $M + 1$ possible states. Hence, we define the conditional probability distribution as

$$D_{ik} = Prob\{action = k \mid state = i\}, \tag{6}$$

where $\sum_k D_{ik} = 1$ and $0 \leq D_{ik} \leq 1$.

The action space[4] is defined as follows:

- Action 1: Go to the PS-ON mode, i.e., shut down BS_C and increase transmission power of Region-W,
- Action 2: Go to the PS-OFF mode, i.e., switch on BS_C and decrease transmission power of Region-W,
- Action 3: Keep the current power mode unchanged.

It is worth mentioning that each of these actions can be taken with certain probability based on the current state and previous action. An action may or may not lead the system to a new state. If the system goes to a new state, the process may or may not go to another state based on the new action-state pair, and so on.

[3]From here on, the terms *action* and *decision* would be interchangeably used in this paper.

[4]For the sake of simplicity, the action space has been deliberately kept small. Larger action space leads to more complex policy, in general.

Hence, our policy can be characterized by a policy matrix, **D**, where each state of **D** is one of the states defined for **P** and each action is one of the actions from the above mentioned action space. Therefore, each element D_{ik} of **D** represents the probability of action k in state i; and each row of **D** is a Probability Distribution Function (PDF) of all the possible actions that can be taken in a state. For example $i = 0$ means state s_0 and $k = 1$ means Action 1; and so on.

A simple example of a policy matrix, **D**, is shown below, given that there are 6 states and 3 actions.

$$\mathbf{D} = \begin{array}{c} \\ s_0 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{array} \begin{array}{c} \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{pmatrix} p_h & 0 & 1-p_h \\ p_h & 0 & 1-p_h \\ p_m & 0 & 1-p_m \\ p_l & p_l & 1-2p_l \\ 0 & 1-p_h & p_h \\ 0 & p_h & 1-p_h \end{pmatrix} \end{array},$$

where the subscripts, h, m, and l of p, respectively indicate the high, medium, and low probability of an action in a state. Actions 1, 2, and 3 are respectively represented by a_1, a_2, and a_3. As an example, the top-left element of **D** means that given no channels are occupied, there is a high probability that the system goes to the PS-ON mode.

D. Cost Optimization and Linear Programming

In order to optimize the transmission power consumption cost through a linear program, we need first to define the decision variables, y_{ik}, to be optimized. Let y_{ik} be the steady-state probability that the system is in state i and action k is taken, i.e.,

$$y_{ik} = Prob\{action = k \text{ and } state = i\}. \tag{7}$$

From the rules of conditional probability, y_{ik} and D_{ik} can be related[5] as

$$y_{ik} = \pi_i D_{ik}. \tag{8}$$

The long-term expected transmission power consumption cost is given as

$$E(C) = \sum_{i=0}^{M} \sum_{k=1}^{K} \pi_i C_{ik} D_{ik} = \sum_{i=0}^{M} \sum_{k=1}^{K} C_{ik} y_{ik}, \tag{9}$$

where C_{ik} denotes the cost (in Watts) incurred when action k is taken in state i, and it can be calculated using Eq. (12).

As mentioned earlier, our goal is to minimize the cost function, $X \equiv E(C)$. The minimization of X means finding optimal values of the decision variables y_{ik} which minimize the expected cost (see Eqs. (7) and (9)). In order

[5]Note the difference between Eqs. (6) and (7). The former is a conditional probability while the latter is a limiting unconditional probability.

to do so, the model is turned into a linear program as follows:

$$\min \quad X = \sum_{i=0}^{M}\sum_{k=1}^{K} C_{ik} y_{ik} \tag{10}$$

$$\text{s.t.} \quad \sum_{i=0}^{M}\sum_{k=1}^{K} y_{ik} = 1,$$

$$\sum_{k=1}^{K} y_{jk} - \sum_{i=0}^{M}\sum_{k=1}^{K} y_{ik} p_{ij}(k) = 0,\ j = 0,1,...,M,$$

$$y_{ik} \geq 0 \ for\ i = 0,1,...,M \ and\ k = 1,2,...,K.$$

In this linear program, argument k in $p_{ij}(k)$ indicates that the appropriate transition probability depends on action k. The first constraint is necessary because as y_{ik} are probabilities of all action-state pairs, their sum must be 1. Hence, this is the normalization equation for the Markov process. The second constraint indicates that the steady-state probability of being in state j is the same as the probability computed by conditioning on the state and action chosen one stage earlier (as per definition of an MC). Hence, this is the balance equation for the Markov process. The third constraint imposes a lower bound on the decision variable values (as they represent probabilities, they should be non-negative).

VI. Base Station Power Consumption

The power consumption of a BS mainly consists of two major components, i.e., a component responsible for the total transmission power of the emitted signal and a component for the fixed power consumption. Below, we first describe how to obtain the total transmission power for a BS. The fixed power consumption component is later accounted for.

