EURECOM

EURECOM

Sophia Antipolis

EURECOM
Department of Mobile Communications
Campus Sophia Tech
Les Templiers
450 route des Chappes
B.P. 193
06410 Biot
FRANCE

Research Report RR-12-274

Analysis of Cognitive User Performance under Generic Primary User Activity

October 1$^{\text{th}}$, 2012
Last update October 12$^{\text{th}}$, 2012

Fidan Mehmeti, Thrasyvoulos Spyropoulos

Tel : (+33) 4 93 00 81 00
Fax : (+33) 4 93 00 82 00
Email : {mehmeti,spyropou}@eurecom.fr

Analysis of Cognitive User Performance under Generic Primary User Activity

Fidan Mehmeti, Thrasyvoulos Spyropoulos

Abstract

Cognitive Networks have been proposed to opportunistically discover and exploit (temporarily) unused licensed spectrum bands. With the exception of TV white spaces, secondary users (SUs) can access the medium only *intermittently*, due to deferring to primary user (PU) transmissions and scanning for new channels. This raises the following questions: (i) what sort of delays can an SU expect on a channel given the PU utilization of this channel? (ii) how do specific characteristics of the PU activity patterns (e.g. burstiness) further affect performance? These questions are of key importance for the design of efficient algorithms for scheduling, spectrum handoff, etc. In this paper, we propose a queueing analytical model to answer them. We model the PU activity pattern as an ON-OFF alternating renewal process with generic ON and OFF durations, and derive a closed form expression for packet delays by solving a variant of the M/G/1 queue. Contrary to the common belief that low utilization channels are good channels, we show that the expected SU delay on a channel, and thus the best channel to use, is a subtle interplay between the ON and OFF duration distributions of the primary users, and the SU traffic load. We validate our analysis against simulations for different PU activity profiles.

Index Terms

Cognitive networks, Queueing, Renewals.

Contents

List of Figures

1   Introduction

Measurements of the utilization of licensed wireless spectrum have (somewhat counter-intuitively) revealed that the available spectrum is rather under-utilized, exhibiting high variability across space, frequency, and time [1]. Yet, the current lack of flexibility in dynamically assigning spectrum to match demand over time and space further limits the service levels offered, in addition to rapidly increasing demand [2].

Cognitive radios and networks have been proposed to address this problem. Cognitive users (also referred to as "Secondary Users (SU)") can sense a range of licensed or unlicensed bands and, if found idle, opportunistically use one or more of them to meet the user/application demands. However, most of these channels are only *temporarily* available, when the licensed user (also referred to as "Primary User (PU)") is not transmitting or receiving on them (one exception are TV white spaces [3], which can be known in advance and available for very long periods of time).

Such interruptions make media access by SUs *intermittent*, especially in densely utilized parts of the spectrum. When the SU cannot transmit anymore on the current channel(s), it will (in the simplest case) either have to wait for the channel to become available again or switch its radio to scanning mode (assuming a single radio) to discover other available channels. This can delay ongoing or new SU transmissions. The probability and duration of such delays depends on the PU's activity characteristics (percentage of time being idle, duration of idle periods, duration of traffic bursts, etc.). These characteristics can be highly variant, since the (PU) channels might belong to different primary wireless systems, be used to carry different types of traffic (e.g. voice, file transfer, web browsing, video streaming, etc.), and be governed by different protocols managing the access to this channel. All of these features impact the PU activity observed on a chosen channel by a cognitive user in non-trivial ways.

These observations lead to some important questions: (i) What is the level of performance (e.g. delay) that a secondary user should expect on a channel, given the PU activity characteristics on this channel? (ii) Does the *average* "amount" of PU activity (e.g. being active for 40% of the time) suffice to predict SU performance, or can differences between PU activity profiles (e.g. distribution of active periods) further affect SU performance and to what extent? These questions have implications for the design of efficient spectrum scheduling, scanning, and handoff strategies, among other things.

