Computer Networks 54 (2010) 2519–2536

Contents lists available at ScienceDirect

Computer Networks

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

Spectrum sharing in cognitive radio networks with imperfect sensing: A discrete-time Markov model

Xavier Gelabert *, Oriol Sallent, Jordi Pérez-Romero, Ramon Agustí

Department of Signal Theory and Communications (TSC), Universitat Politècnica de Catalunya (UPC), 08034 Barcelona, Spain

ARTICLE INFO

Article history:
Received 14 January 2010
Received in revised form 27 March 2010
Accepted 6 April 2010
Available online 11 April 2010
Responsible Editor: M.C. Vuran

Keywords:
Mobile communication systems
Cognitive radio networks
Markov processes
Queuing theory
Model validation and analysis

ABSTRACT

An efficient and utmost utilization of currently scarce and underutilized radio spectrum resources has stimulated the introduction of what has been coined Cognitive Radio (CR) access methodologies and implementations. While the long-established approach has been based on licensed (or primary) spectrum access, this new communication paradigm enables an opportunistic secondary access to shared spectrum resources provided mutual interference is kept below acceptable levels. In this paper we address the problem of primary-secondary spectrum sharing in cognitive radio access networks using a framework based on a Discrete Time Markov Chain (DTMC) model. Its applicability and advantages with respect to other approaches is explained and further justified. Spectrum awareness of primary activity by the secondary users is based on spectrum sensing techniques, which are modeled in order to capture sensing errors in the form of false-alarm and missed-detection. Model validation is successfully achieved by means of a system-level simulator which is able to capture the system behavior with high degree of accuracy. Parameter dependencies and potential tradeoffs are identified enabling an enhanced operation for both primary and secondary users. The suitability of the specified model is justified while allowing a wide range of extended implementations and enhanced capabilities to be considered.

1. Introduction

The key purpose of dynamic spectrum management is to maximize spectrum reuse amongst users while ensuring that mutual interference between them remains at acceptable levels [1]. This notion has been motivated by the sporadic use of particular spectrum bands while others are profusely used. In this sense, the traditional fixed spectrum assignment to a licensee which has exclusive exploitation rights for a particular spectrum range may not be satisfactory to respond to the new radio use context that requires enhancement in spectrum efficiency and can lead to spectrum underutilization [2]. Consequently, new technical advances are focused on the development of strategies and policies aiming to the utmost and efficient access to shared spectrum resources. Such new developments are usually coined under the global term of *cognitive radio networks*, which includes a set of different approaches and implementation alternatives [3].

In this paper we tackle the problem of dynamic spectrum access considering the Hierarchical Access Model [3], where the licensed (or primary) spectrum is opened to secondary users (SUs) provided the interference over the primary users (PUs, or licensees) is kept under acceptable limits. In addition, two approaches for spectrum sharing have been devised: Spectrum Underlay and Spectrum Overlay. Spectrum underlay aims at operating below the floor noise of primary users by using ultra-wideband (UWB) techniques which, on the other hand, limits the transmitted power by secondary users. As for spectrum overlay, it targets at spatio-temporal spectrum holes by allowing secondary users to identify and exploit them in a non-intrusive manner. In the remainder of the paper, it

* Corresponding author. Tel.: +34 625790691.
E-mail addresses: xavier.gelabert@tsc.upc.edu (X. Gelabert), sallent@tsc.upc.edu (O. Sallent), jorperez@tsc.upc.edu (J. Pérez-Romero), ramon@tsc.upc.edu (R. Agustí).

1389-1286/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2010.04.005

will be assumed that spectrum overlay is used as a basis of our model.

From a regulatory perspective, the Federal Communications Commission (FCC) in the U.S. and Ofcom in the U.K. are currently considering the use of cognitive radio technologies [4]. Accordingly, the unlicensed use of VHF and UHF TV bands, provided no harmful interference is caused, was targeted by the FCC in [5]. This was a first milestone in the development of the IEEE 802.22 standard, proposing a cognitive radio-based physical and medium access control (MAC) layer for use of TV spectrum bands by license-exempt devices on a non-interfering basis [6]. Furthermore, the IEEE activities in developing architectural concepts and specifications for network management interoperability, including CR and dynamic spectrum access, are addressed by SCC41/P1900 standardization groups [7]. Finally, many operative standards such as WiFi (IEEE 802.11), Zigbee (IEEE 802.15.4), and WiMAX (IEEE 802.16) already include some degree of CR technology today [4], in the form of coexistence among radios, Dynamic Frequency Selection (DFS) and Power Control (PC).

The primary-secondary (P-S) spectrum sharing operation can take the form of cooperation or coexistence. Cooperation means there is explicit communication and coordination between primary and secondary systems, and coexistence means there is none [8]. When sharing is based on coexistence, secondary devices are essentially invisible to the primary. Thus, all of the complexity of sharing is handled by the secondary and no changes to the primary system are needed. Among the different forms of coexistence, we adopt the opportunistic exploitation of white spaces in spatial–temporal domain sustained on spectrum sensing, coordination with peers and fast spectrum handover, i.e. the spectrum overlay case. As for cooperation, again different forms of P-S interactions are possible. For example, spatial–temporal white spaces can be signaled through a common control channel from the primary network side, such as the Cognitive Pilot Channel (CPC) or the CSCC (Common Spectrum Coordination Channel) [9–14], which would provide primary spectrum usage information to SUs. In addition, the interaction between PUs and SUs provides an opportunity for the license-holder to demand payment according to the different quality-of-service grades offered to SUs.

In the abovementioned context, the use of Markov models becomes an important aid in modeling problems dealing with the dynamic access to shared spectrum resources. In this sense, a significant number of papers in the literature have been devoted to the characterization of such scenarios using Markov models as, e.g., in [15–21][1].

Work in [15–17], which employ similar CTMC-based models considering infinite, [15], and finite, [16,17], population models, assume perfect spectrum sensing conditions, i.e. free from sensing errors. In this respect, our contribution goes further in considering the effect of erroneous sensing given by false-alarm and missed-detection probabilities.

In turn, work in [18,19] also assume CTMC-based models. Therefore, as the transition rates of the CTMC indicate, these works consider that sensing information is instantly available upon user arrival. In this paper, the DTMC allows to capture the sensing instants and the effect of sensing information ageing into the model. This is because in a DTMC we observe the system at discrete time instants which, in our proposed model, correspond to the periodic sensing instants. In addition, work in [18,19] considers false-alarm and missed-detection probabilities as numerical inputs with no explicit reference to any particular sensing mechanism (e.g. energy detection, pilot detection, etc. [23]). Conversely, our work adopts an energy-based detector for sensing implementation in Rayleigh fading, [24,25] from which false-alarm and missed-detection probabilities are extracted, thus offering a wider applicability range and a higher degree of practicality.

Finally, although not strictly related to our work, in [20,21], CTMC models are used to characterize the interactions between primary and secondary users where random spectrum access protocols, as opposed to channelization schemes considered herein, are proposed and evaluated. In this sense, sensing errors along with sensing periodicity and ageing issues are not considered.

In this work, a Markovian framework based on Discrete Time Markov Chains (DTMC) to evaluate the opportunistic spectrum access in a P-S spectrum sharing scenario is proposed. The rationale behind using DTMCs instead of CTMCs is based on the fact that sensing mechanisms operate on a periodic time basis, and where the sensing periodicity is an important design parameter. Therefore, the DTMC models, which observe the state of the system at discrete-time instants, can accurately model the proposed scenarios by considering the observation instants of the DTMC as the sensing instants.

Model validation and evaluation studies considering several parameter dependency issues and tradeoffs are addressed in this paper revealing the usefulness of the proposed model for cognitive radio networks system design, realization and operation. In particular, relevant parameters are identified that influence the performance of the spectrum sharing model. Among these, sensing periodicity (*how often do we sense?*) and sensing accuracy (*how well do we sense?*) are shown to be key parameters that greatly affect the behavior of the system. In addition, this work reflects the importance of time-sharing between spectrum sensing (*for how long do we sense?*) and data transmission (*for how long do we transmit?*), which tradeoffs the sensing accuracy with the obtained throughput, thus leading to possible parameter optimization which will be also addressed in this work. Finally, the awareness of both primary and secondary traffic load distributions also enables to identify optimized parameter values for an overall enhanced network operation as will be shown in the following.

The remainder of the paper is organized as follows. In Section 2 the system model is described along with the considered procedures and the implementation approach. Subsequently, in Section 3, the DTMC model is formulated along with the main hypothesis and considerations. A number of relevant performance metrics are derived in Section 4 which will be evaluated numerically in Section 5. Finally, Section 6 concludes the paper with some final

[1] Work in [15] should be considered along with amendments in [22].

remarks and future considerations. For the sake of readability, in Appendix A, Table A.2 contains the notation used in this paper.

2. System model

The considered system involves a Primary Network (PN), serving PUs, and a Secondary Network (SN), serving SUs. Both the PN and the SN operate autonomously and each network implements efficient protocols for the correct and coordinated operation among their own users (i.e. PUs and SUs respectively). Thus, the PN is aware of the spectrum occupancy by PUs and, correspondingly, the SN is aware of the spectrum occupancy of SUs. The PN has been assigned a total number of C channels, partitioning a certain frequency bandwidth. SUs can make use of free channels; though PUs have strict priority over SUs (i.e. if a SU is using a given channel and this channel is required by a PU, then the SU must release it). Fig. 1 shows an example of the considered bandwidth sharing model. It is worth noting that the present model could be conveniently extended to consider the case of multiple primary networks by considering independent arrival distributions to different channels.

2.1. Procedure

The procedure to be followed for the operation in the SN and the corresponding required functionalities is presented in the following. In the first stage, a frequency band and a specific available channel where a secondary communication can be established have to be identified. Then, both secondary communication ends have to be configured to be able to transmit and receive over the identified channel. While maintaining the secondary communication, it is required that the presence of a primary communication is detected, so that if a PU arrives the secondary communication must evacuate the channel. Spectrum handover (SpHO) procedures will intend to find an appropriate alternative channel where the secondary communication can be continued in order to avoid its interruption.

2.2. Spectrum awareness implementation approach

Among the two aforementioned implementation approaches (i.e. coordination vs. coexistence) the coexistence case will be adopted in the sense that the SN implements spectrum discovery mechanisms in order to exploit unused spectrum in an opportunistic fashion. In this approach, the identification of a candidate frequency band (or channel) for the secondary communication as well as the monitoring of primary's presence is performed within the SN based on sensing mechanisms without any direct interaction with PN. Depending on the secondary network architecture, whether it is infrastructure-based (centralized) or infrastructure-less (i.e. decentralized or *ad hoc*), sensing information may be gathered in different forms. In the centralized case, SUs equipped with sensors may sense the whole spectrum and report to a centralized entity (e.g. located at the Base Station, BS) which may in turn schedule or re-schedule SU transmissions accordingly. Alternatively, a centralized entity at the BS may be responsible for sensing tasks, thus alleviating SUs from sensing capabilities. For the decentralized approach, spectrum sensing and use is entirely handled by SUs, thus information exchange mechanisms among them should be implemented. For the sake of simplicity, and to avoid a further increase in the model complexity, we will assume the centralized case where a centralized entity is responsible for spectrum sensing tasks. Nevertheless, the presented model allows further implementation alternatives to be considered which is left for future work.

Channel occupancy detection performed at the SU's terminal side through sensing mechanisms is affected by a number of aspects (e.g. adverse channel conditions, hidden terminal problem, limited sensitivity on the sensing equipment, etc.) that may limit the reliability of sensing results [24]. Typically, spectrum detection through sensing in the presence of errors performs a binary hypotheses test over a given band (or channel), that is: $\mathcal{H}_0$ if the channel is available and $\mathcal{H}_1$ if the channel is occupied. Accordingly, the miss-detection and false-alarm probabilities, δ and ε, can be defined as:

$$\delta = \Pr[\mathcal{H}_0|\mathcal{H}_1 \text{ is true}] \quad (1)$$

$$\varepsilon = \Pr[\mathcal{H}_1|\mathcal{H}_0 \text{ is true}]. \quad (2)$$

An appropriate selection of the so-called *time-bandwidth product*, defined as

$$m = T \cdot W, \quad (3)$$

where T is the time devoted to sense bandwidth W [24], is of great relevance. In general, the longer we sense the bandwidth W seeking for spectrum opportunities the more reliable are our sensing measures (i.e. lower δ and ε values), however, high T values, as shown further on, will trade-off the achievable throughput experienced by SUs [26]. Moreover, spectrum errors can be improved by means of the cooperation of sensing entities as suggested in [24].