Quality of service can be taken as the acceptable cumulative effect on user satisfaction of all imperfections affecting the service. To ensure a certain QoS level for any specific type of service, a BS is required to maintain a minimum level of Signal to Interference and Noise Ratio (SINR) for a target MS. Hence, for a given noise level, there should be a minimum level of received power, P_{rx}, at the MS. Therefore, in order to ensure this minimum P_{rx}, the BS needs to transmit at least with a power level of P_{tx}. Thus we can write

$$P_{tx} = P_{rx} \cdot d^{\alpha} \cdot L_1. \tag{11}$$

Here $P_{rx} = \gamma \cdot (W \cdot T_0 + I)$, where γ is the required minimum SNR level for the target MS, W is the channel bandwidth, T_o is the thermal noise level, and I is the maximum assumed value of the interference level. Furthermore in Eq. (11), the distance between the MS and the BS is denoted by d, and α is the path-loss coefficient. The losses due to fading and building penetration etc are collectively given by component L_1. The total transmission power consumption by the BS towards a total number of N_A active users in a cell/sector can thus be written as

$$P_{Tot\_tx} = N_A \cdot P_{tx}. \tag{12}$$

Eq. (12) is calculated for the worst case power consumption scenario, i.e., the MSs are considered to be located at the cell/sector boundaries.

The fixed power consumption component, P_{fix}, of a BS includes power consumed due to heating effects, electronic processing of transmitted and received signals, etc. For the sake of simplicity, it is assumed as constant.

Therefore, the total BS power consumption can now be written as

$$P_{BS} = L_2 \cdot N_S \cdot P_{Tot\_tx} + P_{fix}, \tag{13}$$

where scalar N_S represents the number of sectors in a cell. The losses associated with components of a BS, e.g., antenna feeder cable loss, directional antenna gain, power amplifier efficiency etc are collectively denoted as L_2 in Eq. (13).

In the deterministic scheme case, the number of users can be obtained based on the channel threshold values. Similarly, the number of users for the FMDP-based scheme can be obtained based on the state information the system is currently operating in. Then using the value of N_A (i.e., the number of active users), Eq. (12), and Eq. (13), we can obtain the total transmission power consumption of BS for a sector (and for a cell when scaled accordingly).

VII. Numerical Results: Sojourn Time-based Deterministic Scheme

In this and the next sections, we evaluate numerically the performance of the proposed schemes using MATLAB. The evaluation parameters are summarized in Table I [2], [19], and [20].

The total transmission power of a cell before and after the increase in transmission power of the sector antenna is calculated. Similarly the total consumed power of the network before and after the switch-off of BS_C calculated. Thus, the benefit as well as the tradeoff in the consumed power are obtained.

As mentioned earlier, the network power has two components, i.e., the power consumed for transmission of

TABLE I.
ANALYSIS PARAMETERS.

Parameter	Value
Channel bandwidth W	5 MHz
Minimum SNR value for a service γ	-18 dB
Total noise + interference density	-166 dBm/Hz
Receiver sensitivity	-117 dBm
Loss component (Rayleigh fading)	$2 \sim 5$ dB
Loss component (building penetration)	$12 \sim 15$ dB
Loss component (shadowing)	$6 \sim 7$ dB
Path-loss coefficient α	4
Distance between MS and BS d	$(450, 900)$ meters
BS directional antenna gain	10 dB
MS antenna gain	0 dB
P_{fix}	60 W
BS antenna feeder cable loss	-2 dB
Power amplifier efficiency	50%
Number of channels in each sector r	15
Average injected traffic by each user λ/μ	$(0.4, 0.5, 0.6)$ Erlangs
GoS level	$(0.5, 1, 2)$ %

signals, and the fixed BS power consumption component. In the following, we discuss the effects of our scheme on both of these components for different values of GoS, $t_l + t'_l$, $t_u + t'_u$, and λ/μ. In both Figs. 6 and 7, the upper part illustrates total transmission power and the lower part illustrates total network power consumption for the sojourn time-based scheme.

The bars in Set 3 of the upper parts of Figs. 6 and 7 indicate the extra cost in terms of transmission power in the PS-ON mode versus the PS-OFF mode. Similarly, the bars in Set 3 of the lower parts of all of these figures indicate the saving in the network power.

It is worth mentioning that all these results are for the transmission and total power consumption of the network. In other words, MS power consumption is not included in these figures. GoS requirement is given as a percentage value and indicates the blocking probability.