To this end, in this paper we propose a queueing analytic model for the performance of secondary user transmissions for *general* PU activity patterns. Contrary to the common belief that a channel with lower average PU activity is a better channel, our results reveal that the SU delay is in fact a subtle interplay between PU activity statistics (1st, 2nd and 3rd moments of the random durations of PU *active* and *idle* periods), and the secondary user traffic load. As one example, for applications with low traffic intensity (e.g. machine-to-machine communications [4]),

choosing the channel with the lowest PU utilization might lead to higher delays (in fact, arbitrarily worse *in theory*).

Our model and analysis are presented in detail in Section 2. Then, in Section 3 we validate our theory using simulations of different PU activity profiles including some realistic models, recently proposed [5,6]. Related work is given in Section 4. We conclude our work in Section 5.

2 Performance modeling

Primary User Model: Consider a single channel used by one or more primary users. We assume that the state of this channel can be either active ("ON"), i.e. a primary packet is transmitted (it is indifferent to the SU whether this is data, signaling, uplink or downlink traffic) or idle ("OFF"), as depicted in Fig. 1. The exact duration of ON and OFF periods depends first on user behavior (e.g. arrival process of PU traffic flows) [5]. Furthermore, it depends on the type of traffic (e.g. short VoIP packets vs. long file transfer packets), system details (e.g. number of independent users multiplexed on the channel) and intricate protocol interactions (e.g. MAC layer carrier sense, TCP mechanisms etc.). Such details cannot be known or inferred by the secondary user.

In order to allow for the maximum amount of generality, while maintaining analytical tractability, we thus model the ON-OFF activity pattern of PUs as an *alternating renewal process* [7]: $(T_{ON}^{(n)}, T_{OFF}^{(n)}), n \geq 1$, where n denotes the number of $ON - OFF$ cycles elapsed until time t. The duration of any ON period, $T_{ON}^{(n)}$ (OFF period $T_{OFF}^{(n)}$, respectively), is a random variable distributed according to some probability distribution F_{ON} (F_{OFF}), and independently of other ON or OFF periods[1]. This model for PU activity is significantly more flexible in capturing different types of PUs than simple memoryless or "half-memoryless" models used in related work [8,9]. As we will see in Section 3, we can accommodate recently proposed models of PU activity that attempt to mimic real traffic datasets.

Secondary User Model: In this paper, we will assume that SUs can only access a channel at times where there is no PU activity (i.e. during OFF periods). While our model could be extended in other directions, we defer this to future work. [2]

The success of an SU packet transmission depends on the duration of that packet and the (remaining) duration of the OFF period. We assume that if a PU starts transmitting before the SU packet is completely sent, the SU transmission is considered lost (collision), and has to be retransmitted in the next idle (OFF) pe-

---

[1]As usual, we denote probability distribution functions with capital letters, e.g. F(t), and the respective density function with lowercase, e.g. f(t).

[2]Our model does not apply directly to spectrum sharing based on "interference temperature", where SUs are allowed a budget of interference while concurrently transmitting with PUs. Yet, the FCC has recently commented that the interference temperature approach might not be a workable concept, increasing interference in the frequency bands where it would be used [1].
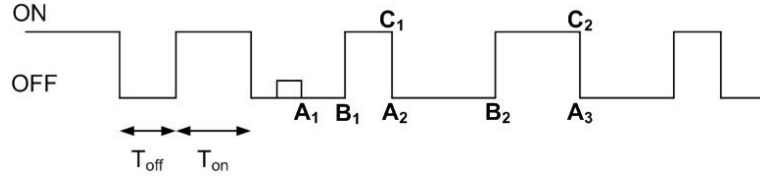
Figure 1: The acivity pattern of the primary user

riod. We also assume that the spectrum sensing ability of cognitive users is perfect. This means that there are no missdetections or false alarms. Such events however are orthogonal to our model and analysis, and could be captured in the PU ON-OFF process, if needed.

