

Cognitive Radio Networks with Elastic Traffic

Jorge Martinez-Bauset, Alexandru Popescu, Vicent Pla and Adrian Popescu

Abstract—We model and evaluate an infrastructure based cognitive radio network where impatient unlicensed or secondary users (SUs) offer elastic traffic. The interference created by SUs to the licensed users is analyzed when the secondary network uses two different spectrum access schemes, the conventional random access scheme and a new scheme we refer to as access with preference. To further control the interference, a limit is set to the number of channels SUs have access to. With this new constraint, the abandonment probability, the mean delay and throughput of the SUs degrades significantly. To improve the QoS perceived by the active SUs, we define and evaluate an admission control scheme for the SUs.

Index Terms—Cognitive radio networks, spectrum access, elastic traffic, performance evaluation.

I. INTRODUCTION

Cognitive radio networks (CRNs) has attracted a considerable attention from the research community in recent years as an enabling technology to solve the spectrum scarcity problem [1]. This problem is the result of, or is exacerbated by, the long-running static spectrum allocation policies, which are based on assigning spectrum bands to license holders on a long-term basis for large geographical regions [2].

The aim of our study is to evaluate the performance of a CRN in a scenario where there is either a central or distributed control entity to perform channel allocations to SUs, and where SUs offer elastic traffic. We approach the study from the traffic perspective, as we believe that the traffic management techniques complement those defined at the physical layer. From this perspective, two traffic types are commonly considered: streaming or real-time traffic and elastic traffic.

Streaming traffic is generated by applications that require a minimum number of channels to be allocated during the session life time. In addition, the number of channels allocated to a streaming session does not affect its duration. On the other hand, elastic traffic adapts its transfer rate according to the available resources, and as a consequence, the duration of an elastic flow transfer depends on the flow size (in bits) and the system load, which is unpredictable. We use the term *flow* to describe a sequence of packets corresponding, for example, to a TCP connection established for the transfer of one element of a Web document or an entire page.

Our study is motivated by the observation that when SUs offer streaming traffic and a certain degree of QoS must be guaranteed to them, measured in terms of their forced termination probability, then some sort of resource reservation must be performed. When *spectrum handover* is supported, a SU vacated by the arrival of a licensed or primary user (PU) will be forced to terminate unless enough free channels are available. To control the forced termination probability of

SUs, the set up of new SU sessions can be rejected by the control entity to leave some free channels for vacated SUs, which is a form of channel reservation. However, channel reservation reduces the traffic carried by SUs [3]. When the SUs offer elastic traffic, and no minimum rate requirements are imposed, in principle, no forced terminations occur. This happens because the SUs can always adapt their transfer rate when new PUs are accepted in the system. Our interest is to evaluate how efficiently can the SUs exploit the resources not used by the PUs in this new scenario.

Scenarios where both the PUs and the SUs offer streaming traffic have been extensively studied, see for example [4], [3]. However, only a small number of studies have evaluated the performance of traffic schemes for CRNs with elastic traffic at the flow level. The study in [5] evaluates a scenario where a finite population of SUs offer elastic traffic. However, it does not consider neither the interference nor the impatience of SUs. It focuses on an approximation to determine the stationary distribution based on the idea of time-scale decomposition. In [6], the authors approach the problem from a learning-based perspective, where a number of agents (coalitions of SUs) that offer elastic traffic contend for a set of channels in a time slotted system without PUs traffic. A minimum QoS, defined in terms of a minimum transfer rate, is expected by the agents. When it falls below the minimum the agent reward decreases exponentially, which might be a form of modelling the impact of impatient SUs. However, both the objective and the performance parameters of the study are different from ours. The scenario studied in [7] considers SUs traffic that is not fully elastic, as it can only adapt between two possible transfer rates, otherwise it is blocked. In addition, SUs impatience is not considered. The study in [8] approaches the problem of spectrum allocation from an economic perspective, where the transmission rate of SUs is a function of the available resources and the price per bandwidth unit. A class of spectrum pricing functions is considered which leads to an efficient spectrum allocation.

The main features of our study can be summarized in the following three points. First, we consider a realistic model where SUs may become impatient. When the total traffic demand of elastic flows exceeds the available capacity, a situation we call SUs overload, some flows might be aborted due to impatience. Flow impatience due to a very low throughput may arise due to human impatience or because TCP or higher layer protocols interpret that the connection is broken. This phenomenon has a negative impact on the efficiency because capacity is wasted by non completed flows. This drop of efficiency led the authors of [9] to suggest that admission control (AC) should be enforced for elastic traffic.

Second, we study the interference caused by SUs to PUs. To measure the interference we determine the so-called *interference probability*, i.e. the fraction of PU arrivals that vacate a SU. This metric has been used, for instance, in [10], [3]. To reduce the interference, we propose a simple spectrum access scheme and compare it with the conventional random access scheme.

Third, to further control the interference we evaluate the impact of limiting the number of channels that the SUs have access to. While this is an effective way to control the interference, it causes the abandonments to increase considerably. To limit abandonments and improve the QoS perceived by the active SUs, an admission control scheme for SUs is incorporated and its effectiveness is evaluated.

We develop a continuous-time Markov chain (CTMC) model that describes the system behavior and evaluate common performance parameters. We also develop a simulation model that mimics the physical behavior of the system and therefore it is completely independent from the CTMC model.

The rest of the paper is organized as follows. Section II presents the model used in our study. It also contains the definition of the spectrum access schemes and the traffic management functions to control the interference and the abandonments. The numerical evaluation of different system scenarios is reported in Section III. Finally, the conclusions are presented in Section IV.