With respect to the availability of updated primary spectrum occupancy information based on sensing, ΔT would represent the time between two consecutive sensing information updates. Considering that a generic underlying MAC level time-frame structure enabling sensing would devote some time, T_{sens}, for sensing purposes, i.e. the *sensing time*, the *sensing efficiency* can be defined as

$$\eta_{sens} = 1 - T_{\text{sens}}/\Delta T. \quad (4)$$

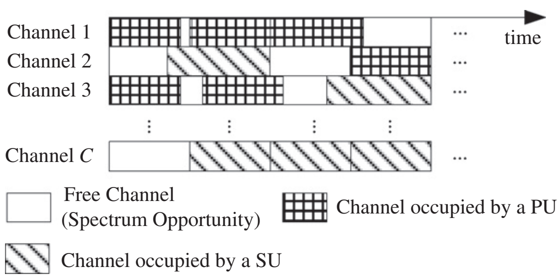


Fig. 1. Bandwidth partitioning model for spectrum sharing between PUs and SUs.

We assume that time T_{sens} will be the time T, devoted to sense a single channel, multiplied by the number of channels with bandwidth W that should be sensed, i.e.

$$T_{sens} = T \cdot C. \quad (5)$$

To account for possible detection errors during spectrum sensing procedures, a probabilistic model will be developed in order to compute the number of detected (or sensed) PUs in the system. This model will consider specific missed-detection and false-alarm probabilities values, δ and ε, obtained from well-known expressions in the literature, see e.g. [24,25].

3. DTMC model formulation

The proposed DTMC model is devoted to determine the statistical occupancy of the shared spectrum by PUs and SUs. It is mainly fed by traffic-related input parameters, such as arrival and departure rates (λ_p and λ_s along with μ_p and μ_s for PUs and SUs correspondingly), and also the number of channels to be shared, C.

It is assumed that arrival processes follow a Poisson distribution and that service times are exponentially distributed. While adopting general distributions for arrival and departure processes would be interesting, this would increase the already complex model notation and formulation. On the other hand, service-types (i.e. voice, web-browsing, streaming, etc.) for PUs and SUs are not specified, since it falls out of the scope of this paper, thus Poisson and exponential distributions can be safely assumed as also considered in many works in the literature such as [15–17,19–21].

The sensing periodicity, ΔT, which denotes the periodic time instants in which updated spectrum occupancy information is made available for secondary communication, is, on the other hand, the operating time-basis of the DTMC.

The proposed DTMC model accounts for the spectrum usage of PUs and SUs in a shared spectrum scenario. For simplicity reasons, it is supposed that the whole spectrum bandwidth is partitioned into a total number of C channels (bands) to be shared among both PUs and SUs. It is further considered that both PUs and SUs demand a single channel for transmission purposes (recall Fig. 1). These assumptions, although simplifying, will keep the algebra at an understandable and tractable level while still capturing the essence of the problem under study. If desirable, more elaborate shared bandwidth models can be easily considered and adapted to the model here presented (e.g., considering different bandwidth requirements for PUs and SUs).

In a DTMC, [27], we observe the system state at discrete-time instants $\{t_0, t_1, t_2, \ldots, t_n, \ldots\}$, with $t_n = t_0 + n \cdot \Delta T$ and periodicity ΔT, which is, on the other hand, assumed to specify the time instants where primary spectrum usage information is made available for secondary communication use. In addition, let $I_n = (t_n, t_{n+1}]$ define the nth time interval between two successive observation times. Note that while ΔT specifies the sensing periodicity (in seconds), I_n refers to a particular time-interval (of length ΔT secs.). The DTMC model formulation involves a number of steps which are presented in the following subsections.

3.1. State space definition

Let $N_p(t_n)$ and $N_s(t_n)$ be stochastic processes indicative of the number of PUs and SUs in the system at time t_n. It is further assumed that users remaining in the system have always data to transmit, i.e. are always active, except for SUs which are refrained from transmission during sensing periods of length T_{sens}. Accordingly, allow $\mathbf{X}_n = S_{(i,j)} = \{N_p(t_n) = i, N_s(t_n) = j\}$ to represent a state of the DTMC at time t_n. Thus, if C channels are available, the considered state space $\mathcal{S}$ must contain all possible sates $S_{(i,j)}$ which fulfill both $i \leqslant C$ and $j \leqslant C$, formally:

$$\mathcal{S} = \{S_{(i,j)} : i \leqslant C, j \leqslant C\}. \quad (6)$$

Nevertheless, for a correct spectrum use (i.e. with no spectrum collisions), the number of PUs (i) plus the number of SUs (j) must not exceed the total number of available channels (C). In addition, due to spectrum detection errors, a SU might be erroneously assigned to a band already in use by a PU. Then, for convenience, we define the following three subsets of $\mathcal{S}$ accounting for those states that necessarily imply spectrum collision, i.e.

$$\mathcal{S}_c = \{S_{(i,j)} : i + j > C\} \subset \mathcal{S}, \quad (7)$$

those states which possibly imply a spectrum collision, i.e.

$$\mathcal{S}_{pc} = \{S_{(i,j)} : i + j \leqslant C, j > 0, i > 0\} \subset \mathcal{S}, \quad (8)$$

and those states that are collision-free, i.e.

$$\mathcal{S}_{nc} = \{(S_{(i,j)} : i = 0) \cup (S_{(i,j)} : j = 0)\} \subset \mathcal{S}. \quad (9)$$

Note that in absence of spectrum sensing errors, transitions to states belonging to $\mathcal{S}_c$ will not apply, since the secondary network is fully aware of primary spectrum occupancy and will consequently block further SUs attempting access. On the contrary, sensing errors may lead transitions to states belonging to $\mathcal{S}_c$ caused by PU miss-detections.

Fig. 2 shows an example of the considered state space for $C = 3$ channels, where it must be regarded that $\mathcal{S} = \mathcal{S}_c \cup \mathcal{S}_{pc} \cup \mathcal{S}_{nc}$ and $\mathcal{S}_c \cap \mathcal{S}_{pc} \cap \mathcal{S}_{nc} = \emptyset$. In addition, the total number of states is given by $N_{states} = (C + 1)^2$.

A clear advantage in considering the total number of PUs (i) and SUs (j) in the system as indices of the Markov state space (as, e.g. in [15]) is that system dimensionality (and consequently complexity) of the model can be reduced. Contrarily, if the status of each channel (that is, free, occupied by PU or occupied by SU) was to be captured by the Markov model, the number of states would dramatically increase thus limiting the applicability of the model. On the other hand, the main drawback of adopted approach is that it remains uncertain how different users are distributed over the channels. Fortunately, this can be solved by considering probabilistic models for the occupation of different channels by PUs and SUs.

3.2. Detection of primary spectrum occupancy

At a particular time t_n, let the state of the DTMC be $\mathbf{X}_n = S_{(i,j)} \in \mathcal{S}$. At this same time instant, spectrum occupancy information is made available to the SN side (either

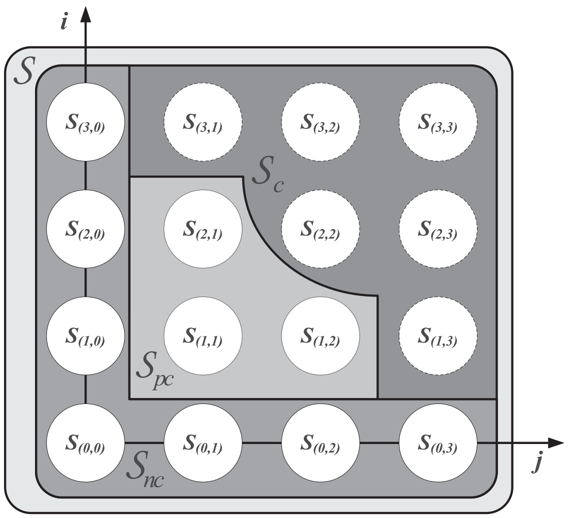


Fig. 2. The state space for $C = 3$, where different shaded regions determine the spectrum collision states $[S_{(i,j)} \in \mathcal{S}_c]$, possible collision states $[S_{(i,j)} \in \mathcal{S}_{pc}]$ and collision-free states $[S_{(i,j)} \in \mathcal{S}_{nc}]$

to some centralized infrastructure-based entity or to a specific SU). Due to spectrum detection errors, the observed state at time t_n using such erroneous information may be $\mathbf{Y}_n = S_{(k,j)} \in \mathcal{S}$, i.e. $\mathbf{Y}_n \neq \mathbf{X}_n$, with k denoting the number of detected PUs (note the number of SUs at time t_n, j, is known by the SN, so it is not subject to errors). Consequently, we are interested in determining the conditional probability of detecting k PUs when there are in fact i PUs in the system at time t_n, which may formally be expressed as:

$$b_{(k,i)} = \Pr[\mathbf{Y}_n = S_{(k,j)} | \mathbf{X}_n = S_{(i,j)}]. \tag{10}$$

Theorem 1. *The conditional primary user detection probability, $b_{(k,i)}$, subject to false-alarm and missed-detection probabilities, ε and δ, is given by*

$$b_{(k,i)} = \sum_{m=\max(0,i-k)}^{\min(i,C-k)} \binom{C-i}{m+k-i} \cdot \varepsilon^{m+k-i} \cdot \bar{\varepsilon}^{C-m-k} \cdot \binom{i}{m} \cdot \delta^m \cdot \bar{\delta}^{i-m}, \tag{11}$$

with $\bar{\varepsilon} = 1 - \varepsilon$ and $\bar{\delta} = 1 - \delta$.

Proof. See Appendix B. □

Then, function $b_{(k,i)}$ provides the application function between the so-called *true state space* given by states $\mathbf{X}_n = S_{(i,j)} \in \mathcal{S}$ and the *detected state space* given by states $\mathbf{Y}_n = S_{(k,j)} \in \mathcal{S}$. Since the SN operation will be based on the knowledge of $\mathbf{Y}_n$ as opposed to $\mathbf{X}_n$, the values of ε and δ will considerably affect the performance of such system and lead, in the worst case, to an ineffective operation.

3.3. Arrival and departure processes

Let $N^A \in \{N^{PA}, N^{SA}\}$ along with $N^D \in \{N^{PD}, N^{SD}\}$ denote the number of arrivals and departures of PUs and SUs respectively in I_n (i.e. in a time interval of duration ΔT).

Given PUs and SUs arrive at the system according to a Poisson distribution with rates λ_p and λ_s respectively, the probability that k arrivals occur in I_n, P_k^A, is given by [27]:

$$P_k^A = \Pr[N^A = k] = [(\lambda \Delta T)^k / k!] e^{-\lambda \Delta T}, \tag{12}$$

where for $\lambda \in \{\lambda_p, \lambda_s\}$ we will refer to $P_k^A \in \left\{P_k^{PA}, P_k^{SA}\right\}$ correspondingly.

If the session duration is exponentially distributed with mean $1/\mu$ (i.e., rate μ), the probability of a session departure in I_n is [27]:

$$P^D = 1 - e^{-\mu \Delta T}. \tag{13}$$

Then, the probability of having k-out-of-m departures in I_n, P_k^D, is given by the binomial distribution [27]:

$$P_k^D = \Pr[N^D = k] = \binom{m}{k} (1 - e^{-\mu \Delta T})^k (e^{-\mu \Delta T})^{m-k}, \tag{14}$$

where for $\mu \in \{\mu_p, \mu_s\}$ we will refer to $P_k^D \in \left\{P_k^{PD}, P_k^{SD}\right\}$ respectively.

Note that enabling multiple arrivals and departures in one ΔT period will affect the decision process on whether a SU can be assigned or not given that detection information is retrieved only at times t_n. This also constitutes a differentiating aspect with respect to other approaches to the same problem such as in [15–17,19–21].

3.4. Transition probabilities

The transition probabilities between each pair of states $S_{(k,l)} \rightarrow S_{(i,j)}$ in the DTMC model can be expressed as [27]:

$$\begin{aligned} P_{(i,j|k,l)} &= \Pr[X_{n+1} = S_{(i,j)} | X_n = S_{(k,l)}] = \Pr[N_p(t_{n+1}) \\ &= i, N_s(t_{n+1}) = j | N_p(t_n) = k, N_s(t_n) = l] \\ &= \Pr[N_p(t_{n+1}) = i | N_p(t_n) = k, N_s(t_n) \\ &= l] \times \Pr[N_s(t_{n+1}) = j | N_p(t_n) = k, N_s(t_n) = l], \end{aligned} \tag{15}$$

where the conditional independence of processes $N_p(t_n)$ and $N_s(t_n)$ has been considered (since primary and secondary arrival/departure processes are also assumed independent). Probabilities $P_{(i,j-k,l)}$ constitute the elements of the transition probability matrix $\mathbf{P}$, from which the steady-state probabilities, $P_{(i,j)}$, of the DTMC will be determined [28]. For the sake of algebra tractability, the following assumptions are considered in the presented expressions:

Hypothesis 1. A primary or secondary session arriving in I_n will not depart in the same I_n. This implies that $\Delta T \ll 1/\mu$ with $\mu \in \{\mu_p, \mu_s\}$ and where $1/\mu$ is the average session duration.