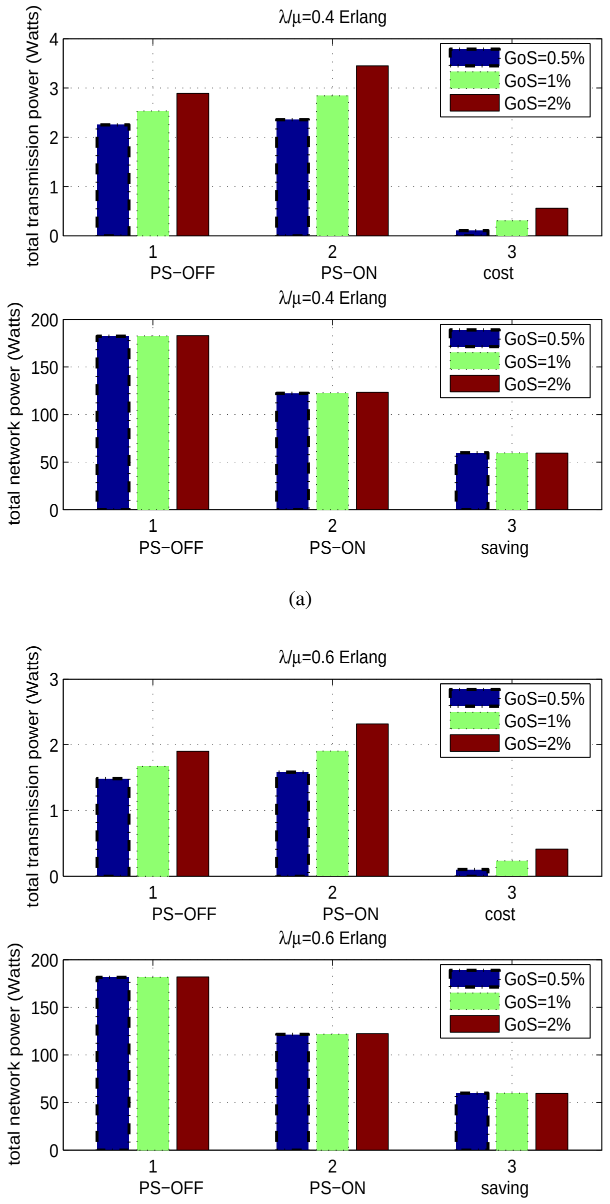
A. Network Transmission Power Consumption

Here, we present the effects of GoS and λ/μ; hysteresis region boundary shifting; and hysteresis region length variation on the network transmission power consumption. Moreover, the effects on the total network power consumption are also described.

1) Effects of GoS and λ/μ: As illustrated in Fig. 6, for λ/μ =0.4, loose requirement on GoS (e.g., 2%) causes more transmission power consumption in both the PS-ON and the PS-OFF modes. This is because that the network can provide service to a higher number of users if GoS requirement is relaxed. That is, fewer number of calls are blocked with higher GoS value. Furthermore, the PS-ON mode consumes more transmission power than the PS-OFF mode. This is because of the reason that in the PS-ON mode, two of the sector antennas (i.e., of BS_A and BS_B) need to cover $cell_C$ as well. Hence, they are required to transmit with higher power to reach the users at distance $2d$, causing further power consumption due to the exponential relation between transmission power and distance. This indicates the extra cost in terms of transmission power that the network operator has to bear in order to save network power.

Comparing Figs. 6(a and b) (upper parts only), similar trend for transmission power consumption is observed. However, the magnitude of transmission power consumed for any GoS requirement in Fig. 6(b) is lower than the corresponding values in Fig. 6(a). For example, in the PS-ON mode, for the same value of GoS (here, GoS = 2%), the total transmission power is about 2.4 Watts for $\lambda/\mu = 0.6$ in Fig. 6(b); whereas it is more than 3 Watts for $\lambda/\mu = 0.4$ in Fig. 6(a). This is because of the reason that as λ/μ is taken 0.6 in Fig. 6(b), the amount of traffic injected by each user into the network is higher. This leads to a lower number of supported users at a given time because a higher number of channels is occupied by each user on average. A similar trend can also be observed for the PS-OFF mode.

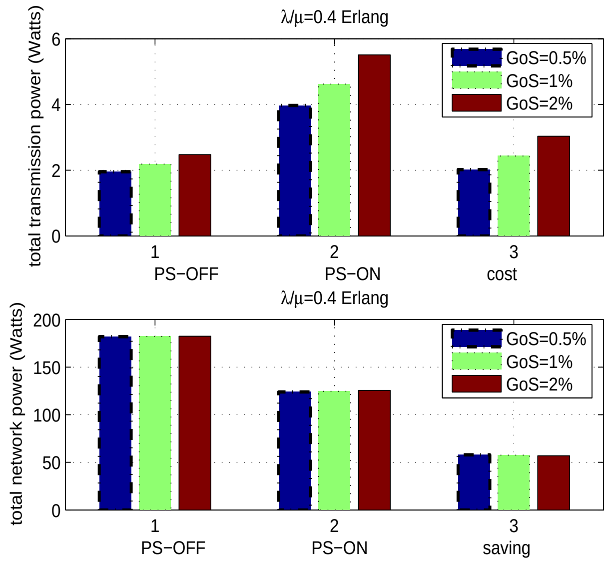
2) Effects of Hysteresis Region Boundary Shifting: By comparing Figs. 7(a and b), we observe a similar trend as discussed in the above paragraphs for Fig. 6. However, a higher amount of transmission power is consumed in the PS-ON mode for all mentioned GoS requirements than the corresponding values in Figs. 6, respectively. This is because that since, in Figs. 7(a and b), the thresholds $t_l + t'_l$ and $t_u + t'_u$ are increased by one, there is a higher number of states in the pure PS-ON region (as illustrated in Fig. 4 and Sec. IV-B). This implies a higher amount of total expected time spent in the PS-ON mode. Hence, the network remains in the PS-ON mode for a relatively longer fraction of time than in the PS-OFF mode. Furthermore, in the PS-ON mode, since two sector antennas (i.e., of BS_A and BS_B) transmit at longer distance (i.e., $2d$), there is higher transmission power