II. MODEL DESCRIPTION

We model SUs traffic at the flow level and ignore interactions at the packet level (scheduling, buffer management, etc.). The flow content is then viewed as a fluid that is transmitted as a continuous stream with rate changes occurring only at PU or SU arrivals and departures. We assume that flows equally share the available bandwidth and therefore we ignore the imprecisions that might be introduced by packet level protocols (e.g., slow start and congestion avoidance algorithms of TCP). This is a widely used traffic model, see for instance [11], [9] and their references. In part, the interest of this model is due to the fact that the QoS perceived by users of elastic applications is related to transfer time of complete files (documents), i.e. the flow duration, rather than to the packet level performance parameters like end-to-end delay.

Note that with this traffic model we are assuming an ideal MAC layer for SUs, that achieves a perfect share of the allocated channels among the active SUs (all active SUs get the same bandwidth portion), introduces zero delay and which control mechanisms consume zero resources. PUs have preemptive access priority over SUs, and SUs support spectrum handover functions. The model we consider does not capture the traffic losses and the delay suffered by the SUs due to the execution of spectrum handover, which occur during interference episodes. In other words, the SUs vacate channels instantaneously upon the arrival of interfering PUs.

We consider a CRN where there is either a central or distributed control entity to perform channel allocations to SUs. Systems with a control entity for the SUs have been

identified as potential deployment scenarios for CRNs [2]. The existence of a control entity for the SUs has been assumed before, see for example [5], [4], [12], [3].

The system has C channels and all of them can be accessed by the SUs. In our model we avoid starvation of SU flows by assigning $(C - 1)$ channels to the PUs. Note that this is just an assumption and not a limitation of the model, as assigning C channels to the PUs will only make the SUs performance worse. The extra channel allocated to the SUs may be obtained, for example, based on an agreement with the primary network, as it might be of commercial interest for the primary and secondary networks to cooperate [2].

For the sake of mathematical tractability, we make the common assumptions of Poisson arrival processes and exponentially distributed service times. The arrival rate of PU (SU) sessions (flows) to the system is denoted by λ_1 (λ_2). A PU request consumes one channel when accepted and we denote its service rate by μ_1 . In Section III we also evaluate the impact of flow inter-arrival times distributions different from the exponential.

We also assume that SU terminals are rate limited and can use at most one channel, achieving a maximum transfer rate of r (in bits/s). For mathematical tractability, we assume exponentially distributed SUs flow sizes with mean $E[\sigma]$ (in bits). We denote by $\mu_2 = r/E[\sigma]$ the nominal flow completion rate. It is commonly accepted that the statistical distribution of Internet document sizes shows a greater variability than the exponential distribution. However, by using an exponential document size we expect to obtain a lower bound on performance [9]. In Section III we also evaluate the impact of flow size distributions different from the exponential.

The control entity is in charge of allocating resources to the SUs. We assume that it is aware of the state of the channels (free or occupied by a PU), either as a consequence of a sensing procedure (performed by itself, by the SUs or both) or by coordinating its operation with the control entity of the primary network.

We consider two channel access schemes. The first one is the common *random access* (RAN) scheme, where both PUs and SUs select a channel randomly upon arrival, if available. PUs are blocked when no available channels are found, i.e. when there are $(C - 1)$ PUs in the system. However, SUs always join the system regardless of its state.

The second access scheme operates as follows. Assume that the primary network assigns the channels in order and that they are numbered according to this order. We consider that, to setup a PU session, the system searches from left (low channel numbers) to right (high channel numbers) until a free channel is allocated to the new session. Conversely, to setup a new SU flow the system searches from right (high channel numbers) to left (low channel numbers). We call this scheme *channel access with preference* (CAP). PUs and SUs cooperate in the sense that the secondary network is aware of the spectrum access rule followed by the primary network. The cooperation is advantageous for both because the interference is reduced, as shown later.

We denote by $\mathbf{x} = (x_1, x_2)$ the system state, when there are x_1 ongoing PUs sessions and x_2 SUs flows. During overload episodes, we assume that the patience time of a SU is exponentially distributed with rate

$$\mu_p(\mathbf{x}) = \begin{cases} 0 & x_2 \leq (C - x_1), \\ K_p \left(\frac{x_2}{C - x_1} - 1 \right) & x_2 > (C - x_1), \end{cases}$$

where K_p is a constant. The patience times of the SUs are independent and identically distributed. Note that the mean patience time in state \mathbf{x} is $1/\mu_p(\mathbf{x})$, and $\mu_p(\mathbf{x})$ increases from zero upwards when the bandwidth share per SU decreases from one channel (nominal share) downwards.

A. Random Access

For CTMC describing the system dynamics, let $\pi(\mathbf{x})$ be the stationary probability of state \mathbf{x} . The global balance equations for $0 \leq x_1 \leq (C - 1)$ and $0 \leq x_2 \leq CM$ are:

$$\begin{aligned} & \pi(\mathbf{x}) [\lambda_1 I(x_1 < C - 1) + \lambda_2 I(x_2 < CM) \\ & \quad + x_1 \mu_1 + x_2 \mu_2(\mathbf{x})] \\ & = \pi(\mathbf{x} - \mathbf{e}_1) \lambda_1 I(x_1 > 0) + \pi(\mathbf{x} - \mathbf{e}_2) \lambda_2 I(x_2 > 0) \quad (1) \\ & \quad + \pi(\mathbf{x} + \mathbf{e}_1) (x_1 + 1) \mu_1 I(x_1 < C - 1) \\ & \quad + \pi(\mathbf{x} + \mathbf{e}_2) (x_2 + 1) \mu_2(\mathbf{x} + \mathbf{e}_2) I(x_2 < CM) \end{aligned}$$

where, M is the maximum number of SUs that can share a channel, $\mu_2(\mathbf{x}) = f(\mathbf{x}) \mu_2 + \mu_p(\mathbf{x})$,

$$f(\mathbf{x}) = \min \left(1, \frac{C - x_1}{x_2} \right),$$

$I(A)$ is an indicator function, which value is 1 when condition A is satisfied and 0 otherwise, and \mathbf{e}_k denotes a two element vector whose k -th component is equal to 1 and the other component is 0.