Hypothesis 2. We disregard the order in which session arrivals and departures occur in a given I_n by considering the resulting net number of users, i.e. those obtained after subtracting the departures and adding the new arrivals.

Note that both previous hypotheses rely on the relation between ΔT and traffic-related parameters λ and μ, which can be adjusted to fit our needs. The validation and justification of these hypotheses will be addressed in Section 5 when we deal with the model validation. In particular for

Hypothesis 2, and to be consistent with the fact that the order of arrivals and departures may significantly affect the performance of the system, we shall provide specific values for the sensing periodicity ΔT for which the assumption of Hypothesis 2 can be safely considered.

For convenience we define the following set of lemmas:

Lemma 1. *The probability of assigning k PUs when we have also l PU de-assignments in state $S_{(i,j)}$, $a^P_{(i,j,k,l)}$, is given by:*

$$a^P_{(i,j,k,l)} = \begin{cases} P^{PA}_k, & \text{if } i-l+k<C \\ 1-\sum_{m=0}^{k-1} P^{PA}_m, & \text{if } i-l+k=C \end{cases}. \tag{16}$$

Proof. See Appendix C.1. □

Lemma 2. *The probability of de-assigning k PUs in state $S_{(i,j)}$, $d^P_{(i,j,k,l)}$, is given by:*

$$d^P_{(i,j,k)} = P^{PD}_k. \tag{17}$$

Proof. See Appendix C.2. □

Lemma 3. *The probability of assigning k SUs when we have also l SU de-assignments in state $S_{(i,j)}$, $a^S_{(i,j,k,l)}$, is given by:*

$$a^S_{(i,j,k,l)} = \begin{cases} \sum_{m=0}^{C-k-j+l} \bar{a}^S_{(m,j,k,l)} \cdot b_{(m,i)} & \text{for } k>0 \\ \sum_{m=0}^{C-j+l} \bar{a}^S_{(m,j,0,l)} \cdot b_{(m,i)} + \sum_{m=C-j+l}^{C} b_{(m,i)} & \text{for } k=0 \end{cases}, \tag{18}$$

with

$$\bar{a}^S_{(m,j,k,l)} = \begin{cases} P^{SA}_k, & \text{if } m+j-l+k<C \\ 1-\sum_{r=0}^{k-1} P^{SA}_r, & \text{if } m+j-l+k=C \end{cases}. \tag{19}$$

Proof. See Appendix C.3. □

Lemma 4. *The probability of de-assigning k SUs in state $S_{(i,j)}$, $d^S_{(i,j,k,l)}$, is given by:*

$$d^S_{(i,j,k)} = \sum_{r=0}^{k} d^{S,S}_{(i,j,k-r,r)} \cdot d^{S,SC}_{(i,j,r)}, \tag{20}$$

with

$$d^{S,S}_{(i,j,k,l)} = \begin{cases} b_{(C+k-j+l,i)} & \text{if } 0<k\leqslant j-l \\ 1-\sum_{r=1}^{j-l} b_{(C+r-j+l,i)} & \text{if } k=0 \end{cases}, \tag{21}$$

and

$$d^{S,SC}_{(i,j,k)} = P^{SD}_k. \tag{22}$$

Proof. See Appendix C.4. □

Then, it follows that:

Theorem 2. *The transition probability between states $S_{(i,j)} \to S_{(i+N,j+M)}$, $P_{(i+N,j+M|i,j)}$, with $-i \leqslant N \leqslant C-i$ and $-j \leqslant M \leqslant C-j$, is given by:*

$$P_{(i+N,j+M|i,j)} = \left(\sum_{k=\max(-N,0)}^{i} a^P_{(i,j,N+k,k)} \cdot d^P_{(i,j,k)}\right) \cdot \left(\sum_{k=\max(-M,0)}^{j} a^S_{(i,j,M+k,k)} \cdot d^S_{(i,j,k)}\right), \tag{23}$$

Proof. See Appendix D. □

4. Performance metrics

From the resulting transition probability matrix $\mathbf{P}$ defined through (23), we obtain the true steady-state probabilities, $P_{(i,j)} = \lim_{n\to\infty}\Pr[\mathbf{X}_n = S_{(i,j)}]$, for each true state $S_{(i,j)}$ in the state space $\mathcal{S}$. The knowledge of such statistical distribution enables the definition of several performance metrics which are addressed in the following.

On the other hand, it is also relevant to determine the steady-state probabilities of the detected states (i.e. including possible sensing errors): $P'_{(i,j)} = \lim_{n\to\infty}\Pr[\mathbf{Y}_n = S_{(i,j)}]$, which are computed as:

$$P'_{(i,j)} = \sum_{n=0}^{C} b_{(i,n)} \cdot P_{(n,j)}. \tag{24}$$

In this case, $P'_{(i,j)}$ is the steady-state probability observed by the SN, i.e. without true knowledge of PU activity and, therefore, sensible to sensing errors. Then, by considering $P'_{(i,j)}$instead of $P_{(i,j)}$, metrics computed from the SN side, which account for possible sensing errors, can be obtained.

In order to obtain probabilities $P_{(i,j)}$, we apply numerical methods [28] so as to solve the matrix equation given by $v = v\cdot\mathbf{P}$, where

$$v = [P_{(0,0)},\ldots,P_{(0,C)},P_{(1,0)},\ldots,P_{(1,C)},\ldots,P_{(C,0)},\ldots,P_{(C,C)}]$$

is the steady-state probability vector.

4.1. Average number of users

The average number of PUs and SUs (i.e. average served traffic) is computed as:

$$N_p = \sum_{S_{(i,j)}\in\mathcal{S}} i \cdot P_{(i,j)}, \tag{25}$$

and

$$N_s = \sum_{S_{(i,j)}\in\mathcal{S}} j \cdot P_{(i,j)}. \tag{26}$$

In addition, we can compute the average number of sensed PUs, N'_p, by considering $P'_{(i,j)}$, thus leading to:

$$N'_p = \sum_{S_{(i,j)}\in\mathcal{S}} i \cdot P'_{(i,j)}. \tag{27}$$

4.2. Blocking probability

Blocking occurs whenever a new user cannot be assigned a channel given all channels are occupied or thought to be occupied. Consequently, blocking probability for SUs can be computed as:

$$P_B^S = \sum_{i=0}^{C} \sum_{j=C-i}^{C} P'_{(i,j)}, \tag{28}$$

with $P'_{(i,j)}$ defined in (24).

4.3. Interruption probability

Interruption of secondary service occurs whenever a SU is forced to release a channel before its session has ended due to the appearance of a PU and no other channel is sensed to be free. To compute the interruption probability the average number of secondary users defined in (26) can be regarded as the average served SU traffic $T_s^{\text{served}} = N_s$. Then, we can express

$$T_s^{\text{served}} = T_s \cdot \left(1 - P_B^S\right) \cdot (1 - P_D), \tag{29}$$

indicating that the served traffic is the offered traffic which is not blocked nor interrupted. From (29), we express the interruption probability as:

$$P_D = 1 - \frac{T_s^{\text{served}}}{T_s\left(1 - P_B^S\right)} = 1 - \frac{N_s}{\frac{\lambda_s}{\mu_s}\left(1 - P_B^S\right)}. \tag{30}$$

4.4. Interference probability: an upper-bound

At a given sensing instant in state $S_{(i,j)} \in \mathcal{S}$, the probability of miss-detecting n PUs out of i PUs in the system can be expressed as the binomial distribution with δ the probability of miss-detection (refer to Appendix B). Consequently, the average number of missed-detections in state $S_{(i,j)} \in \mathcal{S}$ is given by

$$N_{MD}(i,j) = i \cdot \delta. \tag{31}$$

The average number of collisions in state $S_{(i,j)} \in \mathcal{S}_{pc}$ (recall from Section 3.1) may be initially upper-bounded by

$$N_c(i,j) \leqslant N_{MD}(i,j), \tag{32}$$

indicating that at most (i.e. in the worst case) we will have a collision for each missed-detection. In addition, this upper-bound can be tightened by considering that the average number of collisions will be also less than the average number of SUs in state $S_{(i,j)}$, i.e. $N_c(i,j) \leqslant j$. Hence, we have

$$N_c(i,j) \leqslant \min[i \cdot \delta, j] \triangleq N_c^*(i,j). \tag{33}$$

The number of collisions in state $S_{(i,j)} \in \mathcal{S}_c$ (i.e. with $i + j > C$) can be initially lower-bounded as

$$N_c(i,j) \geqslant (i + j - C) \triangleq \kappa_c, \tag{34}$$

where we know that at least κ_c channels are being simultaneously shared by both a PU and a SU.

In addition, remaining $(i - \kappa_c)$ PUs and $(j - \kappa_c)$ SUs may be also in a collision situation, hence the number of collisions can be upper-bounded, similar to (33), as

$$N_c(i,j) \leqslant \kappa_c + \min[(i - \kappa_c) \cdot \delta, (j - \kappa_c)] \triangleq N_c^*(i,j). \tag{35}$$

As for the case where $S_{(i,j)} \in \mathcal{S}_{nc}$, the number of collisions is zero, i.e. $N_c(i,j) = 0$.

Given the maximum possible number of collisions would be C, we define the collision probability ratio, hereon the *interference probability*, in state $S_{(i,j)}$ as

$$P_c(i,j) = \frac{N_c^*(i,j)}{C}, \tag{36}$$

where we have used collision upper-bounds $N_c^*(i,j)$ provided in (33) and (35) as worst cases, along with $N_c^*(i,j) = 0$ for $S_{(i,j)} \in \mathcal{S}_{nc}$.

Finally, the *average interference probability* can be then expressed as

$$P_c = \sum_{S_{(i,j)} \in \mathcal{S}} P_c(i,j) \cdot P_{(i,j)}. \tag{37}$$

4.5. Throughput

In a given state $S_{(i,j)} \in \mathcal{S}$, the number of channels that are being simultaneously used by both a PU and a SU, $N_c^*(i,j)$, has been computed in the previous subsection.

It is considered that a channel being shared by both a PU and a SU does not contribute to throughput; consequently we can define the throughput of PUs or SUs in state $S_{(i,j)} \in \mathcal{S}$ as:

$$\Gamma^p_{(i,j)} = \left[i - N_c^*(i,j)\right] \cdot R_p, \tag{38}$$

along with

$$\Gamma^s_{(i,j)} = \left[j - N_c^*(i,j)\right] \cdot R_s, \tag{39}$$

measured in bits per second (bps), where R_p and R_s are the average bit rate per channel for PUs and SUs respectively. Given that SUs need some time to sense the medium (or to remain silent so that the sensing entity can sense the medium) the net throughput per channel achieved by SUs, R_s, will be lower than the net throughput per channel achieved by PUs, R_p, under our assumption that the same amount of bandwidth is devoted to both users. Then, we can write,

$$R_s = R_p \cdot \eta_{sens}, \tag{40}$$

where $\eta_{sens} \in [0,1]$ is the sensing efficiency defined in (4). The granted bit rate for a single PU will be simply computed using the Shannon bound as

$$R_p = W \cdot \log_2(1 + \bar{\gamma}), \tag{41}$$

where W (Hz) is the bandwidth of a single channel and $\bar{\gamma}$ (dB) is the average signal-to-noise ratio (SNR).

Finally, the average throughput for both PUs and SUs is computed as

$$\Gamma^p = \sum_{S_{(i,j)} \in \mathcal{S}} \Gamma^p_{(i,j)} \cdot P_{(i,j)}, \tag{42}$$

along with

$$\Gamma^s = \sum_{S_{(i,j)} \in \mathcal{S}} \Gamma^s_{(i,j)} \cdot P_{(i,j)}. \tag{43}$$

5. Performance evaluation

In the following, the considered parameter setup for the numerical evaluation and also some implementation aspects for the uncoordinated spectrum awareness case are introduced. Subsequently, numerical results address, in the first place, the model validation by means of a system-level simulator. Secondly, numerical results will be given so as to capture the tradeoff between the time devoted to sensing and the throughput, the impact on the number of considered channels (C) and, finally, the effect on the spectrum awareness periodicity ΔT.