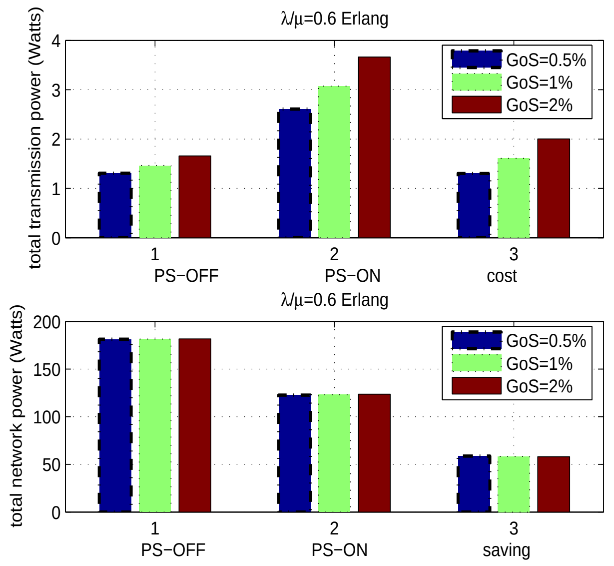
(a)

(b)

Figure 6. Deterministic scheme: Total transmission and network power consumptions for (a) $\lambda/\mu = 0.4$, $t_l + t'_l = 5$, $t_u + t'_u = 10$; (b) $\lambda/\mu = 0.6$, $t_l + t'_l = 5$, $t_u + t'_u = 10$.



(a)



(b)

Figure 7. Deterministic scheme: Total transmission and network power consumptions for (a) $\lambda/\mu = 0.4$, $t_l + t'_l = 6$, $t_u + t'_u = 11$; (b) $\lambda/\mu = 0.6$, $t_l + t'_l = 6$, $t_u + t'_u = 11$.

consumption even though the antennas of BS_C are not transmitting. Thus, for a given λ/μ and GoS requirement, *shifting the thresholds to the higher values results in a higher transmission power consumption in the PS-ON mode.*

Furthermore, comparing Figs. 7(a and b) with Figs. 6(a and b) respectively, for the PS-OFF mode, little difference in transmission power consumption for the same λ/μ is observed. The reason for this effect is because that, as in the PS-OFF mode all the antennas are transmitting at normal distance, shifting the hysteresis region to just one higher value (and thus reducing the pure PS-OFF region) only decreases the transmission power slightly.

B. Total Network Power Consumption

The effects of the proposed scheme on the total network power consumption are illustrated in the lower parts of Figs. 6 and 7. As can be observed in all the figures, *the proposed scheme saves a considerable amount of network power for all sets of GoS, λ/μ, $t_l + t'_l$, and $t_u + t'_u$*. In all cases, the amount of power saved is more than 50 Watts out of about 185 Watts. It is worth mentioning that the values of saved network power for different parameters are somewhat different. However, this difference is not clearly visible in the figures due to the scale on vertical axes.

The major reason for the large amount of network power saving is due to BS switch-off. Even when the transmission range of an antenna (i.e., for example, in Region-W) is increased from d to $2d$ while the remaining sector antennas of the corresponding BS are transmitting at distance d, the total transmission power value remains much lower than the fixed power consumption component of a BS, due to lower traffic intensity in Region-C. Hence, a BS's fixed power consumption always overwhelms the corresponding transmission power consumption for reasonable values of d. Furthermore, when a BS is switched off (i.e., BS_C in this case), its own transmission power is also saved. Thus, switching off just one BS in our network results in a large amount of network power saving. Nevertheless, this saving is achieved at a cost of transmitting extra power from two sector antennas (of BS_A and BS_B). Numerically with our scheme, *more than 50 Watts of network power can be saved by sacrificing about 4 - 5 Watts in terms of extra transmission power.* This implies a cost-to-benefit ratio of about 1:10.

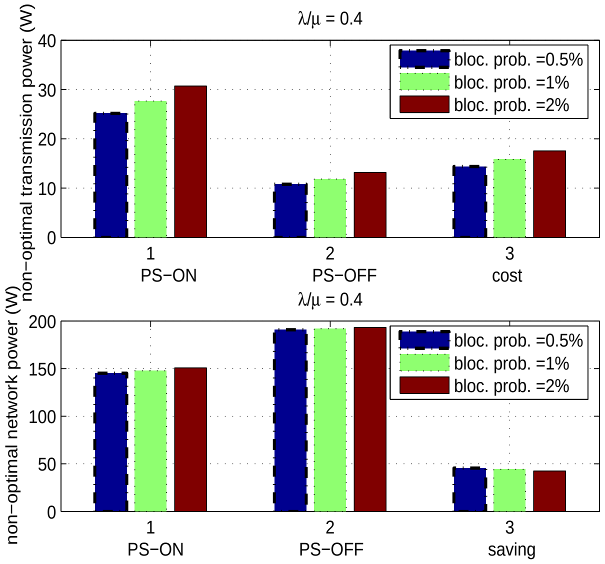
VIII. Numerical Results: FMDP-based Probabilistic Scheme

This section presents the results achieved by numerical evaluation of the FMDP-based probabilistic scheme. The linear program is solved by using the well-known simplex algorithm [21].