As PUs have preemptive priority over the SUs, the PUs perceive a $M/M/C - 1/C - 1$ system and their blocking probability P_1 can be obtained using the well-known Erlang B formula. For the SUs, we are interested in the following performance parameters:

$$P_{if} = \frac{\lambda_1}{\Lambda_1} \sum_{x_1=0}^{C-2} \sum_{x_2=0}^{\alpha(x_1)} \beta(\mathbf{x}) \pi(\mathbf{x}), \quad (2)$$

$$P_{ab} = \frac{1}{\Lambda_2} \sum_{x_1=0}^{C-1} \sum_{x_2=0}^{\alpha(x_1)} x_2 \mu_p(\mathbf{x}) \pi(\mathbf{x}), \quad (3)$$

$$D_2 = \frac{\mu_2}{\Lambda_2} \sum_{x_1=0}^{C-1} \sum_{x_2=0}^{\alpha(x_1)} x_2 \pi(\mathbf{x}), \quad (4)$$

$$Th_2 = \frac{1}{\lambda_2} \sum_{x_1=0}^{C-1} \sum_{x_2=0}^{\alpha(x_1)} x_2 f(\mathbf{x}) \mu_2 \pi(\mathbf{x}). \quad (5)$$

where, P_{if} is the interference probability (the fraction of PU arrivals that vacate a SU), P_{ab} is the abandonment probability (the fraction of accepted SUs that abandon), D_2 , by Little's law, is the actual mean delay normalized to the nominal mean delay ($1/\mu_2$), and Th_2 is the throughput of successfully

completed flows normalized to the offered rate (λ_2). Note also that, $\Lambda_1 = \lambda_1 (1 - P_1)$ and $\Lambda_2 = \lambda_2$ are the PUs and SUs acceptance rates, $\alpha(x_1) = (C - x_1)M$ is the maximum number of SUs in system when there are x_1 active PUs and

$$\beta(\mathbf{x}) = \min \left(1, \frac{x_2}{C - x_1} \right)$$

is the probability that a PU arrival vacates a SU in state \mathbf{x} .

By limiting M we are truncating the state space. However, a good precision is obtained for small values of M . Let us denote by $Th_2(m)$ the value of Th_2 computed with $M = m$. As an example, for the system studied in Section III, with no admission control and no limit on the number of channels used by SUs, $|Th_2(15) - Th_2(16)|/Th_2(15) \approx 10^{-13}$. Note that there can be system limitations that can bound M . For example, hardware or protocol efficiency limitations.

B. Access with Preference

Given that spectrum handover is supported, the selection of the access scheme only changes P_{if} , while the other performance parameters are not affected. Then, when CAP is used, P_{ab} , D_2 and Th_2 can still be obtained from (3) to (5).

To determine P_{if} for the CAP scheme, the state representation must describe not only the number of ongoing PU and SU sessions, but also their physical placement on the spectrum. This makes the cardinality of the state space of the CTMC that models the system to grow very quickly with C . As an example, consider a system with $C = 4$ where the state is represented by (c_1, c_2, c_3, c_4, n) , being $c_i = 0, 1, 2$, and $0 \leq n \leq CM$. Each channel (c_i) can be idle (0), occupied by a PU (1) or by SUs (2), and n is the number of SUs in the system. In a system with $M = 15$, the number of states is 528 for $C = 4$, while it is 1362 for $C = 5$. This makes the solution of the CTMC intractable for any practical scenario. Therefore, we developed a simulation model and validated it with the analytical model of a simple scenario.

C. Control Schemes

To control the interference probability, we limit to $R \leq C$ the number of physical channels that the elastic flows have access to. To control abandonments, we enforce AC. If the number of active flows is equal to a threshold F , then new flow arrivals are rejected. Otherwise, they are accepted. The balance equations for the RAN scenario are the same as (1), but now $0 < x_2 \leq F$, $LR(x_1) = \min((C - x_1), R)$ and

$$f(\mathbf{x}) = \min \left(1, \frac{LR(x_1)}{x_2} \right),$$

where $LR(x_1)$ is the number of channels SUs have access to when there are x_1 PUs.

The performance parameters can be obtained from (2) to

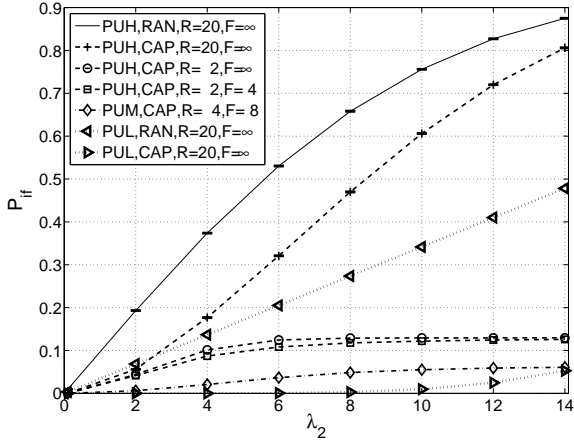


Fig. 1. Interference probability P_{if} for the RAN and CAP schemes.

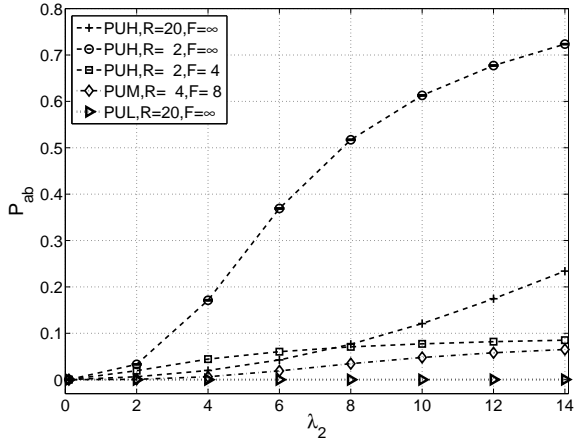


Fig. 2. Abandonment probability P_{ab} for the CAP scheme.