5.1. Parameter setup

As an initial reference, it is considered a total bandwidth partitioned into $C = 16$ channels to be shared amongst PUs and SUs. The offered primary traffic load is $T_p = \lambda_p/\mu_p = \{5, 10\}$ Erlangs (E), while secondary offered traffic, defined as $T_s = \lambda_s/\mu_s$, will span over a specified range of values. Note that for $T_p = 10$E and $C = 16$ channels the system is significantly loaded by PUs, so that the system is evaluated under rather unfavorable conditions for SUs. Average session duration is assumed to be equal for both PUs and SUs with $1/\mu = 1/\mu_p = 1/\mu_s = 120$ s (i.e. assuming a time-based service). Sensing periodicity is, unless otherwise stated, $\Delta T = 100$ ms. The characterization of sensing errors will be provided by means of missed-detection and false-alarm probabilities, i.e. δ and ε respectively (the following subsection addresses such implementation in detail). In addition, for comparative purposes, we will also consider the case where perfect sensing (errors-free) is available. For this case, not only δ and ε are zero but also the time devoted to sensing purposes is considered to be zero.

5.2. Spectrum detection in rayleigh fading

The detection of unknown signals by means of sensing has captured a lot of attention in the past, and the advent of cognitive radio has indeed contributed to increase the work devoted to this matter [24,25].

In general, sensing in the presence of errors due to imperfect channel conditions can be characterized through miss-detection and false-alarm probabilities (i.e. δ and ε respectively) extracted from so-called Receiver Operating Curves (ROC) which relate both magnitudes. These curves mainly depend on the considered channel model (Rayleigh, Rician, etc.), the average SNR ($\bar{\gamma}$) at the sensor's end and the time-bandwidth product (m) previously introduced in (3). Fig. 3 shows the ROC curves for the energy detector under Rayleigh fading considering a number of time-bandwidth products (m) which have been computed according to analytical expressions in [24].

Note that, for a fixed observed bandwidth W, the increase of the time devoted to sensing, T, i.e. also meaning an increase in the time-bandwidth product (recall that $m = T{\cdot}W$), will mean that a more accurate sensing information is retrieved. In this sense, setting different target values for the miss-detection probability δ^* we obtain the corresponding values for the false-alarm probability ε^* for the different time-bandwidth products m which define the working point of the receiver. In practice, this is achieved by setting appropriate values for the decision threshold at the sensor's end [24]. Without loss of generality we assume a fixed value of $W = 200$ kHz and a target miss-detection probability value of $\delta^* = 0.01$. Unless otherwise stated the average SNR is chosen to be $\bar{\gamma} = 10$ dB. Then, and for the sake of representation, performance results will refer to particular values of m (which in turn refer to specific values of δ^* and ε^* as shown in Table 1).

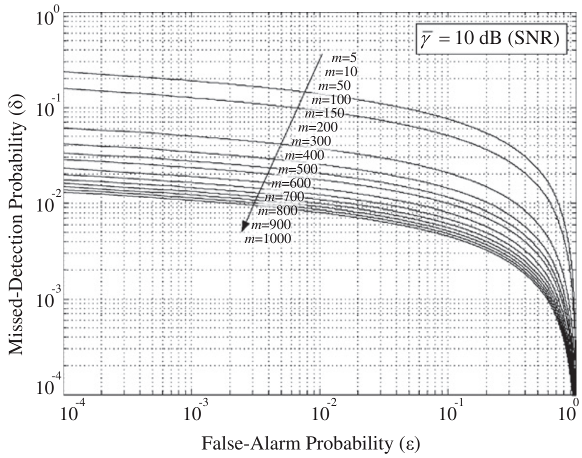


Fig. 3. Receiver Operating Characteristic (ROC) curves for the energy detector in Rayleigh fading considering $\bar{\gamma} = 10$ dB and time-bandwidth products m between 5 and 1000.

Table 1
ROC values for $\bar{\gamma} = 10$ dB and $W = 200$ kHz extracted from Fig. 3.

m	δ^*	ε^*	$T = m/W$
5	0.01	0.7277	0.000025
10	0.01	0.6278	0.00005
50	0.01	0.3378	0.00025
100	0.01	0.2073	0.0005
150	0.01	0.1391	0.00075
200	0.01	0.0974	0.001
300	0.01	0.0511	0.0015
400	0.01	0.0281	0.002

5.3. Numerical results

5.3.1. Validation and preliminary model assessment

One main objective of analytical system modeling is to retain the basic interactions between the main parameters affecting the performance of a real system. How close is the performance behavior of the model with respect to the real system will largely depend on the considered model assumptions and hypothesis. In this sense, in this section we intend to identify those parameters that may limit the applicability of the model, which is, on the other hand, inherent to all analytical models.

For the sake of a complete model validation, a system-level simulator has been developed so as to extract relevant metrics and compare them to those defined for the model in Section 4. Specifically, this simulator operates on a discrete-time-driven basis. Time granularity is set several orders of magnitude smaller than the sensing periodicity in order to capture higher degree of dynamism between arrivals and departures occurring between two sensing instants. Moreover, arrivals and departures of PUs and SUs are generated using Monte Carlo methods assuming Poisson and exponential distributions respectively. All users are considered to be located in the area of coverage of their respective networks. This simulator is somewhat more realistic than the model in a way that it is not constrained by Hypothesis 1 and 2 given in Section 3.4. Instead, the order in which PUs and SUs arrive and depart from the system is taken into account, contrarily to what Hypothesis 1 states; and a user arriving after a sensing period may depart before the next sensing period starts, which is opposite to what Hypothesis 2 states. This is particularly useful in order to assess the range of applicability of the proposed model. Simulation runs are long enough to guarantee statistical goodness and, for the considered cases herein, values in the order of 10^5 s are used.

In order to assess the validation of the proposed model we compare the steady-state probabilities obtained through the Markov chain with the computed values through simulation. Since the desired metrics (see Section 4) are directly computed from the steady-state probabilities, $P_{(i,j)}$ and $P'_{(i,j)}$, a positive validation at this point results in a good indicator about the validity of the proposed model. In this respect, Fig. 4 plots the steady-state probabilities resulting from the Markov model against the same probabilities computed by means of simulation. In addition it is also plotted the curve $y = x$ for reference purposes (note that for a perfect match, plotted data in Fig. 4 should lay over the curve $y = x$). In addition, log–log scaling enables a detailed comparison between both magnitudes. It can be observed that a good match between the model and the simulation is attained, in particular for higher probability values (i.e. in the range between 10^{-2} and 10^{-1}). For probabilities below this range, differences become more evident indicating that a larger number of samples (i.e. larger simulation times) are needed in order to obtain a good statistics of simulated data.

Figs. 5 and 6 show the average number of users and the interruption probability, as defined in (25), (26) and (30) respectively, where simulation results appear as circles and the DTMC model performance appear as lines. At a first glance, note how the system-level simulator values (i.e. circles) closely match those obtained by means of the DTMC model.

Regarding Fig. 5, when spectrum sensing is subject to errors, performance evaluation is shown for different considered values of δ and ε through specified values of m (regard Table 1). Provided that PUs have spectrum access priority over SUs, the average number of served PUs remains constant (note that we have fixed $T_p = 10$E). As for the average number of SUs, as expected, the better the spectrum sensing information (i.e. higher m values) the better spectrum opportunities can be exploited. For the case of perfect sensing the highest utilization of free bands is attained by SUs.

Fig. 6 shows the interruption probability (i.e. the probability that a SU is forced to suspend its session due to primary activity) defined in (30) for the different values of m. While the perfect-sensing case exhibits the most favorable behavior, a decrease in m results into higher false-alarm probabilities which in turn triggers SUs to release occupied channels. In addition, the interruption probability increases with the offered secondary load, which is, on the other hand, somewhat expected.

Additionally, it is necessary to determine when hypotheses 1 and 2 (provided in Section 3.4) pose some risk on the model's validity. Indeed, such hypotheses rely in one way or another on the relation of arrival and departure process parameters with the DTMC observation period ΔT. As for Hypothesis 1, the relation between $1/\mu$ and ΔT will determine if a call/session arriving within a period ΔT will depart in the same ΔT. On the other hand, the longer the period ΔT the more important becomes the order in which arrival and departures occur, thus Hypothesis 2 will be less true as ΔT increases. To that end, the behavior of the average number of users and the interruption probability against several $\Delta T \cdot \mu$ values is presented in Fig. 7(a) and (b) respectively. The offered traffic is fixed to $T_p = 5$E and $T_s = 20$E. It can be observed (see Fig. 7(b)) that the interruption probability decreases when ΔT values increase. Not

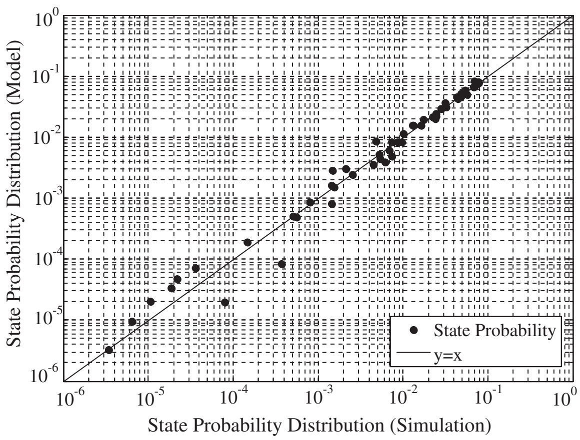


Fig. 4. Model validation of the steady-state probabilities by plotting the obtained simulation value against the theoretical (model) value. Perfect match implies data laying on $y = x$ curve.

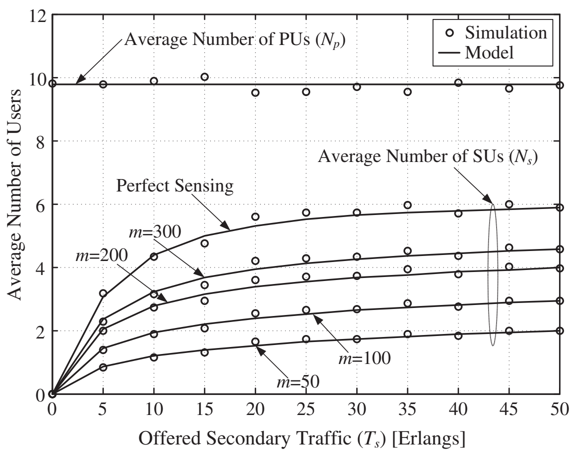


Fig. 5. Average number of users against offered secondary traffic load (T_s). The offered PU load is $T_p = 10$E.

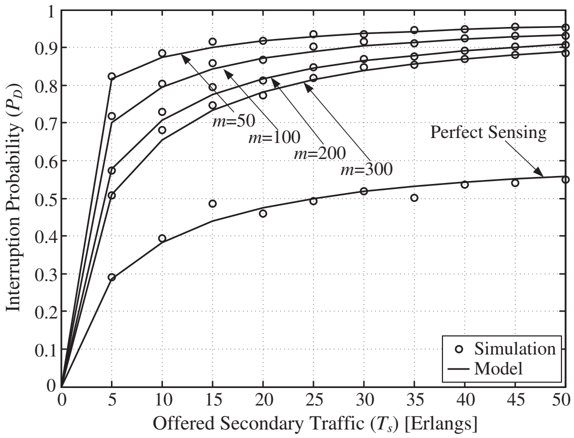


Fig. 6. Interruption probability against offered secondary traffic load (T_s). The offered PU load is $T_p = 10$E.

surprisingly, very frequent tests on the occupancy of a channel (i.e. low ΔT values) will translate into higher chances that a SU is refrained from transmitting specially for the cases that the false-alarm probability is significant (i.e. low m values). Consequently, see Fig. 7(a), the number of secondary users increases with ΔT.

Note that for $\Delta T{\cdot}\mu$ values over $2{\cdot}10^{-2}$, the model and the simulated data start to drift apart. This drift is more noticeable in the interruption probability case, see shaded region in Fig. 7(b), than for the average number of users. Nevertheless, for values of $\Delta T{\cdot}\mu$ which remain in the range 10^{-4} to 10^{-2}, the model is able to capture the behavior exhibited by the system-level simulator, thus validating the proposed hypothesis in these cases. Note, that this dependency between the model validation and ΔT was somewhat expected when expressing hypotheses 1 and 2. In terms of model applicability, practical system parameter values can be evaluated bearing in mind the relation between ΔT and μ.