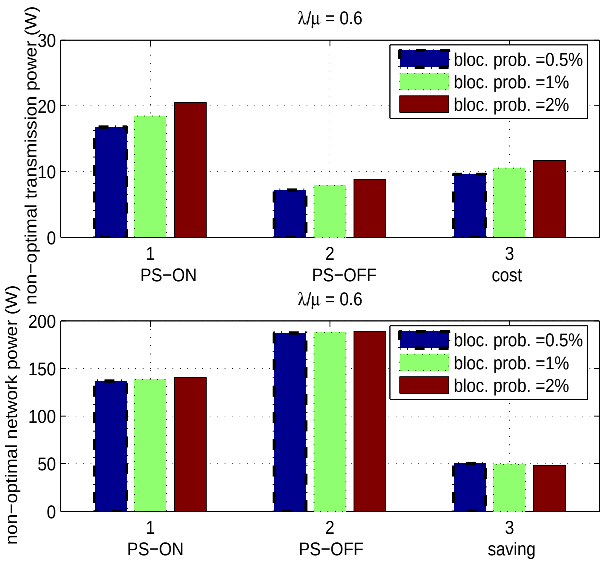
A. Power Consumption without Optimization

For a fair comparison with the obtained optimization results, we first need to know the non-optimal power consumption picture for our FMDP-based scheme. With non-optimal, we mean the normal power consumption without minimization algorithm applied on the system, neither in the PS-ON nor in the PS-OFF mode. The upper and lower parts of Fig. 8(a) present, respectively, the non-optimal network transmission and total power consumption in both the PS-ON and the PS-OFF modes for $\lambda/\mu = 0.4$ and different levels of blocking probability. Similarly, Fig. 8(b) illustrates the same set of results for $\lambda/\mu = 0.6$.

As illustrated in Fig. 8(a) (the upper part), the values of power consumption are lower in the PS-OFF mode than the corresponding values in the PS-ON mode. This result is expected and it is because that in the PS-ON mode,



(a)



(b)

Figure 8. Probabilistic scheme: Non-optimal network transmission and total power consumption for (a) $\lambda/\mu = 0.4$, and (b) $\lambda/\mu = 0.6$.

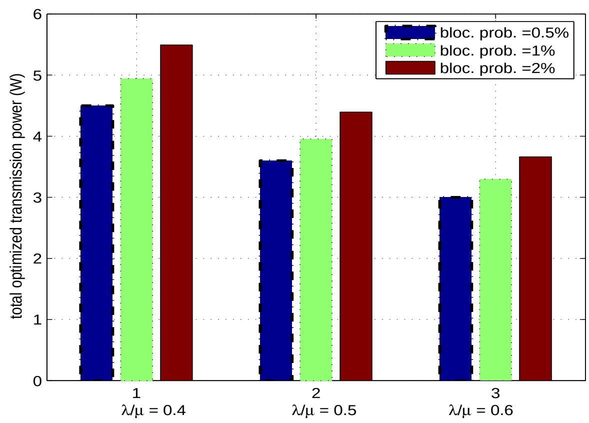


Figure 9. Optimum network transmission power consumption.

two of the sector antennas (i.e., of $cell_A$ and $cell_B$) are required to transmit with a higher power level to reach the users in $cell_C$ at distance $2d$. This causes further power consumption due to the exponential relation between the transmission power and the distance (as implied in Eq. (11)).

Moreover, by comparing the upper and lower parts in Fig. 8(a), we observe a clear advantage in terms of total network power consumption. That is, *the system can save a considerable amount of network power in the PS-ON mode at certain extra cost of the transmission power*. This is because that in the PS-ON mode, only two BSs are consuming power and the BS_C is switched off. This saves a reasonable amount of network power because P_{fix} for BS_C as well as its antenna transmission power are not contributing to the network power consumption. However, comparing Set 3 in the upper and lower parts of Fig. 8(a), we observe that the ratio of the network power saving to the extra transmission power cost decreases. This is because that when the transmission power is higher, there is a relatively lower amount of saved total network power.

Fig. 8(b) illustrates the total non-optimal transmission and network power consumption given $\lambda/\mu = 0.6$. The trend of power consumption is in accordance with that of Fig. 8(a). However, the values of the consumed network power (both transmission and total) are lower in Fig. 8(b) compared with those in Fig. 8(a). This effect is due to the reason that λ/μ is higher in Fig. 8(b), meaning that more traffic is injected by each user on average. Hence, a fewer number of users can be supported for a given number of occupied channels, resulting in a lower amount of transmission (and network) power consumption.