(5), with

$$\Lambda_2 = \lambda_2 \sum_{x_1=0}^{C-1} \sum_{x_2=0}^{\alpha(x_1)-1} \pi(\mathbf{x}),$$

$$\beta(\mathbf{x}) = \min\left(1, \frac{\gamma(x_2)}{C-x_1}\right),$$

$\alpha(x_1) = \min((C-x_1)M, F)$, and $\gamma(x_2) = \min(x_2, R)$, where $\gamma(x_2)$ is the number of channels occupied by the SUs when there are x_2 of them.

As mentioned in the previous Section II-B, for a scenario that deploys the CAP scheme and the interference and abandonment control schemes, P_{ab} , D_2 and Th_2 can still be obtained from (3) to (5) with the appropriate changes described above in this Section.

III. NUMERICAL EVALUATION

We consider a system with $C = 20$ and with three different loads of PUs, that we refer to as high (PUH), medium (PUM) and low (PUL). For these loads, the PUs arrival rates λ_1 are 15, 10, and 5, respectively. For $\mu_1 = 1$, the PUs blocking probabilities P_1 are 0.0637, 0.0037 and 10^{-6} , respectively.

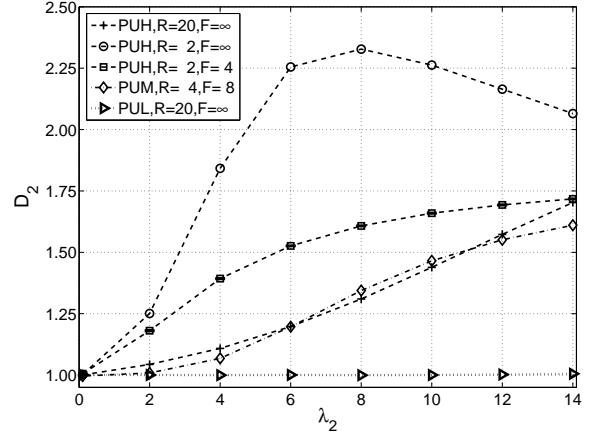


Fig. 3. Normalized SUs delay D_2 for the CAP scheme.

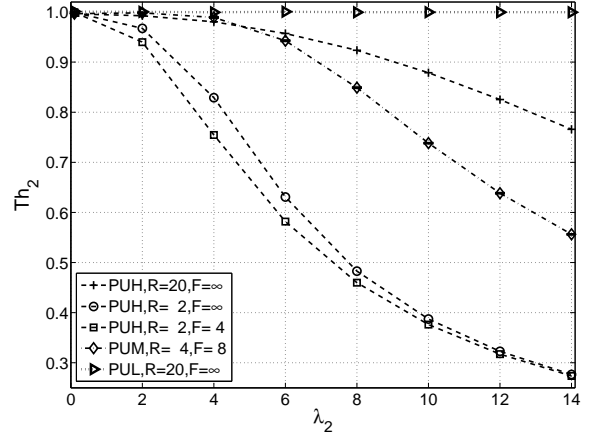


Fig. 4. Normalized SUs throughput Th_2 for the CAP scheme.

The traffic carried by the PUs are 14.04, 9.96 and 5.0 Erlangs, respectively. For the SUs, we set $\mu_2 = 2$ and $K_p = 0.1$. As an example, when the SUs get a channel share of $1/5$, their mean patience time is $5/2$ (5 times their nominal service time). While when they get a share of $1/100$, their mean patience time is $\approx 1/10$ (their nominal service time divided by 5).

Figures 1, 2, 3 and 4 display P_{if} , P_{ab} , D_2 and Th_2 as a function of λ_2 . The analytical results obtained for the systems that use the RAN scheme, are shown with a continuous line in Fig. 1. To validate them, we simulated the system and determined the 95% confidence intervals. Observe the excellent agreement between the analytical and simulation results, i.e. they practically overlap.

For the systems that use the CAP scheme, we determined the P_{if} by simulation. To determine the other performance parameters we used the analytical model. As discussed in Section I, the simulation model is completely independent from the analytical model. Obviously, for the simulation model we used the same inter-arrival and service time distributions (exponential) and the same means that in the analytical model. The analytical results for the CAP scheme have also been validated by simulation. For clarity we only show the

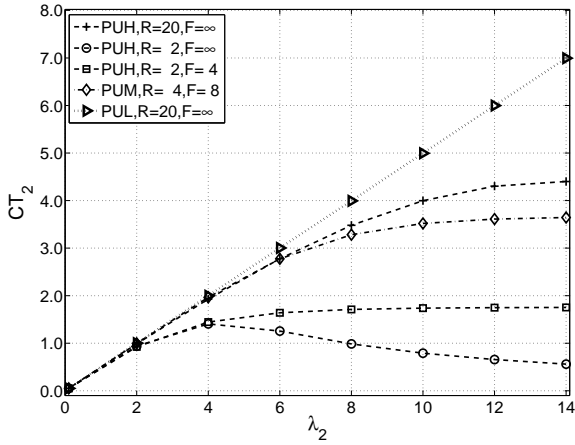


Fig. 5. Traffic carried by successfully completed SU flows.

95% confidence intervals of simulation results for the curves “PUH, $R = 2, F = \infty$ ” of Fig. 2, “PUH, $R = 2, F = 4$ ” of Fig. 3 and “PUM, $R = 4, F = 8$ ” of Fig. 4. Observe again the excellent agreement between the analytical and simulation results, i.e. they practically overlap. Note also that the widths of the confidence intervals are negligible.