5.3.2. Sensing time/throughput trade-off

Fig. 8 shows the existing trade-off between the time-bandwidth product (m) and the throughput experienced by SUs for several offered secondary traffic loads (T_s) and average SNRs ($\bar{\gamma}$). It can be seen that following an initial throughput increase due to accurate sensing information, throughput degradation starts to rise when the sensing time increases to values that are high compared to the time devoted to data transmission (i.e. the sensing efficiency $\eta_{sens} \to 0$). The extreme case is when all transmission time is devoted to sensing purposes, in this case when $T_{sens} = \Delta T$ then $\eta_{sens} = 0$ which occurs for $m = 1250$ considering $\Delta T = 0.1$ s. For the particular study case in Fig. 8, several optimum values for the time-bandwidth product can be selected in order to maximize the average secondary throughput. These values are represented by m^* in Fig. 8. Note that increasing the secondary offered load requires a better knowledge of spectrum occupancy, thus increasing the time-bandwidth product is convenient. In the same way, higher SNR values require less sensing time, i.e. lower m values, to attain acceptable accuracy in terms of missed-detection and false-alarm probabilities. In addition, a decrease in the average SNR results in a dramatic decrease of the average throughput as could be expected by regarding (41).

5.3.3. Channel number impact

Fig. 9 shows the impact of the number of total channels (C) on the experienced secondary throughput for two different values of ΔT and constant offered traffic load of $T_p = T_s = 5$E along with average SNR of $\bar{\gamma} = 10$ dB. In Fig. 9(a), for $\Delta T = 0.1$ s, note that between 2 and 10 channels, the better the sensing accuracy (i.e. higher m) the higher the number of SUs are getting assigned, thus higher

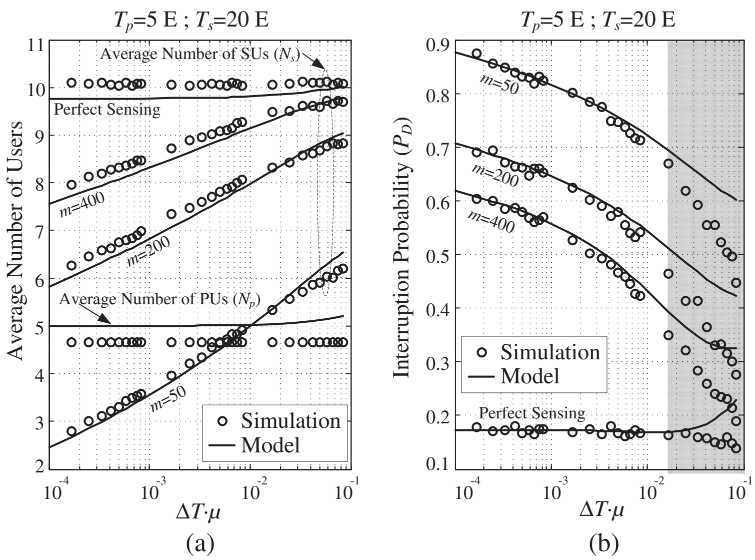


Fig. 7. The impact of $\Delta T{\cdot}\mu$ in model validation: (a) average number of users and (b) interruption probability.

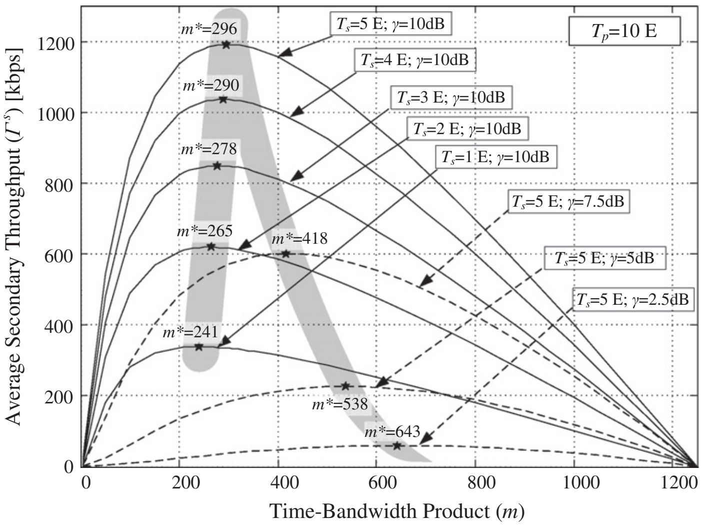


Fig. 8. Secondary throughput against the time-bandwidth product (m) for different traffic (solid) and SNR (dashed) conditions. Shaded regions emphasize optimum m values (m^*) for which secondary throughput is maximum.

throughput is attained. However, given that sensing requires a time $T_{sens} = C\cdot T$ (see (5)); as the number of channels increases, high sensing accuracy (i.e. high m which implies also high T values) does not payoff the increased time devoted to sensing and its consequences on throughput (as already observed in subsection 5.3.2)). Then, it is observed in Fig. 9(a) that at some value of C it is better to reduce the sensing accuracy, i.e. decrease m, which in turn produces higher false-alarm. The grey shadowed zones indicate regions where a suitable time-bandwidth product, among those considered in Fig. 9 and denoted as m^*, maximizes average secondary throughput. Given that the offered load is constant and more channels are available by rising C, the increased false-alarm as a result of lowering m remains bearable. On the other hand, in Fig. 9(b) for $\Delta T = 5$ s, the time devoted to sensing can be larger in this case with no observed throughput degradation. This can be followed by observing the sensing efficiency in (4), where if ΔT increases, we may tolerate higher T_{sens} values so that the sensing efficiency is still acceptable. Then, for $\Delta T = 5$ s, accurate sensing does payoff the time devoted to sensing procedures and, thus, better throughput is observed for the case of higher values of m.

5.3.4. Spectrum awareness periodicity

Fig. 10 shows the impact of spectrum occupancy information periodicity ΔT in terms of interference probability (P_c) as defined in Section 4.4. It is again assumed that $C = 16$ channels are available. Results indicate that high values of ΔT cause the secondary system to take decisions

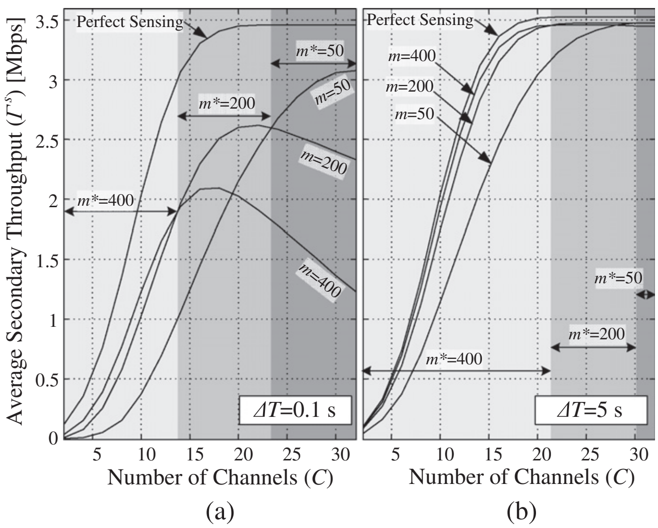


Fig. 9. Secondary throughput against number of channels (C) for (a) $\Delta T = 0.1$ s and (b) $\Delta T = 5$ s. Shadowed regions indicate the suitability of particular time-bandwidth products (m^*) such that throughput is maximized.

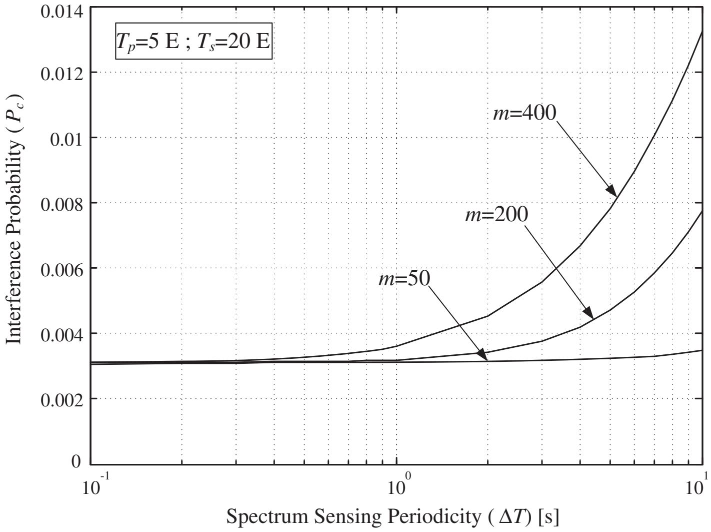


Fig. 10. Interference probability against sensing periodicity (ΔT) when T_p = 5E and T_s = 20E for several time-bandwidth products (m).

with out-of-date primary spectrum occupancy information which translates into higher interference probabilities, i.e. higher chances that a PU and a SU are assigned the same channel. In this sense, given high time-bandwidth products maximize spectrum utilization (see Fig. 5), higher chances that SUs interfere with PUs arise. On the contrary, for low m values the higher false-alarm probability prevents from assigning SUs thus less interference is observed as opposed to higher m values.

The impact of interference on primary user throughput, as revealed by (38), is given in Fig. 11(a), where primary throughput benefits from high false-alarm (i.e. low m values) since interference is lowered as already shown in Fig. 5.

As for the secondary throughput, Fig. 11(b), the sensing periodicity ΔT conditions the time devoted to sensing purposes (T_{sens}) and, hence, the time-bandwidth product. Then, for low ΔT values, large time-bandwidth products causes sensing efficiency to decrease and thus reduced throughput is attained. On the contrary, increased ΔT values allow a longer sensing period and consequently better spectrum awareness which in turn improves secondary spectrum usage and throughput. Shaded regions in Fig. 11(b) reflect the suitable time-bandwidth values (m^*), among those considered, indicating the existing trade-off between sensing efficiency and spectrum awareness quality.

5.4. Computational considerations

In general, the computational complexity of solving the matrix equation $v = v \cdot \mathbf{P}$ depends on the size of transition probability matrix $\mathbf{P}$ which in turn depends on the number of states in the Markov chain given by $N_{states} = (C + 1)^2$ (see Section 3.1). Thus, it is straightforward that increasing the number of channels C directly impacts the computational

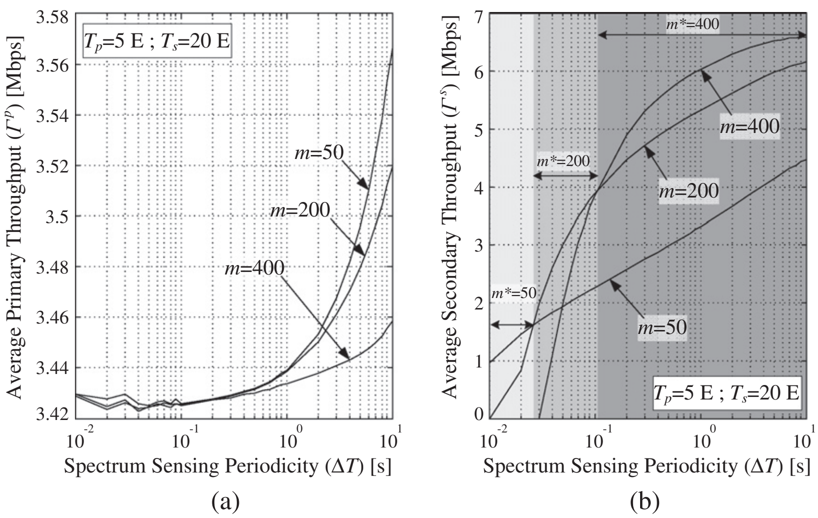


Fig. 11. Primary (a) and secondary (b) throughput against ΔT when T_p = 5E and T_s = 20E for several time-bandwidth products (m).

time required to solve $v = v{\cdot}\mathbf{P}$ (since increased vector–matrix products are required), and also the memory storage requirements involved in storing matrix **P**. In this sense, numerical iterative methods, such as Successive Over-Relaxation (SOR) which has been used in this work [28], converges towards the solution with relatively low number of iterations (typically less than 50 for 10^{-6} solution accuracy). Commercial off-the-shelf computers have been used exhibiting reasonable processing times (in the order of several minutes) and sufficient storage capabilities being capable of handling the performance evaluation considered herein. In addition, the proposed model offers reduced complexity with respect to similar approaches using Partially Observable Markov Decision Processes (POMDP), as e.g. in [29], while still being capable of capturing sensing error uncertainty in opportunistic spectrum access scenarios.