B. Optimum Power Consumption

The optimized[6] transmission power consumption of the system for different values of GoS level and λ/μ is illustrated in Fig. 9. The linear program (Eq. (10) and associated constraints) and Eq. (9) are respectively used to obtain the minimized y_{ik} and the minimized expected transmission power consumption cost. As expected, the optimized (i.e., minimized) transmission power values for lower values of λ/μ are higher for given blocking probability levels. This is because that for higher λ/μ, on average, more traffic is injected by individual users. Thus a higher number of channels is more quickly occupied, pushing the system towards the PS-OFF mode. Hence, a lower amount of transmission power is consumed because antennas only need to transmit over distance d.

A comparison of Fig. 9 with the upper parts of Fig. 8(a and b) illustrates that the optimized transmission power consumption is the lowest. This is because that *instead of switching to the PS-ON/OFF mode definitely, we switch to*

[6]As the goal of the optimization is to minimize only the transmission power consumption, the network power consumption results will not be discussed.

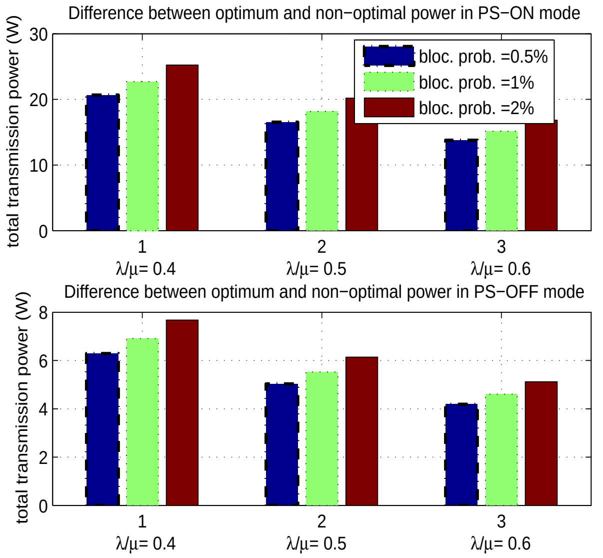


Figure 10. Difference between non-optimal and optimum transmission power consumption in the PS-ON mode (upper part) and the PS-OFF mode (lower part).

the PS-ON/OFF mode with certain probability depending on the state-action pair in the FMDP, so the power consumption is optimized. This ensures minimum long-term transmission power consumption due to the time intervals when very few channels are occupied.

In Fig. 10 (the upper part), the difference between the non-optimal transmission power consumption in the PS-ON mode and the optimum transmission power is illustrated. That is, e.g., Set 3 in Fig. 10 (upper part) illustrates the corresponding difference between Set 1 of Fig. 8 (upper part) and Set 3 of Fig. 9. The trend for this result is in accordance with the results described in the above paragraphs of this section. Similarly, Fig. 10 (lower part) illustrates the same for the PS-OFF mode. That is, e.g., Set 3 in Fig. 10 (lower part) illustrates the corresponding difference between Set 2 of Fig. 8 (upper part) and Set 3 of Fig. 9. From these results, we observe that by using the proposed optimization scheme, even in the PS-OFF mode, a further saving of about 5 Watts is achieved in terms of total transmission power, for the given levels of blocking probability and λ/μ.

1) Effects of Flipped ***P****:* Figs. 11 and 12 respectively present the same phenomena as those in Fig. 9 and 10, however, for a flipped version of **P** along its rows. Flipping **P** along its rows has the effect of reversing the channel occupancy probabilities from states s_0 to s_5. That is, the steady-state probability of s_5 in **P** becomes steady-state probability of s_0 in the flipped **P**, and so on. As expected, the trends are similar if we compare Fig. 9 with Fig. 11 or Fig. 10 with Fig. 12. However, as observed by comparing Fig. 11 with Fig. 9, the optimized transmission power consumption is lower in Fig. 11. The reason for this effect is because that as **P** is flipped, the transition probabilities between states are also changed. For example, if the transition probability between s_0 and s_4 was high in the first place, it becomes low in the flipped **P**. Therefore, arguing on these lines, in general in the flipped **P**, the probabilities that a higher number of channels are occupied are relatively lower than probabilities that a lower number of channels are occupied. Thus, a flipped version of **P** results in an even lower amount of optimized transmission power.

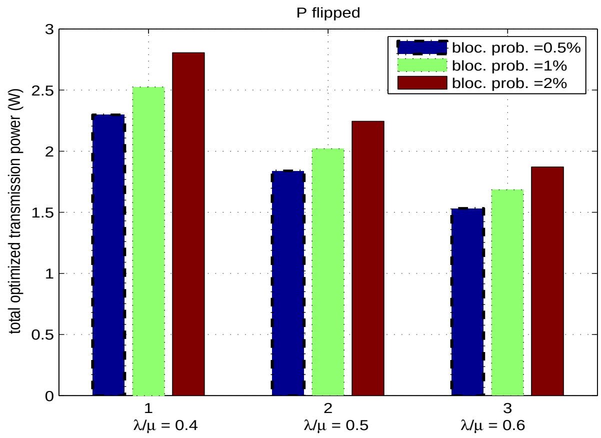


Figure 11. **P**-flipped: Optimum network transmission power consumption.