The effectiveness of the CAP scheme over the RAN scheme to reduce the interference is shown in Fig. 1 for the PUH and PUL loads and for $R = 20, F = \infty$, which refer to configurations with the interference control and the AC disabled. As an example, compare the curves for the scenarios “PUH, CAP, $R = 20, F = \infty$ ” and “PUH, RAN, $R = 20, F = \infty$ ”. Also, compare the curves for the same scenarios but with a small PUs load (PUL). Note that the effectiveness of CAP increases as the PUs load decreases. As CAP achieves better interference performance, we only discuss scenarios with this access schemes in the sequel.

In order to reduce the interference around 10% in the PUH scenario, we set $R = 2$ and keep $F = \infty$. As observed in the curve “PUH, $R = 2, F = \infty$ ” of Fig. 2, P_{ab} increases considerably. The objective $P_{if} \approx 10\%$ has been chosen arbitrarily to show the effectiveness of the suggested traffic management scheme. To control both the interference and the abandonments, we set $R = 2$ and $F = 4$. As observed in the curve “PUH, $R = 2, F = 4$ ” of Fig. 2, P_{ab} can be effectively controlled by the proposed AC scheme. A similar procedure is followed for the PUM scenario, but here we set $R = 4$ and $F = 8$. For the PUL scenario, the interference and the abandonments are very small and therefore controls are not considered.

By comparing the curves, for example, “PUH, CAP, $R = 2, F = \infty$ ” and “PUH, CAP, $R = 2, F = 4$ ” of Fig. 1, it is also observed that the configuration of the AC scheme (i.e. the value of F) has a negligible impact on the interference. This is an interesting result as it indicates that F can be set independently to control the abandonments.

The impact of the AC scheme on D_2 and Th_2 can be observed in Fig. 3 and Fig. 4. In the scenario “PUH, $R = 2, F = \infty$ ”, D_2 increases rapidly with λ_2 up to a maximum

and then decreases. This behavior shows that abandonments help to stabilize the system. However, it also indicates that a significant number of flows abandon (see Fig. 2), and therefore that a significant waste of resources occurs as λ_2 increases. When the AC is activated (“PUH, $R = 2, F = 4$ ”), D_2 is considerably reduced. The penalty paid by enforcing AC is that Th_2 decreases as new flows can be blocked. If we consider that the feasible range for λ_2 is the one displayed in the figures, then an aggregated perception of the Th_2 drop can be obtained by the expression $(STh_2(F = \infty) - STh_2(F = 4)) / STh_2(F = \infty)$, where $STh_2(F = \infty)$ and $STh_2(F = 4)$ are the surfaces below the corresponding Th_2 curves. With this metric, the Th_2 drop is not significant.

Figure 5 shows the traffic carried by successfully completed flows, which has been obtained by simulation. Note that for a small PUs load (PUL), the PUs operation is transparent to the SUs. However, as the PUs load increases the traffic carried by the SUs decreases. In particular, note that the traffic carried by the SUs in the system scenario (“PUH, $R = 2, F = \infty$ ”) decreases as the SUs load increases. This behavior clearly shows the negative impact that abandonments have on the system performance.

To estimate the effective capacity that the SUs can get from a system using the CAP scheme, we compare two scenarios: “PUH, $R = 2, F = 4$ ” and “PUM, $R = 4, F = 8$ ”. Assume by now that their differences in P_{if} , P_{ab} and D_2 are not significant. In the PUH scenario, the PUs carried traffic is $\approx 14Er$. As a consequence, there would be 6 channels available for SUs in average. To limit the interference we set $R = 2$. The effective capacity is then $2/6 \approx 0.33$ of the available one. Likewise, for the PUM scenario the effective capacity is $4/10 = 0.4$ of the available one. Note that the curves for P_{if} , P_{ab} and D_2 in the scenario “PUM, $R = 4, F = 8$ ” are below the corresponding ones in the scenario “PUH, $R = 2, F = 4$ ”. If we set the scenario “PUH, $R = 2, F = 4$ ” as a reference in terms of P_{if} and P_{ab} , then, comparatively, the effective capacity fraction of the scenario “PUM, $R = 4, F = 8$ ” is slightly higher than 0.4. As discussed before, the impact of the AC on the effective capacity is not significant.

In summary, when the interference and the abandonments are limited by the suggested traffic management mechanisms, the effective capacity obtained by the SUs is only a fraction of the available capacity, and this fraction decreases with the PUs load.

Finally, note that in the scenarios “RAN, $R = C, F = \infty$ ”, that are the ones studied in [5], P_{ab} , D_2 and Th_2 show acceptable values. However, the interference goes out of control as the PUs load increases. Then, limiting the analysis to the SUs transfer delay, as in [5], provides only a partial understanding of the behavior of a cognitive radio system with elastic traffic.

A. Sensitivity to the Flow Size Distribution

In this section we evaluate the sensitivity of the system performance parameters to the distribution of the flow size σ . Figure 6, 7 and 8 show the evolution of different system

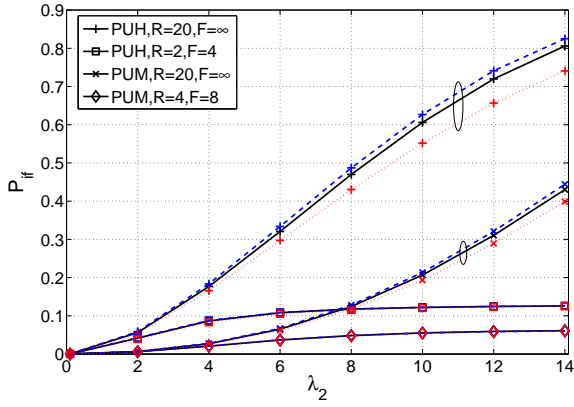


Fig. 6. Sensitivity of P_{if} to the flow size for the CAP scheme.