6. Concluding remarks and future work

In this work a generalized and flexible framework for the definition and evaluation of opportunistic shared spectrum scenarios has been presented. This framework is capable of supporting a wide range of implementation possibilities and functionalities. In this sense, the suitability of a DTMC model as the core of the framework has been suggested and further justified. The DTMC model has been formulated with a high degree of generality and some performance metrics extracted. An uncoordinated operation between primary and secondary networks has been assumed where primary spectrum occupancy information is retrieved through sensing mechanisms. A first goal was to determine the validity of the proposed model which has been assessed by means of its comparison with a system-level simulator. In addition, the limitations of the model have also been determined and the parameters influencing such limitation have been identified. It was shown that, for a correct model operation, the value $\Delta T{\cdot}\mu$ should be appropriately chosen. Consequently, practical system parameters values can be evaluated bearing in mind such limitations. In this case, in our particular scenario, values of $\Delta T{\cdot}\mu$ below $2{\cdot}10^{-2}$ are suggested in order to achieve a good match between the model and the simulations.

The existing tradeoff between the sensing accuracy and the exhibited secondary throughput has also been studied. As expected, an increased sensing accuracy through longer sensing periods will, at a given point, not payoff the degradation obtained in terms of throughput since less time is then devoted to the actual data transmission. Results revealed that the sensing time (equivalently, the time-bandwidth product) can be conveniently adjusted in order to maximize throughput. In our numerical analysis (refer to Fig. 8) we observed how for decreasing offered secondary traffic loads (from T_s = 5E to T_s = 1E) the sensing accuracy can be slightly decreased (from m^* = 296 to m^* = 241) in order to favor secondary throughput. In addition, an improved SNR condition allows a decrease in sensing time since energy detection is better. In this case, see Fig. 8, suitable time-bandwidth values span from m^* = 296 (for $\bar{\gamma} = 10$ dB) to m^* = 643 (for $\bar{\gamma} = 2.5$ dB). In addition, if the number of channels to be sensed is large, sensing procedures will take longer to determine the spectrum occupancy of the whole band, consequently reducing the sensing efficiency. Then, the time-bandwidth product should be reduced when increasing the bandwidth on which PUs and SUs operate. The impact of the spectrum awareness periodicity ΔT has also been evaluated. As expected, the longer the time between sensing instants (i.e. ΔT) the higher the chances of collision events between PUs and SUs happen. In addition, improved sensing accuracy degrades the experienced interference by allowing an increased number of SUs in the system. As shown, secondary operation can be optimized by choosing adequate values for the time-bandwidth product (i.e. the time devoted to sensing) in such way that the throughput is maximized.

Future work will be devoted to extend the presented model to include a flexible bandwidth partition scheme in which PUs and SUs do not necessarily occupy the same amount of spectrum. In addition, bandwidth requirements could be dynamically adjusted according to specific demands which may vary over time. Furthermore, besides the considered time-based service type of SUs in this work, a volume-based service type which intends to transmit a certain volume of data will be also considered.

Acknowledgment

This work has been supported by the Spanish Research Council under COGNOS Grant (ref. TEC2007-60985.

Appendix A. Notation

Table A.2 summarizes the notation used throughout this article.

Appendix B. Proof of Theorem 1

Given a total number of C channels we aim to compute the probability of having n detected PUs when actually i PUs are assigned. It is assumed that the sensing of each channel is independently affected by errors in the form of missed-detection and false-alarm probabilities. Let N_{FA} and N_{MD} represent the total number of false-alarms and missed-detections after scanning the C channels. The number of possible miss detections, N_{MD}, will be consequently in the range $0 \leqslant N_{MD} \leqslant i$, i.e. we cannot miss-detect more than the number of "true" assigned PUs. In the same way, the number of false-alarms is in the range $0 \leqslant N_{FA} \leqslant C - i$.

On the other hand, we can express the number of detected PUs, n, as

$$n = i + N_{FA} - N_{MD}, \tag{B.1}$$

indicating that the total number of detected PUs are those actually assigned (i) plus those we believe are assigned (N_{FA}) minus those we have missed (N_{MD}).

From (B.1), we may re-write

$$N_{FA} = N_{MD} + n - i, \tag{B.2}$$

Table A.2
Summary of used notation.

Symbol	Description
C	number of channels
δ (ε)	miss-detection (false-alarm) probability
m	time-bandwidth product
T	time devoted to sense a channel
W	channel bandwidth
ΔT	sensing periodicity
T_{sens}	time devoted to sense the whole spectrum
η_{sens}	sensing efficiency
λ_p (λ_s)	arrival rate for PUs (SUs)
μ_p (μ_s)	departure rate for PUs (SUs)
t_n	nth observation/sensing time
I_n	nth time interval between two successive observation times
$N_p(t_n)$ ($N_s(t_n)$)	number of PUs (SUs) at time t_n
$\mathbf{X}_n$	true state of the Markov Chain at time t_n
$\mathbf{Y}_n$	detected state of the Markov Chain at time t_n
$S_{(i,j)}$	state of the Markov chain with i PUs and j SUs
$\mathcal{S}$	state space
$\mathcal{S}_c$	collision state space
$\mathcal{S}_{pc}$	possible collision state space
$\mathcal{S}_{nc}$	collision-free state space
N_{states}	number of states in the state space $\mathcal{S}$
$b_{(k,i)}$	conditional PU detection probability
$N^{PA}(N^{SA})$	number of PU (SU) arrivals in I_n
$N^{PD}(N^{SD})$	number of PU (SU) departures in I_n
$P_k^{PA}\left(P_k^{SA}\right)$	prob. of k PU (SU) arrivals in ΔT
$P_k^{DA}\left(P_k^{DA}\right)$	prob. of k PU (SU) departures in ΔT
$P_{(i,j-k,l)}$	transition prob. from state $S_{(k,l)}$ to state $S_{(i,j)}$
$\mathbf{P}$	transition probability matrix
$P_{(i,j)}$	steady-state probability of being in state $S_{(i,j)}$
$P'_{(i,j)}$	detected steady-state probability of being in state $S_{(i,j)}$
$a^P_{(i,j,k,l)}(a^S_{(i,j,k,l)})$	prob. of assigning k PUs (SUs) when also l PU (SU) de-assignments in state $S_{(i,j)}$
$d^P_{(i,j,k,l)}\left(d^S_{(i,j,k,l)}\right)$	prob. of de-assigning k PUs (SUs) in state $S_{(i,j)}$
v	steady-state probability vector
$N_p(N_s)$	average number of PUs (SUs) in the system
N'_p	average number of sensed PUs
$P_B^P\left(P_B^S\right)$	blocking probability for PUs (SUs)
$T_p(T_s)$	offered PU (SU) traffic
T_s^{served}	average served SU traffic
P_D	interruption probability
$N_{MD}(i,j)$	average number of missed-detections in state $S_{(i,j)}$
$N_c(i,j)$	average number of collisions in state $S_{(i,j)}$
N_c^*	upper-bound on average number of collisions
$P_c(i,j)$	interference probability in state $S_{(i,j)}$
P_c	average interference probability
$\Gamma^p_{(i,j)}\left(\Gamma^s_{(i,j)}\right)$	PU (SU) throughput in state $S_{(i,j)}$
$R_p(R_s)$	PU (SU) net throughput per channel
$\bar{\gamma}$	average signal-to-noise ratio (SNR)
$\Gamma^p(\Gamma^s)$	average PU (SU) throughput
$\delta^*(\varepsilon^*)$	target miss-detection (false-alarm) probability
$N_{MD}(N_{FA})$	number of missed-detections (false-alarms) after sensing C channels
$N_a^P\left(N_a^S\right)$	number of spectrum assignments for PUs (SUs)
$N_d^P\left(N_d^S\right)$	number of spectrum de-assignments for PUs (SUs)

where if $N_{FA} \geqslant 0$ it follows that $N_{MD} \geqslant i-n$. Then, given that also $N_{MD} \geqslant 0$, we have that $N_{MD} \geqslant \max(0, i-n)$. On the other hand, from $0 \leqslant N_{FA} \leqslant C-i$ and (B.2), we have that $0 \leqslant N_{MD}+n-i \leqslant C-i$, thus $N_{MD} \leqslant C-n$. Provided $0 \leqslant N_{MD} \leqslant i$ must be also satisfied we finally obtain $N_{MD} \leqslant \min(i, C-n)$.

The probability of detecting n PUs given i PUs are assigned can be expressed using the total probability definition for conditional probabilities as

$$b_{(n,i)} = \sum_{N_{MD}} \Pr(N_{FA} = N_{MD}+n-i|N_{MD})\Pr(N_{MD}), \tag{B.3}$$

where, by considering binomial distributions for both false-alarm and miss-detection, we have

$$\begin{aligned}\Pr(N_{FA} &= N_{MD}+n-i|N_{MD}) \\ &= \binom{C-i}{N_{MD}+n-i}\varepsilon^{N_{MD}}+n-i(1-\varepsilon)^{C-i-N_{MD}}-n+i, \\ &= \binom{C-i}{N_{MD}+n-i}\varepsilon^{N_{MD}}+n-i(1-\varepsilon)^{C-N_{MD}}-n\end{aligned} \tag{B.4}$$

along with

$$\Pr(N_{MD}) = \binom{i}{N_{MD}}\delta^{N_{MD}}(1-\delta)^{i-N_{MD}}, \tag{B.5}$$

from which expression (11) can be obtained.

Appendix C. Proof of Lemmas 1–4

The number of PU/SU spectrum assignments and de-assignments in I_n, $N_a \in \left\{N_a^P, N_a^S\right\}$ and $N_d \in \left\{N_d^P, N_d^S\right\}$, will depend on the spectrum occupancy given by the true or detected states at time t_n, i.e. $\mathbf{X}_n$ or $\mathbf{Y}_n$, and on the number of arrivals N^A and departures N^D in time interval I_n. These number of arrivals and departures will eventually lead to a number of channel assignments and de-assignments depending on the true or detected spectrum occupancy at time t_n. Fig. C.12 illustrates the arrival and departure process which conditions the state transition probabilities.

In the following, expressions for Lemmas 1–4 are derived.

C.1. Proof of Lemma 1

Let the true state be $\mathbf{X}_n = S_{(i,j)}$; we intend to compute the probability of assigning $N_a^P = k$ PUs in I_n given we have $N_d^P = l \leqslant i$ PU de-assignments in I_n, $a^P_{(i,j,k,l)}$. In words, it is the probability of assigning exactly k PUs in the case that least k channels are available (i.e. having exactly k PU arrivals), and the probability of having at least k PU arrivals if exactly k channels are available. Thus, we may write:

$$\begin{aligned}a^P_{(i,j,k,l)} &= \Pr\left[N_a^P = k|\mathbf{X}_n = S_{(i,j)}, N_d^P = l\right] \\ &= \begin{cases}\Pr[N^{PA} = k] = P_k^{PA}, & \text{if } i-l+k < C \\ \Pr[N^{PA} \geqslant k] = 1 - \sum_{m=0}^{k-1} P_m^{PA}, & \text{if } i-l+k = C\end{cases},\end{aligned} \tag{C.1}$$

with P_k^{PA} given in (12). For the case of assigning more PUs than available channels the probability in (C.1) is zero. Implicitly in (C.1) we consider that k PU assignments are made upon l PU de-assignments, thus using the assumption of disregarding the order in which arrivals

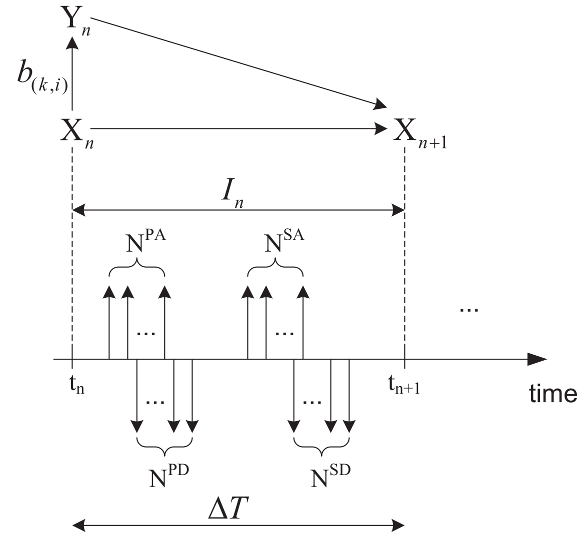


Fig. C.12. Spectrum assignments/de-assignments due to arrival/departure of PUs and SUs.

and departures occur in I_n which will be an assumed hypothesis in our model (see Hypothesis 2 in Section 3.4).