In practice, a cognitive (SU) user can have access to multiple channels. A number of architectures and protocols have been proposed [1] to discover and access such channels. To avoid including fine architectural details that could make our analysis intractable and reduce the generality of our results, we choose to maintain the abstraction of a single stochastic ON-OFF process. When the state of the process is ON, it means that the SU cannot transmit any new or queued traffic. This ON period might correspond, among other things, to: (i) the SU waiting on a busy channel until it becomes available again, in situations where spectrum mobility is expensive or unlikely to yield better results; (ii) a scanning period during which the SU tries to gain or regain one (or more) idle channel(s). Similarly, an OFF period (during which the SU can transmit) may correspond to using different spectrum holes "back-to-back", e.g. when backup channels are maintained [10].

2.1 Arrival and Service Processes for SU Traffic

Our goal is to use a queueing model to derive the expected delay SU packets will experience, depending on the type of PU channel activity they "compete" with. We will assume that packets are generated at the cognitive user according to a Poisson process with rate λ. This will allow us to focus on the effect of PU activity on the service process, and get exact expressions for queueing delays. We defer the treatment of generic arrivals to future work. Without loss of generality, we also assume that packet sizes are fixed and equal to Δ (size normalized for transmission rate). A generic packet size distribution can be easily integrated in our derivations.

The service time that an SU packet experiences depends on the packet size, but also on the state of the PU (ON or OFF) when the packet arrives. While this points to an M/G/1 system [11], this is not the appropriate model. In fact, there are two service time distributions, *S'* and *S"*, depending on the (SU) queue length at the time of arrival:

Service S': Consider SU packets arriving to find the queue empty (no SU packets in queue or being transmitted). Then,

  (i) if the PU is ON, the packet will have to wait until the next OFF cycle;

(ii) if the PU is OFF, it can immediately start transmission.

The success of this (first) transmission attempt depends on the duration (case (i)) or the remaining duration (case (ii)) of this OFF cycle. In case of failure (PU restarts too soon), the SU packet will have to wait and retry in the next OFF period(s).

Service S": Consider an SU packet that finds the SU system busy. This packet will have to queue until all packets in front of it are successfully transmitted. However, it is guaranteed that its "service" will start in an OFF period, since the packet in front of it *just* finished transmitting (successfully). [3]

The next two Lemmas derive the expected service times $E[S']$ and $E[S'']$, as a function of the PU activity profile (ON-OFF statistics). Theorem 3 then combines the two service time distributions to derive the total system delay for arriving SU packets, including both queueing and service time. We assume throughout unlimited queue size and FCFS service order.

Before we proceed, we summarize in Table 1 a number of variables and useful (shorthand) notations that we will use in our results and derivations.

Table 1: Variables and Shorthand Notation

Variable	Definition/Description
T_{ON}	Duration of ON periods
T_{OFF}	Duration of OFF periods
$T_{OFF}^{(f)}$	$T_{OFF}\|T_{OFF} < \Delta$: OFF period with duration smaller than Δ
$T_{OFF}^{(e)}$	Excess OFF period
$f_{OFF}^{(e)}(x)$	$\frac{1-F_{OFF}(x)}{E[T_{OFF}]}$
p_1	$\int_0^\infty e^{-\lambda t_{off}} f_{OFF}(t_{off}) dt_{off}$
p_2	$\int_0^\infty e^{-\lambda t_{on}} f_{ON}(t_{on}) dt_{on}$
P	$\int_0^\infty e^{-\lambda t_{off,e}} f_{OFF}^{(e)}(t_{off,e}) dt_{off,e}$
p	Prob.of transmission success in an OFF period: $p = P[T_{OFF} > \Delta]$
N	Number of (extra) ON-OFF cycles until successful transmission
T_1	Duration of (extra) ON-OFF cycles (for packet arriving in ON period)
T_2	Duration of (extra) ON-OFF cycles (for packet arriving in OFF period)