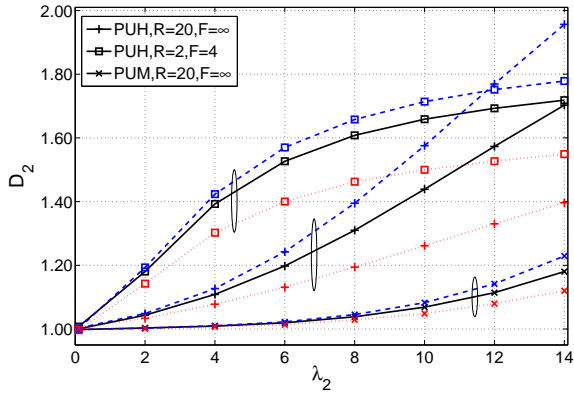


Fig. 7. Sensitivity of D_2 to the flow size for the CAP scheme.

parameters with the arrival rate of the SUs when the flow size follows an Erlang (discontinuous line), exponential (continuous line) or hyper-exponential (dotted line) distribution. We used hyper-exponential distributions as they are more realistic [9]. However, as a reference, we use the Erlang and exponential distributions in the figures of this Section for illustration purposes.

The mean flow size $E[\sigma]$ is kept constant for all distributions. The coefficient of variation CV_σ for the Erlang

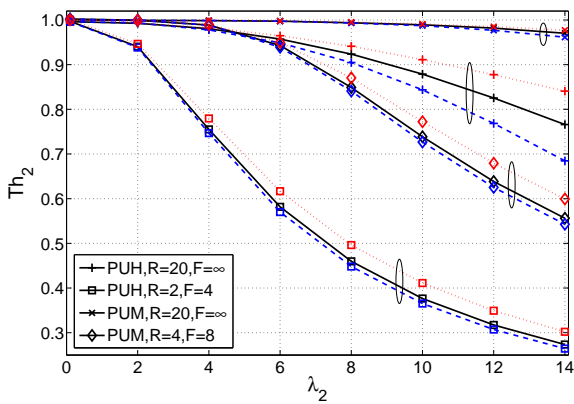


Fig. 8. Sensitivity of Th_2 to the flow size for the CAP scheme.

and hyper-exponential distributions are set to 0.5 and 2, respectively. We use a special type of a two stage hyper-exponential distribution that requires only two parameters (mean and standard deviation) for its characterization. This is achieved by using *balanced means*, i.e., $p_1/\gamma_1 = p_2/\gamma_2$, where γ_1 and γ_2 are the parameters of the two exponentials, and p_1 and p_2 are the fraction of flows with mean sizes $1/\gamma_1$ and $1/\gamma_2$, respectively. In that case,

$$p_1 = \frac{1}{2} \left(1 + \sqrt{\frac{CV_\sigma^2 - 1}{CV_\sigma^2 + 1}} \right), \quad p_2 = 1 - p_1,$$

$$\gamma_1 = \frac{2p_1}{E[\sigma]}, \quad \gamma_2 = \frac{2p_2}{E[\sigma]}.$$

Then, approximately 89% of the flows are small and the remaining 11% are large, which size is in average approximately 7.87 times larger than the small ones.

As observed in Fig. 6 the probability of interference P_{if} is almost insensitive to the flow size distribution, particularly when the AC scheme is active. It seems to depend only on the admitted PUs and SUs load. The abandonment probability P_{ab} (not shown), is almost insensitive to the flow size distribution, except when the two traffic management schemes are inactive and the PUs load is high, which is the scenario where abandonments are more frequent. There, smaller flows will be less affected by impatience. As the population of small flows is larger when the flow size is hyper-exponentially distributed, then the fraction of accepted flows that abandon P_{ab} is smaller than for the other two distributions.

For clarity, in Fig. 7 the scenario “PUM, $R = 4, F = 8$ ” is omitted, as its results are close to the ones obtained for “PUH, $R = 20, F = \infty$ ”. Observe that the normalized flow transfer delay is quite sensitive to the flow size distribution and its sensitivity increases with the system load, both PUs and SUs load. As before, the larger the short flow population is (hyper-exponential distribution) the smaller D_2 will be, as their weight on D_2 is more significant. This result suggests that enforcing deliberate discriminatory sharing might improve system performance. A similar reasoning can be applied to explain the behavior of the normalized throughput Th_2 in Fig. 8. Although not shown, the results for CT_2 reinforce previous reasoning about the dominant impact that small flows have on performance, as the traffic carried by the SUs completed flows is slightly smaller for the hyper-exponential distribution.

B. Sensitivity to the Impatience

In this section we evaluate the sensitivity of the system performance parameters to the mean impatience time. All results have been obtained for exponentially distributed impatience times and hyper-exponentially distributed flow sizes. Figure 9 shows the evolution of the interference probability P_{if} with the arrival rate of the SUs when the parameter of the impatience time takes the values: $K_p = 0.01$ (discontinuous line), 0.1 (continuous line) and 1.0 (dotted line). Note that 0.1 was chosen as the default value. Also, $K_p = 1.0$ might represent

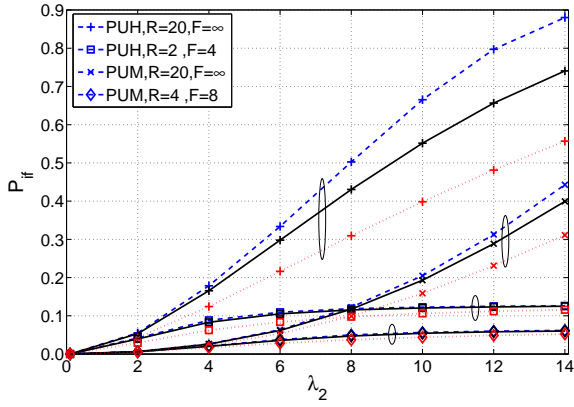


Fig. 9. Sensitivity of P_{if} to the impatience duration for the CAP scheme.