C.2. Proof of Lemma 2

Again, let the true state be $\mathbf{X}_n = S_{(i,j)}$; the probability of de-assigning $N_d^P = k$ PUs in I_n, where $0 \leqslant k \leqslant i$ (i.e. we can only de-assign those already assigned prior to t_n), depends on the number of PU departures in I_n:

$$d^P_{(i,j,k)} = Pr\left[N_d^P = k|\mathbf{X}_n = S_{(i,j)}\right] = \Pr[N^{PD} = k] = P_k^{PD}, \qquad (C.2)$$

with P_k^{PD} given in (14). Note that we have made use of the assumption that a new arrival in I_n will not depart in I_n by specifying that the number of de-assignments is bounded as $0 \leqslant k \leqslant i$ in I_n. For any other value of k, the probability in (C.2) is zero.

C.3. Proof of Lemma 3

Let the true state be $\mathbf{X}_n = S_{(i,j)}$; the probability of assigning $N_a^S = k > 0$ SUs in I_n given we have $N_d^S = l \leqslant j$ SU de-assignments in I_n, $a^S_{(i,j,k,l)}$, will depend on the detected state at t_n, $\mathbf{Y}_n = S_{(m,j)}$ and on the number of SU arrivals as:

$$\begin{aligned} a^S_{(i,j,k,l)} &= \Pr\left[N_a^S = k|\mathbf{X}_n = S_{(i,j)}, N_d^S = l\right] \\ &= \sum_{m=0}^{C-k-j+l} \Pr\left[N_a^S = k|\mathbf{Y}_n = S_{(m,j)}, N_d^S = l\right] \cdot b_{(m,i)} \\ &= \sum_{m=0}^{C-k-j+l} \bar{a}^S_{(m,j,k,l)} \cdot b_{(m,i)}, \end{aligned} \qquad (C.3)$$

where the total probability formula has been used to relate the true state $\mathbf{X}_n = S_{(i,j)}$with the detected state $\mathbf{Y}_n = S_{(m,j)}$through probabilities $b_{(k,i)}$. In particular, (C.3) states that $N_a^S = k$ SUs will be assigned provided the detected number of PUs, m, fulfills $m + j + k - l \leqslant C$, i.e., there are at least k detected free channels for secondary use provided that we also have l SU de-assignments. In addition, the number of k SU assignments in state $\mathbf{X}_n = S_{(i,j)}$is bounded by $0 < k \leqslant C - i - j + l$, omitting the case $k = 0$ which will be treated separately. Finally, $\bar{a}^S_{(m,j,k,l)}$ in (C.3) is obtained similar to (C.1) as:

$$\bar{a}^S_{(m,j,k,l)} = \begin{cases} P_k^{SA}, & \text{if } m + j - l + k < C \\ 1 - \sum_{r=0}^{k-1} P_r^{SA}, & \text{if } m + j - l + k = C \end{cases}. \qquad (C.4)$$

For the specific case of no SU assignments (i.e. $k = 0$), the probability of assigning $k = 0$ SUs is the probability of assigning $k = 0$ SUs when there is at least one free detected channel or the probability that there are no detected free channels, i.e.:

$$\begin{aligned} a^S_{(i,j,0,l)} &= \Pr\left[N_a^S = 0|\mathbf{X}_n = S_{(i,j)}, N_d^S = l\right] \\ &= \sum_{m=0}^{C-j+l} \bar{a}^S_{(m,j,0,l)} \cdot b_{(m,i)} + \sum_{m=C-j+l}^{C} b_{(m,i)}. \end{aligned} \qquad (C.5)$$

Combining (C.3) with (C.5), along with the definition in (C.4), expression (18) is obtained.

C.4. Proof of Lemma 4

As for the de-assignment processes of SUs, there are mainly two independent events which imply an SU de-assignment: in the first place, a number of $N_d^{S,S}$ SUs may be de-assigned provided detection at time t_n determines that there are $N_d^{S,S}$ SUs sharing the same channel with PUs. Secondly, a number of $N_d^{S,SC}$ SUs may be de-assigned in I_n simply because their sessions have ended (here, SC stands for Service Completion).

Then, let the true state be $\mathbf{X}_n = S_{(i,j)}$; the probability of de-assigning $N_d^{S,S} = k$ SUs in I_n due to detection of state $\mathbf{Y}_n = S_{(m,j)}$ at time t_n, given the number of de-assignments due to service completion is $N_d^{S,SC} = l$, is given by:

$$\begin{aligned} \Pr[N_d^{S,S} = k|\mathbf{X}_n = S_{(i,j)}, N_d^{S,SC} = l] &= \Pr[m + j - l = C + k] \\ &= b_{(C+k-j+l,i)}, \end{aligned} \qquad (C.6)$$

provided that $0 < k \leqslant j - l$. Accordingly, the probability of no SU de-assignments due to detection of state $\mathbf{Y}_n = S_{(m,j)}$ is:

$$\Pr\left[N_d^{S,S} = 0|\mathbf{X}_n = S_{(i,j)}, N_d^{S,SC} = l\right] = 1 - \sum_{k=1}^{j-l} b_{(C+k-j+l,i)}. \qquad (C.7)$$

Then, from (C.6) and (C.7), we can write:

$$\begin{aligned} d^{S,S}_{(i,j,k,l)} &= \Pr\left[N_d^{S,S} = k|\mathbf{X}_n = S_{(i,j)}, N_d^{S,SC} = l\right] \\ &= \begin{cases} b_{(C+k-j+l,i)} & \text{if } 0 < k \leqslant j - l \\ 1 - \sum_{r=1}^{j-l} b_{(C+r-j+l,i)} & \text{if } k = 0 \end{cases}. \end{aligned} \qquad (C.8)$$

On the other hand, the probability of de-assigning k SUs in I_n due service completions is given by (similar to (C.2)):

$$d^{S,SC}_{(i,j,k)} = \Pr\left[N_d^{S,SC} = k|\mathbf{X}_n = S_{(i,j)}\right] = \Pr[N^{SD} = k] = P_k^{SD}. \qquad (C.9)$$

Finally, we can express the global probability of de-assigning k SUs in I_n (i.e. without specifying if the de-assignment is due to detection or due to session completion) as:

$$d^S_{(i,j,k)} = \Pr\left[N^S_d = k|\mathbf{X}_n = S_{(i,j)}\right] = \sum_{r=0}^{k} d^{S,S}_{(i,j,k-r,r)} \cdot d^{S,SC}_{(i,j,r)}. \quad \text{(C.10)}$$

The above expressions proof Lemma 4.

Appendix D. Proof of Theorem 2

In this appendix, a detailed derivation of all possible transition probabilities in the DTMC model is presented.

D.1. Transition $S_{(i,j)} \to S_{(i+N,j)}$ with $0 < N \leqslant C - i$

The probability associated to transition $S_{(i,j)} \to S_{(i+N,j)}$, with $0 < N \leqslant C - i$, is the probability to have N more PU assignments in than PU de-assignments in I_n; and equal number of SU assignments and de-assignments in I_n. Then, applying conditional independence between PU and SU processes, we can write:

$$P_{(i+N,j|i,j)} = \Pr\left[N^P_a - N^P_d = N|i,j\right] \cdot \Pr\left[N^S_a = N^S_d|i,j\right], \quad \text{(D.1)}$$

where multiplicative probabilities in (D.1) are given by

$$\begin{aligned}\Pr\left[N^P_a - N^P_d = N\right] &= \Pr\left[N^P_a = N|N^P_d = 0\right] \cdot \Pr[N^P_d = 0]\\ &+ \Pr\left[N^P_a = N+1|N^P_d = 1\right] \cdot \Pr\left[N^P_d = 1\right]\\ &+ \Pr\left[N^P_a = N+2|N^P_d = 2\right] \cdot \Pr\left[N^P_d = 2\right]\\ &+ \cdots + \Pr\left[N^P_a = N+i|N^P_d = i\right] \cdot \Pr\left[N^P_d = i\right]\\ &= a^P_{(i,j,N,0)} \cdot d^P_{(i,j,0)} + a^P_{(i,j,N+1,1)} \cdot d^P_{(i,j,1)}\\ &+ a^P_{(i,j,N+2,2)} \cdot d^P_{(i,j,2)} + \cdots + a^P_{(i,j,N+i,i)} \cdot d^P_{(i,j,i)}\\ &= \left(\sum_{k=0}^{i} a^P_{(i,j,N+k,k)} \cdot d^P_{(i,j,k)}\right), \end{aligned} \quad \text{(D.2)}$$

along with

$$\begin{aligned}\Pr\left[N^S_a = N^S_d\right] &= \Pr\left[N^S_a = 0|N^S_d = 0\right] \cdot \Pr\left[N^S_d = 0\right]\\ &+ \Pr\left[N^S_a = 1|N^S_d = 1\right] \cdot \Pr\left[N^S_d = 1\right]\\ &+ \Pr\left[N^S_a = 2|N^S_d = 2\right] \cdot \Pr\left[N^S_d = 2\right]\\ &+ \cdots + \Pr\left[N^S_a = j|N^S_d = j\right]\\ &\cdot \Pr\left[N^S_d = j\right]\\ &= a^S_{(i,j,0,0)} \cdot d^S_{(i,j,0)} + a^S_{(i,j,1,1)} \cdot d^S_{(i,j,1)}\\ &+ a^S_{(i,j,2,2)} \cdot d^S_{(i,j,2)} + \ldots + a^S_{(i,j,j,j)} \cdot d^S_{(i,j,j)}\\ &= \left(\sum_{k=0}^{j} a^S_{(i,j,k,k)} \cdot d^S_{(i,j,k)}\right). \end{aligned} \quad \text{(D.3)}$$

Note that, for the sake of notation relief, we have intentionally omitted the conditioning to i and j stated in (D.1).

Then, by replacing (D.2) and (D.3) in (D.1), we may write

$$P_{(i+N,j|i,j)} = \left(\sum_{k=0}^{i} a^P_{(i,j,N+k,k)} \cdot d^P_{(i,j,k)}\right) \cdot \left(\sum_{k=0}^{j} a^S_{(i,j,k,k)} \cdot d^S_{(i,j,k)}\right). \quad \text{(D.4)}$$

D.2. Transition $S_{(i,j)} \to S_{(i-N,j)}$ with $0 < N \leqslant i$

The probability associated to transition $S_{(i,j)} \to S_{(i-N,j)}$, with $0 < N \leqslant i$, is the probability to have N more PU de-assignments than PU assignments in I_n; and equal number of SU assignments and de-assignments in I_n. Then, we have

$$P_{(i-N,j|i,j)} = \Pr\left[N^P_d - N^P_a = N|i,j\right] \cdot \Pr\left[N^S_a = N^S_d|i,j\right], \quad \text{(D.5)}$$

where the first multiplicative term in (D.5) is given by[2]

$$\begin{aligned}\Pr\left[N^P_d - N^P_a = N\right] &= \Pr\left[N^P_d = N\right] \cdot \Pr\left[N^P_a = 0|N^P_d = N\right]\\ &+ \Pr\left[N^P_d = N+1\right] \cdot \Pr\left[N^P_a = 1|N^P_d = N+1\right]\\ &+ \Pr\left[N^P_d = N+2\right] \cdot \Pr\left[N^P_a = 2|N^P_d = N+2\right]\\ &+ \cdots + \Pr\left[N^P_d = i\right] \cdot \Pr\left[N^P_a = i-N|N^P_d = i\right]\\ &= d^P_{(i,j,N)} \cdot a^P_{(i,j,0,N)} + d^P_{(i,j,N+1)} \cdot a^P_{(i,j,1,N+1)}\\ &+ d^P_{(i,j,N+2)} \cdot a^P_{(i,j,2,N+2)} + \cdots + d^P_{(i,j,i)} \cdot a^P_{(i,j,i-N,i)}\\ &= \left(\sum_{k=0}^{i-N} d^P_{(i,j,N+k)} \cdot a^P_{(i,j,k,N+k)}\right), \end{aligned} \quad \text{(D.6)}$$

and, with $\Pr\left[N^S_a = N^S_d|i,j\right]$ given in (D.3), we have

$$P_{(i-N,j|i,j)} = \left(\sum_{k=0}^{i-N} d^P_{(i,j,N+k)} \cdot a^P_{(i,j,k,N+k)}\right) \cdot \left(\sum_{k=0}^{j} a^S_{(i,j,k,k)} \cdot d^S_{(i,j,k)}\right). \quad \text{(D.7)}$$

D.3. Transition $S_{(i,j)} \to S_{(i+N,j)}$ with $-i < N \leqslant C - i$

In subsections D.1 and D.1 we have separately presented transitions involving an increase and decrease of (positive) PUs respectively. We may generalize the case of transition $S_{(i,j)} \to S_{(i+N,j)}$ when N can be either positive or negative for $-i \leqslant N \leqslant C - i$.