Lemma 1. *The mean service time of non-queued packets $E[S']$ is given by*

$$
\begin{aligned}
E[S'] \;&= \Delta + \frac{P}{1-p_1 p_2}\left(\Delta \cdot p_2 + \left(e^{\lambda\Delta}-1\right)\left(\Delta + E[T_1] - \frac{1}{\lambda}\right)\right) \\
&+ \frac{P}{1-p_1 p_2}\left(E[T_{ON}] + \left(E[T_2] - \frac{1}{\lambda}\right)(1-p_2)\right).
\end{aligned}
\tag{1}
$$

---

[3]We remind the reader that the term "system time" is typically used for the total delay (queueing + service) of a packet, while "service" begins when the packet arrives at the front of the SU queue and lasts until the packet is successfully transmitted.

Proof. The delay of an S' packet (finding no other SU packets queueing or in transmission) depends on its arrival time, relative to the PU state during and after that time. The key to deriving this delay is to notice the following: If we considered a single, isolated packet, we could use the *inspection paradox* to derive the expected delay [7]; e.g. renewal theory tells us that the stationary probability of arriving during an ON period is $\frac{E[T_{ON}]}{E[T_{ON}]+E[T_{OFF}]}$. However, this is only the limiting case, when the SU traffic arrival rate λ goes to 0 (i.e. SU traffic is very sporadic). In fact, the time until the arrival of the next S' packet starts counting *from the point the last queued packet got transmitted, which can only occur during an OFF period.* The situation is depicted in Fig. 1. The higher λ the higher is the probability of the packet arriving in the same OFF period.

To account for this effect, we assume a given realization ("sample path") of the ON-OFF process, represented by a vector t_s of ON and OFF durations: $t_s = \{t_1^{OFF,e}, t_1^{ON}, t_2^{OFF}, t_2^{ON}, \dots \}$, where index 1 corresponds to the (OFF) cycle when the last packet of the previous busy period got successfully transmitted. Note that for the first OFF cycle we consider the remaining ("excess") time $t_1^{OFF,e}$ right *after* the end of the last packet transmission (that ended a "busy" cycle). The delay S' on this sample path can then be expressed as follows (we will later take the expectation over all sample paths):

$$S' = I_{OFF}^1 S'_{OFF} + I_{ON}^1 S'_{ON} + I_{OFF}^2 S'_{OFF} + I_{ON}^2 S'_{ON} + \dots, \qquad (2)$$

where I_{OFF}^i and I_{ON}^i are indicator random variables, which have value 1, only if the S' packet arrival happens in that OFF (ON) period. Clearly, only one such term can be non-zero for a given sample path. We separate this sum into two terms

$$Y_{OFF} = \sum_{i=1}^{\infty} I_{OFF}^{(i)} S'_{OFF} \text{ and } Y_{ON} = \sum_{i=1}^{\infty} I_{ON}^{(i)} S'_{ON}. \qquad (3)$$

For this sample path, delay S' will depend on the time until the next SU packet arrival, which is exponential with rate λ. The expectation of terms in Y_{OFF} is then

$$
\begin{aligned}
E[I_{OFF}^{(i)} S'_{OFF}] &= \int_{A_i}^{B_i} \Delta \lambda e^{-\lambda \cdot x} dx \\
&+ \int_{B_i-\Delta}^{B_i} (B_i - x + T_1) \lambda e^{-\lambda \cdot x} dx, \qquad (4)
\end{aligned}
$$

where $A_1 = 0, A_2 = t_1^{OFF,e} + t_1^{ON}, A_3 = t_1^{OFF,e} + t_1^{ON} + t_2^{OFF} + t_2^{ON}, \dots$ and $B_1 = t_1^{OFF,e}, B_2 = t_1^{OFF,e} + t_1^{ON} + t_2^{OFF}, \dots$, as depicted in Fig. 1.