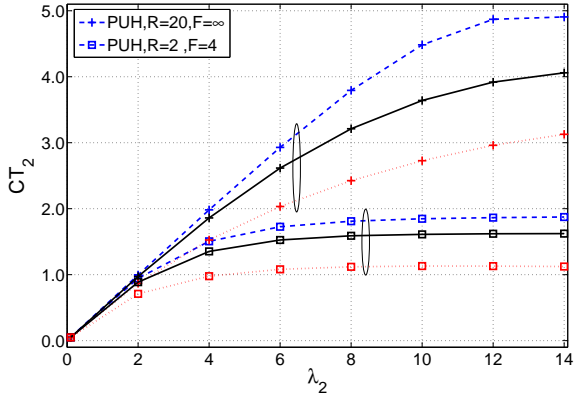


Fig. 10. Sensitivity of CT_2 to the impatience duration for the CAP scheme.

a possibly unrealistic value, as when the SUs get a channel share of $1/5$, their mean patience time would be $1/4$ (half their nominal service time). As expected, P_{if} increases with the patience time, as the SUs carried traffic CT_2 increases, as shown in Fig. 10. However, in systems with the traffic management controls active, the impact of the patience time on P_{if} is negligible.

Although not shown, the abandonment probability P_{ab} is quite sensitive to the impatience time, which was also ex-

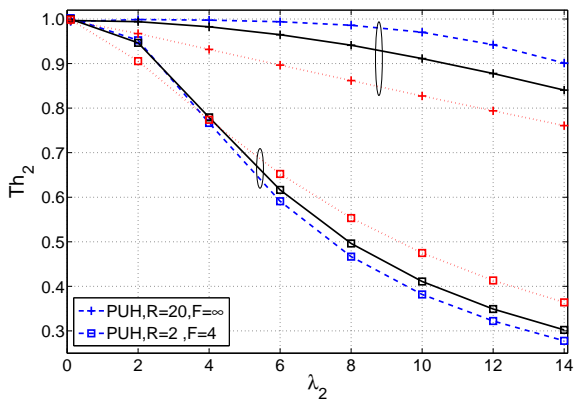


Fig. 11. Sensitivity of Th_2 to the impatience duration for the CAP scheme.

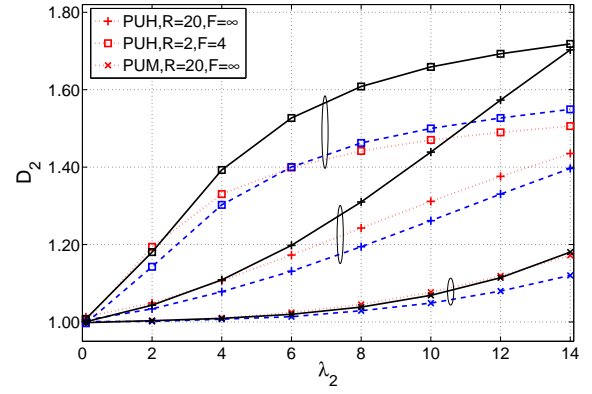


Fig. 12. Sensitivity of D_2 to the flow arrival distribution for the CAP scheme.

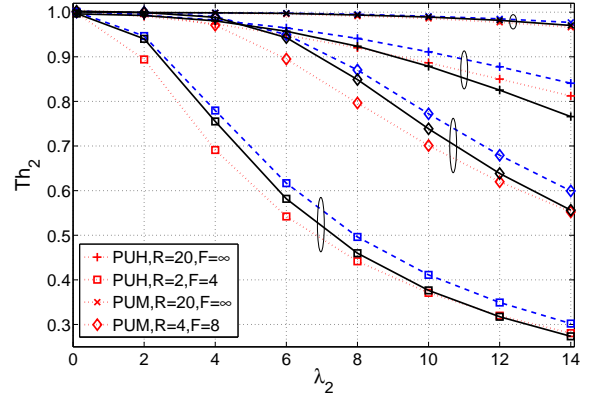


Fig. 13. Sensitivity of Th_2 to the flow arrival distribution for the CAP scheme.

pected. For the scenarios “PUH” and “PUM”, the normalized transfer delay of the successfully completed flows D_2 takes values below 1 for the very impatient flows ($K_p = 1.0$). This clearly shows that most of the completed flows are short flows. Finally, the scenario “PUM, $R = 2, F = 4$ ” of the Fig 11 shows that for low SUs loads, the rate of successful completions increases with the patience time. However, this behavior is reversed beyond certain SUs load. A similar behavior was described in [9].

C. Sensitivity to the Flow Arrival Distribution

We evaluate now the sensitivity of the system performance parameters to the SUs flow arrival distribution. If we denote by τ the SUs inter-arrival time and recalling that σ was the flow size, then we study three system configurations that correspond to: τ and σ being exponentially and hyper-exponentially distributed (discontinuous line), τ and σ being both hyper-exponentially distributed (dotted line), and τ and σ being both exponentially distributed (continuous line). Note that the curves (continuous lines) for the last configuration are only displayed as a reference and will not be discussed along this Section. We use a coefficient of variation of 2 for the hyper-exponential distributions, and we keep the means constant for the same system loads. Following the same reasoning used for the flow sizes, when the inter-arrival times are hyper-

exponentially distributed, the arrival pattern is composed of groups of flows, with individual flows in a group arriving very close, and long inter group times.