From (D.6), we have $P_{(i-N-i)}$, i.e. the transition probability from having i PUs to having $i - N$ PUs with $0 < N \leqslant i$. Note that this is equivalent to consider the probability $P_{(i+N-i)}$ with $-i < N \leqslant 0$, i.e. negative N. Then we can rewrite (D.6) as:

$$P_{(i+N|i)} = \left(\sum_{k=0}^{i+N} d^P_{(i,j,k-N)} \cdot a^P_{(i,j,k,k-N)}\right), \quad \text{(D.8)}$$

for $-i < N \leqslant 0$.

With the change of variable $t = k - N$ in (D.8), we have:

$$P_{(i+N|i)} = \left(\sum_{t=-N}^{i} d^P_{(i,j,t)} \cdot a^P_{(i,j,t+N,t)}\right), \quad \text{(D.9)}$$

for $-i \leqslant N \leqslant 0$, which resembles transition probability $P_{(i+N-i)}$ for $0 < N \leqslant C - i$ given in (D.2), except for the summation index starting value. Nevertheless, we can state the general expression for the transition probability $S_{(i,j)} \to S_{(i+N,j)}$ for $-i \leqslant N \leqslant C - i$ as:

[2] Where conditioning to i and j has been intentionally dropped.

$$P_{(i+N,j|i,j)} = \left(\sum_{k=\max(-N,0)}^{i} a^{P}_{(i,j,N+k,k)} \cdot d^{P}_{(i,j,k)} \right) \cdot \left(\sum_{k=0}^{j} a^{S}_{(i,j,k,k)} \cdot d^{S}_{(i,j,k)} \right). \tag{D.10}$$

For the case of transition probabilities due to the assignment or de-assignment of SUs, an analogous expression to (D.10) can be derived, yielding, for the transition $S_{(i,j)} \to S_{(i,j+M)}$ with $-j \leqslant M \leqslant C - j - i$:

$$P_{(i,j+M|i,j)} = \left(\sum_{k=\max(-M,0)}^{j} a^{S}_{(i,j,M+k,k)} \cdot d^{S}_{(i,j,k)} \right) \cdot \left(\sum_{k=0}^{i} a^{P}_{(i,j,k,k)} \cdot d^{P}_{(i,j,k)} \right). \tag{D.11}$$

Finally, from (D.10) and (D.11), expression (23) is obtained.

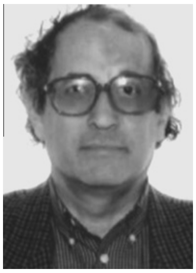
References

[1] M. Cave et al., Essentials of Modern Spectrum Management, Cambridge University Press, 2007.
[2] I.F. Akyildiz, W.Y. Lee, M.C. Vuran, S. Mohanty, Next generation/dynamic spectrum access/cognitive radio wireless networks: a survey, Computer Networks 50 (13) (2006) 2127–2159, doi:10.1016/j.comnet.2006.05.001.
[3] Q. Zhao, B.M. Sadler, A survey of dynamic spectrum access, IEEE Signal Processing Magazine 24 (3) (2007) 79–89, doi:10.1109/MSP.2007.361604.
[4] M. Sherman et al., IEEE standards supporting cognitive radio and networks, dynamic spectrum access, and coexistence, IEEE Communications Magazine 46 (7) (2008) 72–79, doi:10.1109/MCOM.2008.4557045.
[5] Federal Communications Commission (FCC), Notice of Proposed Rule Making, ET Docket no. 04-113, May 25, 2004.
[6] C. Cordeiro, K. Challapali, D. Birru, N. Sai Shankar, IEEE 802.22: the first worldwide wireless standard based on cognitive radios, in: First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN 2005, 2005, pp. 328–337, doi:10.1109/DYSPAN.2005.1542649.
[7] Standards Coordinating Committee (SCC) 41. Available from:<http://www.scc41.org/>.
[8] J.M. Peha, Emerging technology and spectrum policy reform, in: International Telecommunications Union (ITU) Workshop on Market Mechanisms for Spectrum Management, 2007.
[9] M.M. Buddhikot et al., DIMSUMnet: new directions in wireless networking using coordinated dynamic spectrum, in: Sixth IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks, WoWMoM 2005, 2005, pp. 78–85, doi:10.1109/WOWMOM.2005.36.
[10] J. Zhao, H. Zheng, G.-H. Yang, Distributed coordination in dynamic spectrum allocation networks, in: First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN 2005, 2005, pp. 259–268, doi:10.1109/DYSPAN.2005.1542642.
[11] D. Raychaudhuri, X. Jing, A spectrum etiquette protocol for efficient coordination of radio devices in unlicensed bands, in: 14th IEEE Proceedings on Personal, Indoor and Mobile Radio Communications, PIMRC 2003, vol. 1, 2004, pp. 172–176, doi:10.1109/PIMRC.2003.1264255.
[12] D. Bourse et al., The E2R II Flexible Spectrum Management (FSM) – Technical, Business and Regulatory Perspectives, E2R II White Paper, July 2007.
[13] P. Cordier et al., E2R cognitive pilot channel concept, in: 15th IST Mobile and Wireless Communications Summit, 2006.
[14] J. Pérez-Romero, O. Sallent, R. Agustí, L. Giupponi, A novel on-demand cognitive pilot channel enabling dynamic spectrum allocation, in: Second IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN 2007, 2007, pp. 46–54, doi:10.1109/DYSPAN.2007.14.
[15] X. Zhu, L. Shen, T.S.P. Yum, Analysis of cognitive radio spectrum access with optimal channel reservation, IEEE Communications Letters 11 (4) (2007) 304–306, doi:10.1109/LCOM.2007.348282.
[16] M. Raspopovic, C. Thompson, Finite population model for performance evaluation between narrowband and wideband users in the shared radio spectrum, in: Second IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN 2007, 2007, pp. 340–346, doi:10.1109/DYSPAN.2007.52.
[17] S. Tang, B.L. Mark, An analytical performance model of opportunistic spectrum access in a military environment, in: IEEE Wireless Communications and Networking Conference, WCNC 2008, 2008, pp. 2681–2686, doi:10.1109/WCNC.2008.470.
[18] S. Tang, B.L. Mark, Modeling and analysis of opportunistic spectrum sharing with unreliable spectrum sensing, IEEE Transactions on Wireless Communications 8 (4) (2008) 1934–1943, doi:10.1109/T-WC.2008.080213.
[19] D.T.C. Wong, A.T. Hoang, Y.-C. Liang, F.P.S. Chin, Dynamic spectrum access with imperfect sensing in open spectrum wireless networks, in: IEEE Wireless Communications and Networking Conference, WCNC 2008, 2008, pp. 2765–2770, doi:10.1109/WCNC.2008.484.
[20] Y. Xing, R. Chandramouli, S. Mangold, S. Sankar, Dynamic spectrum access in open spectrum wireless networks, IEEE Journal on Selected Areas in Communications 24 (3) (2006) 626–637, doi:10.1109/JSAC.2005.862415.
[21] B. Wang, Z. Ji, K.J.R. Liu, Primary-prioritized markov approach for dynamic spectrum access, in: Second IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN 2007, 2007, pp. 507–515, doi:10.1109/DYSPAN.2007.73.
[22] W. Ahmed, J. Gao, H.A. Suraweera, M. Faulkner, Comments on analysis of cognitive radio spectrum access with optimal channel reservation, IEEE Transactions on Wireless Communications 8 (9) (2009) 4488–4491, doi:10.1109/TWC.2009.090446.
[23] D. Cabric, A. Tkachenko, R.W. Brodersen, Spectrum sensing measurements of pilot, energy, and collaborative detection, in: IEEE Military Communications Conference, MILCOM 2006, 2007, pp. 1–7, doi:10.1109/MILCOM.2006.301994.
[24] A. Ghasemi, E.S. Sousa, Collaborative spectrum sensing for opportunistic access in fading environments, in: First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN 2005, 2005, pp. 131–136, doi:10.1109/DYSPAN.2005.1542627.
[25] F.F. Digham, M.S. Alouini, M.K. Simon, On the energy detection of unknown signals over fading channels, in: IEEE International Conference on Communications, ICC '03, vol. 5, 2003, pp. 3575–3579, doi:10.1109/ICC.2003.1204119.
[26] Y. Pei, A.T. Hoang, Y.-C. Liang, Sensing-throughput tradeoff in cognitive radio networks: How frequently should spectrum sensing be carried out?, in: IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2007, 2007, pp. 1–5, doi:10.1109/PIMRC.2007.4394632.
[27] D. Bertsekas, R. Gallager, Data Networks, second ed., Prentice Hall, 1992.
[28] W.J. Stewart, Introduction to the Numerical Solution of Markov Chains, Princeton University Press, 1994.
[29] Q. Zhao, L. Tong, A. Swami, Y. Chen, Decentralized cognitive mac for opportunistic spectrum access in ad hoc networks: a POMDP framework, IEEE Journal on Selected Areas in Communications 25 (3) (2007) 589–600, doi:10.1109/JSAC.2007.070409.

Xavier Gelabert (Alaior, 1978) received the Telecommunications Engineering degree (equivalent to B.S. plus M.S.) from the Universitat Politécnica de Catalunya (UPC), Barcelona, in 2004. He also holds an M.S. degree in electrical engineering, with a major in wireless communications, from the Royal Institute of Technology (KTH), Stockholm, 2003. In 2004, he joined the Radio Communication Research Group in the Department of Signal Theory and Communications, UPC, where he is pursuing his PhD. From August to December 2008 he was a visiting researcher at the Broadband Wireless Networking Laboratory (BWN-Lab) at Georgia Institute of Technology. From January to August 2009, he was a visiting researcher at the Instituto de Telecomunicaciones y Aplicaciones Multimedia (iTEAM), Universidad Politécnica de Valencia (UPV). His current research interests are in the field of mobile radio communication systems, with a special emphasis on Common Radio Resource Management (CRRM) strategies in multi-access networks, quality-of-service provisioning, and opportunistic/cognitive spectrum management. He has been actively involved in European-funded projects EVEREST, AROMA, and E3 along with Spanish projects COSMOS and COGNOS. He is a member of the IEEE.

Oriol Sallent is Associate Professor at the Universitat Politécnica de Catalunya (UPC). His research interests are in the field of radio resource and spectrum management for heterogeneous cognitive wireless networks, where he has published 100 + papers in IEEE journals and conferences. He has participated in many research projects and consultancies funded by either public organizations or private companies.

Jordi Pérez-Romero received the Telecommunications Engineering and Ph.D. degrees from the Universitat Politécnica de Catalunya (UPC), Barcelona, Spain, in 1997 and 2001, respectively. He is currently an Associate Professor with the Department of Signal Theory and Communications, UPC. He has been involved in several different European projects as well as projects for private companies. He has authored papers in international journals and conference proceedings and has coauthored one book on mobile communications. His research interests are mobile communication systems, particularly packet radio techniques, radio resource and quality-of-service management, heterogeneous wireless networks, and cognitive networks.



Ramon Agustí received the Engineer of Telecommunications degree from the Universidad Politécnica de Madrid, Spain, in 1973, and the Ph.D. degree from the Universitat Politécnica de Catalunya (UPC), Spain, 1978. In 1973 he joined the Escola Técnica Superior d'Enginyers de Telecomunicació de Barcelona, Spain, where he became Full Professor in 1987. After graduation he was working in the field of digital communications with particular emphasis on transmission and development aspects in fixed digital radio, both radio relay and mobile communications. For the last 15 years he has been mainly concerned with the performance analysis, development of planning tools and equipment for mobile communication systems and he has published about 200 papers in that areas. He participated in the European program COST 231 and in the COST 259 as Spanish representative delegate. He has also participated in the RACE, ACTS, IST European research programs as well as in many private and public funded projects. He received the Catalonia Engineer of the year prize in 1998 and the Narcís Monturiol Medal issued by the Government of Catalonia in 2002 for his research contributions to the mobile communications field. He is part of the editorial board of several Scientific International Journals and since 1995 is conducting a post graduate annual course on mobile communications. He co-authored two books on Mobile communications.