The first integral is the case when the OFF period the packet arrives in is long enough for the packet to be transmitted immediately (i.e. delay $S' = \Delta$). The second integral is the case when the (remaining) OFF period is smaller than the packet size: then, transmission fails, a delay equal to that remaining time is "paid" (note

5

that this delay is between 0 and Δ, otherwise transmission would be successful), and additional ON-OFF periods must be experienced before successfull transmission. The number of such periods is a random variable, denoted by N. The total duration of these periods is thus $T_1 = \sum_{i=1}^{N} T_{ON} + \sum_{i=1}^{N-1} T_{OFF}^{(f)}$. N is a stopping time, so the expectation $E[T_1]$ can be found using Wald's equation [7]:

$$E[T_1] = \frac{1}{p} E[T_{ON}] + \left(\frac{1}{p} - 1 \right) E[T_{OFF} \mid T_{OFF} < \Delta], \qquad (5)$$

For the conditional expectation that appears in Eq.(5) we have $E[T_{OFF} \mid T_{OFF} < \Delta] = \int_0^\Delta \frac{x \cdot f_{OFF}(x)}{F_{OFF}(\Delta)} \mathrm{d}x$. Since the duration of OFF cycles is independent and distributed as $F_{OFF}(x)$, N is geometrically distributed with probability $p = P[T_{OFF} > \Delta] = 1 - F_{OFF}(\Delta)$. Calculating the integrals in Eq.(4) yields

$$E[I_{OFF}^{(i)} S'_{OFF}] = \Delta e^{-\lambda A_i} + \left(e^{\lambda \Delta} - 1 \right) e^{-\lambda B_i} \left(\Delta + E[T_1] - \frac{1}{\lambda} \right). \qquad (6)$$

Summing over all the OFF terms of Y_{OFF} in Eq.(3),

$$E[Y_{OFF} \mid t_s] = \Delta \sum_i e^{-\lambda A_i} + \left(e^{\lambda \Delta} - 1 \right) \left(\Delta + E[T_1] - \frac{1}{\lambda} \right) \sum_i e^{-\lambda B_i}$$

This is the expectation of Y_{OFF}, conditional on the ON-OFF sample path t_s. Finally, we take the expectation over all possible sample paths

$$E[Y_{OFF}] = \int_0^\infty E[Y_{OFF} \mid t_s] f_{t_s} \left(x_1^{OFF,e}, x_1^{ON}, \dots \right) dx_1^{OFF,e} dx_1^{ON} \dots \qquad (7)$$

Since ON and OFF periods are IID, we can split this integral into a product of expectations, and after some calculus we get

$$E[Y_{OFF}] = \Delta \left(1 + P p_2 \sum_i (p_1 p_2)^i \right) + \left(e^{\lambda \Delta} - 1 \right) \left(\Delta + E[T_1] - \frac{1}{\lambda} \right) P \sum_i (p_1 p_2)^i,$$

where $P = \int_0^\infty e^{-\lambda t_{off,e}} f_{T_{off,e}}(t_{off,e}) dt_{off,e}$. Calculating the geometric sums we have

$$E[Y_{OFF}] = \Delta + \frac{P}{1 - p_1 p_2} \left(\Delta p_2 + \left(e^{\lambda \Delta} - 1 \right) \left(\Delta + E[T_1] - \frac{1}{\lambda} \right) \right). \qquad (8)$$

Using similar steps, we can calculate the term Y_{ON} of Eq.(3), related to packets S' arriving during an ON period

$$E[I_{ON}^{(i)} S'_{ON}] = \int_{B_i}^{C_i} (C_i - x + T_2) \lambda e^{-\lambda \cdot x} dx, \qquad (9)$$

6

where $C_i \in \left\{ t_1^{ON} + t_1^{OFF,e}, t_1^{ON} + t_1^{OFF,e} + t_2^{OFF} + t_2^{ON}, \dots \right\}$. C_i are also shown in Fig. 1. T_2 is the additional delay caused by unsuccessful packet transmissions, after the first excess ON period $T_2 = \sum_{i=1}^{N-1} \left(T_{ON} + T_{OFF}^{(f)} \right) + \Delta$. Hence,

$$E[T_2] = \left(\frac{1}{p} - 1 \right) \left(E[T_{ON}] + E[T_{OFF} \mid T_{OFF} < \Delta] \right) + \Delta. \tag{10}$$