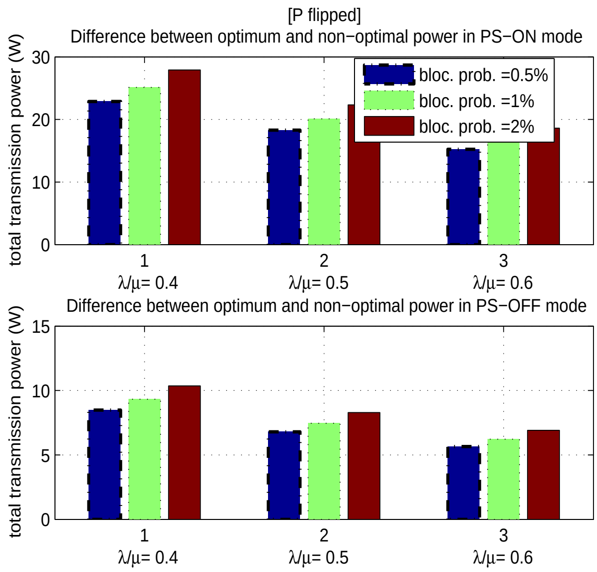


Figure 12. **P**-flipped: Difference between non-optimal and optimum transmission power consumption in the PS-ON mode (upper part) and the PS-OFF mode (lower part).

However, it is worth mentioning that this flipped version of **P** is discussed here merely for the purpose of comparison. Furthermore, this also indicates that the optimization policy may be improved further based on whether the system operator requires higher/lower probabilities for the states in **P**. How such policies can be formed and improved is, nevertheless, beyond the scope of this paper.

IX. Comparison and Future Discussions

A comparison between Figs. 6 and 8 indicates that, for transmission power consumption, the values in any mode for the FMDP-based scheme (for non-optimal power only) are higher than those for the sojourn time-based scheme, for given values of λ/μ. The reason for this effect is due to the fact that in the sojourn time-based scheme, the power consumption is associated with the values of the thresholds $t_l + t'_l$ and $t_u + t'_u$. Since the values of $t_l + t'_l$ are much lower than the maximum value of r, a few channels are occupied meaning that a few users are active. So the transmission power consumption value is also correspondingly low.

Intuitively, one may expect that the optimized transmission power would be the least when the two schemes are compared. Surprisingly, by comparing Fig. 9 with any of the Figs. 6 and 7, we observe that the optimized transmission power consumption is in fact relatively higher. The reason is, again, due to the lower values of the thresholds ($t_l+t'_l$ and $t_u+t'_u$) selected in our earlier work [12] which bring the power consumption in the PS-ON and the PS-OFF modes lower, compared with the optimum values of power in the FMDP-based scheme. However, the selection of higher values of $t_l+t'_l$ and $t_u+t'_u$ will generate higher transmission power consumption in the sojourn time-based scheme than the optimum consumption generated in the FMDP-based scheme.

Furthermore, as power consumption is sensitive to the selection of optimum threshold levels as well as the hysteresis region length, they need to be explored further for an optimum level of transmission power consumption. On the other hand, in the FMDA-based scheme, if the lower channel occupancy levels are assigned with higher probabilities (in $\mathbf{D}$), the corresponding total transmission power consumption will be lower due to the fact that the system spends more time in low channel occupancy states. Therefore, a joint optimization, targeted at threshold selection in the deterministic scheme as well as probability assignment to channel occupancy states in the probabilistic scheme, is required.

Moreover, as our proposed schemes are not restricted to any particular time intervals of a day, they are applicable to all times and all traffic intensities. That is, during the rush hours the schemes may more often go to the PS-OFF mode, and during the low-traffic hours they more often remain in the PS-ON mode. Thus they can save power according to the traffic intensities at all times than just low-traffic times of a day.

To summarize, we claim that both proposed schemes are useful for network operators. The sojourn time-based is preferable in reducing the network transmission power consumption. The FMDP-based scheme can be preferable by the operators in situations where estimating the long-term power consumption cost of their networks is a priority.

X. Conclusions and Future Work

In this paper, we have proposed two independent teletraffic-based schemes to analyze the transmission and network power consumption for a simple network with micro-cells. Numerical results demonstrate that in the sojourn time-based scheme, our hysteresis-based approach saves overall network power for an acceptable amount of extra transmission power consumption cost. In the FMDP-based scheme, the analysis can be utilized to minimize the long-term transmission power consumption in a cellular network.

As our future work, we will extend the current scenario to a more realistic network with multi-tier cells. As in such a scenario, more than one cells can enter into sleep mode, we will investigate the optimum number and locations of the switched-off cells and compensating sectors. Hence, interference from other cells, inter-cell interaction, co-frequency deployment etc will also be thoroughly investigated.

Another further extension of this work will be the inclusion of IP-traffic based analysis with asymmetric traffic distribution and non-uniform MS distribution.

References

[1] J.T. Louhi, “Energy Efficiency of Modern Cellular Base Stations,” *in Proc. International Telecommunications Energy Conference (INTELEC)*, Rome, Italy, October 2007.