We note in Fig. 12 that D_2 increases with the variability of the SUs arrivals, particularly for the systems without controls. However, for the systems with controls, a SUs load can be found that defines a change in system behavior. Below this load, D_2 is bigger for highly variable SUs arrivals (hyper-exponential), but above this load, D_2 is bigger when the SUs arrivals are less variable (exponential). When the SUs load is low, most of the arrivals will be accepted. Then, if SU arrivals are less variable, the number of flows sharing a channel will be small and D_2 will be small. However, if arrivals occur in “groups”, it is highly probable that most of the flows arriving to the system will have to share a channel, at least with the other flows of the same group, and D_2 increases. On the other hand, when the SUs load is high, many arrivals will be rejected due to the AC mechanism. Then, when arrivals occur in “groups” the rejection ratio will be higher than when the arrivals occur with less variability. This has a beneficial impact on D_2 , as a high rejection ratio leads to a reduction on the number of flows sharing a channel and, therefore, to a reduction on the transfer delay of successfully completed flows.

Although not shown, a similar pattern is followed by P_{ab} . We also observe that CT_2 is almost insensitive to the SUs arrival distribution, and that CT_2 slightly decreases when the variability of the SU arrivals increases. The interference probability P_{if} is almost insensitive to the inter-arrival time distribution for the scenario “PUH, $R = 20, F = \infty$ ” and for the scenarios with active traffic management controls. However, for the scenario “PUM, $R = 20, F = \infty$ ”, P_{if} slightly increases with the variability of the SUs arrivals.

We observe in Fig. 13 that Th_2 decreases with the variability of the SUs arrivals. Also, we note that the biggest difference between the discontinuous and dotted line curves occur for those SUs load points discussed before that define a change in the system behavior. This might be due to the combined effect of abandonments and the AC scheme rejections.

Finally, note that although for D_2 the exponential distributions for τ and σ (continuous line) represent approximately a performance bound, for Th_2 this is not always the case. However, note that regardless of the distributions used for the different random variables, the performance parameters studied follow a similar trend.

IV. CONCLUSIONS

We studied a cognitive radio network where the SUs are impatient. We analyzed the interference created by the SUs when the system uses two different access schemes: RAN and CAP. Our results show that the CAP scheme outperforms the RAN scheme, and that, comparatively, the interference reduction increases as the PUs load decreases. To further reduce the interference, a scheme that limits the number of channels that SUs have access to is studied. Additionally, an

admission control scheme for SUs is proposed and evaluated to improve the QoS perceived by the active SUs.

We determined that when the interference and abandonments are limited by the suggested traffic management mechanisms, the effective capacity obtained by the SUs is only a fraction of the available capacity, and this fraction decreases with the PUs load. We also found that the interference management and admission control mechanisms can be tuned independently. This fact simplifies the adjustment of the schemes to varying traffic conditions.

We studied by simulation the sensitivity of the system performance parameters to the flow size distribution and flow inter-arrival time. We used hyper-exponential distributions instead of exponential ones, as they are more realistic. We observed that, in general, when the traffic management schemes are active, the system is less sensitive to the distributions of these random variables. We also studied the sensitivity of the system performance parameters to the impatience time and found a high sensitivity even with the traffic management schemes active.

REFERENCES

- [1] Y. Liang, K. Chen, G. Li, and P. Mahonen, “Cognitive radio networking and communications: An overview,” *Vehicular Technology, IEEE Transactions on*, vol. 60, no. 7, pp. 3386–3407, 2011.
- [2] J. Peha, “Sharing spectrum through spectrum policy reform and cognitive radio,” *Proceedings of the IEEE*, vol. 97, no. 4, pp. 708–719, 2009.
- [3] J. Martinez-Bauset, V. Pla, M. Domenech-Benlloch, and D. Pacheco-Paramo, “Admission control and interference management in dynamic spectrum access networks,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, Article ID 708029, 2010.
- [4] H. Al-Mahdi, M. A. Kalil, F. Liers, and A. Mitschele-Thiel, “Increasing spectrum capacity for ad hoc networks using cognitive radios: an analytical model,” *IEEE Communications Letters*, vol. 13, no. 9, pp. 676–678, Oct. 2009.
- [5] E. Wong and C. Foh, “Analysis of cognitive radio spectrum access with finite user population,” *IEEE Communications Letters*, vol. 13, no. 5, 2009.
- [6] M. NoroozOliaee, B. Hamdaoui, and K. Tumer, “Achieving optimal elastic traffic rewards in dynamic multichannel access,” in *2011 International Conference on High Performance Computing and Simulation (HPCS)*, July 2011, pp. 155–161.
- [7] F. Yu, Z. Qinyu, W. Ye, and G. Xin, “Variable bandwidth spectrum access for secondary user in cognitive radio networks,” in *2011 International Conference on Wireless Communications and Signal Processing (WCSP)*, Nov. 2011, pp. 1–5.
- [8] J. Elias and F. Martignon, “Joint spectrum access and pricing in cognitive radio networks with elastic traffic,” in *2010 IEEE International Conference on Communications (ICC)*, May 2010, pp. 1–5.
- [9] T. Bonald and J. W. Roberts, “Congestion at flow level and the impact of user behaviour,” *Computer Networks*, vol. 42, no. 4, pp. 521–536, 2003.
- [10] L. Yang, L. Cao, and H. Zheng, “Proactive channel access in dynamic spectrum networks,” *Physical Communication*, vol. 1, no. 2, pp. 103–111, 2008.
- [11] T. Bonald and A. Proutière, “Insensitive bandwidth sharing in data networks,” *Queueing Systems: Theory and Applications*, vol. 44, no. 1, pp. 69–100, 2003.
- [12] A. T. Hoang, Y.-C. Liang, and M. Islam, “Power control and channel allocation in cognitive radio networks with primary users’ cooperation,” *IEEE Transactions on Mobile Computing*, vol. 9, no. 3, pp. 348–360, Mar. 2010.