[2] iCell COMPAC IP-RAN EV-DO NETWORK: Technical Specifications. Homepage: http://www.starsolutions.com/index.php/radio-access-network-components/bts-compact-micro-s

[3] *Node B datasheets*, http://www.motorola.com/, 2008.

[4] “Vodafone Using Ericssons New Power-saving Base Station Feature,” http://www.3g.co.uk/PR/Dec2007/5524.htm, (website accessed on Sept. 02, 2011).

[5] B. Debaillie, A. Giry, M. J. Gonzales, L. Dussopt, M. Li, D. Ferling, and V. Giannini, “Opportunities for Energy Savings in Pico/Femtocell Base-Stations, ” *in Proc. Future Network and Mobile Summit*, Warsaw, Poland, June 2011.

[6] S. Bhaumik, G. Narlikar, S. Chattopadhyay, and S. Kanugovi, “Breath to Stay Cool: Adjusting Cell Sizes to Reduce Energy Consumption,” *in Proc. First ACM SIGCOMM Workshop on Green Networking*, New Delhi, India, August 2010.

[7] M. A. Marsan and M. Meo, “Energy Efficient Management of Two Cellular Access Networks,” *in Proc. GreenMetrics 2009 Workshop*, Seattle, WA, USA, June 2009.

[8] M. A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, “Optimal Energy Savings in Cellular Access Networks,” *in Proc. GreenComm’09 - First International Workshop on Green Communications*, Dresden, Germany, June 2009.

[9] M. A. Marsan, L. Chiaraviglio, and D. Ciullo, “A Simple Analytical Model for the Energy-efficient Activation of Access Points in Dense WLANs,” *in Proc. 1st International Conference on Energy-efficient Computing and Networking, e-Energy 2010*, New York, USA, April 2010.

[10] J. Zhou, M. Li, L. Liu, X. Sge, and L. Chen, “Energy Source Aware Target Cell Selection and Coverage Optimization for Power Saving in Cellular Networks,” *in Proc. IEEE/ACM GreenCom-CPSCom*, Hangzhou, China, December 2010.

[11] Z. Chong and E. Jorswieck, "Analytical Foundation for Energy Efficiency Optimisation in Cellular Networks with Elastic Traffic," *in Proc. of 3rd International ICST Conference on Mobile Lightweight Wireless Systems (Mobilight)*, Bilbao, Spain, May 2010.

[12] Z. H. Abbas and F. Y. Li, "A Novel Teletraffic-based Power Saving Scheme for Cellular Networks with Microcells," *in Proc. 7th IEEE Performance and Management of Wireless and Mobile Networks P2MNET Workshop in conjunction with 36th IEEE LCN*, Bonn, Germany, October 2011.

[13] W. C. Y. Lee, "Overview of Cellular CDMA," *IEEE Trans. on Vehicular Technology*, Vol. 40, No. 2, May 1991.

[14] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd Edition, Prentice-Hall Inc, 2002.

[15] W. J. Stewart, *Probability, Markov Chains, Queues, and Simulation: the Mathematical Basis for Performance Modeling*, Princeton University Press, 2009.

[16] S. M. Ross, *Introduction to Probability Models*, 9th edition, Elsevier Inc., 2007.

[17] A. Popoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th Edition, McGraw-Hill Companies, Inc., New York, USA, 2002.

[18] F. S. Hillier and G. J. Lieberman, *Introduction to Operations Research*, 8th edition, McGraw-Hill Companies Inc., New York, USA, 2005.

[19] T. Ojanperä and R. Prasad, *Wideband CDMA for Third Generation Mobile Communications*, Artech House, 1998.

[20] B. T. Ahmed, "WCDMA Downlink Capacity of Cigar-shaped Microcells for Underground Metro Service," *European Trans. on Telecommunications,*, Vol. 20, No. 4, June 2009.

[21] I. Maros, *Computational Techniques of the Simplex Method*, Kluwer Academic Publishers, The Netherlands, 2003.

Ziaul Haq Abbas received his M.Phil. degree in Electronics from Quaid-e-Azam University, Pakistan, in 2001. Afterwards, he joined GIK Institute of Engineering Sciences and Technology, Pakistan, and worked as a Research Associate. His academic interests have been in the field of communications and DSP. He joined Agder Mobility Lab at the Department of Information and Communication Technology, University of Agder, Norway, as a Ph.D. candidate in August 2008. Currently, his research focus is on energy efficiency in hybrid Mobile and wireless networks. Abbas is a student member of the IEEE.

Frank Y. Li holds a Ph.D. degree from the Norwegian University of Science and Technology (NTNU). He worked as a senior researcher at UniK - University Graduate Center, University of Oslo before joining the Department of Information and Communication Technology, University of Agder as an Associate Professor in August 2007 and is currently a Professor in the same department. Dr. Li is a senior member of the IEEE. His research interest includes 3G and beyond mobile systems and wireless networks; mesh and ad hoc networks; cooperative communications; cognitive radio networks; green wireless communications; QoS, resource management and traffic engineering in wired and wireless IP-based networks; analysis, simulation and performance evaluation of communication protocols and networks. Dr. Li is a senior member of the IEEE.