After solving the integral in (9), we get

$$E[I_{ON}^{(i)} S'_{ON}] = (C_i - B_i) e^{-\lambda B_i} + \left(E[T_2] - \frac{1}{\lambda} \right) \left(e^{-\lambda B_i} - e^{-\lambda C_i} \right). \tag{11}$$

Summing over all terms in Eq.(3)

$$E[Y_{ON} \mid t_s] = \sum_i (C_i - B_i) e^{-\lambda B_i} + \left(E[T_2] - \frac{1}{\lambda} \right) \sum_i \left(e^{-\lambda B_i} - e^{-\lambda C_i} \right).$$

This is the expectation of Y_{ON}, conditioned on the ON-OFF sample path t_s. Finally, we take the expectation over all possible sample paths

$$E[Y_{ON}] = \int_0^\infty E[Y_{ON} \mid t_s] f_{t_s} \left(x_1^{OFF,e}, x_1^{ON}, \dots \right) dx_1^{OFF,e} dx_1^{ON} \dots \tag{12}$$

As before, this integral can be splitted into a product of expectations, which after some calculus yields

$$E[Y_{ON}] = P \cdot E[T_{ON}] \sum_i (p_1 p_2)^i + \left(E[T_2] - \frac{1}{\lambda} \right) \left(P \sum_i (p_1 p_2)^i - P p_2 \sum_i (p_1 p_2)^i \right) \tag{13}$$

Calculating the geometric sums in Eq.(13) we have

$$E[Y_{ON}] = \frac{P}{1 - p_1 p_2} \left(E[T_{ON}] + \left(E[T_2] - \frac{1}{\lambda} \right) (1 - p_2) \right). \tag{14}$$

Finally, by summing Eq.(14) with Eq.(8) we have Eq.(1).

$\square$

As mentioned previously, if the arriving packet finds other packets in the system, it has to be queued and wait until its turn. A type 2 packet can start its service only during an OFF period. The following Lemma gives the average service time of a type 2 packet.

Lemma 2. *The mean service time of queued packets $E[S'']$ is given by*

$$E[S''] = \Delta + \int_0^\Delta x f_{OFF}^{(e)}(x) dx + \frac{1}{p} \left(E[T_{ON}] + E[T_{OFF} \mid T_{OFF} < \Delta] \right) \int_0^\Delta f_{OFF}^{(e)}(x) dx. \tag{15}$$

7

Proof. The service time of a packet that finds the system busy is

$$S'' = \Delta + I\left(T_{OFF}^{(e)(f)} + \sum_{i=1}^{N} T_{ON} + \sum_{i=1}^{N-1} T_{OFF}^{(f)}\right). \tag{16}$$

$T_{OFF}^{(e)(f)}$ is the random variable that denotes the first OFF period in which a type 2 packet enters service (excess OFF period) and is smaller than the packet size. I is an indicator random variable that describes the inability of the packet to be transmitted in the first attempt. It is defined as

$$I = \begin{cases} 1 & \text{if } T_{OFF}^{(e)} < \Delta \\ 0 & \text{otherwise} \end{cases}$$

Its expectation is $E[I] = P[T_{OFF}^{(e)} < \Delta]$. Taking the expectations in Eq.(43)

$$E[S''] = \Delta + E\left[IT_{OFF}^{(e)(f)}\right] + E[I]E\left[\sum_{i=1}^{N} T_{ON}\right] + E[I]E\left[\sum_{i=1}^{N-1} T_{OFF}^{(f)}\right] \tag{17}$$

and rearranging we obtain

$$E[S''] = \Delta + E\left[IT_{OFF}^{(e)(f)}\right] + P\left(T_{OFF}^{(e)} < \Delta\right)\frac{1}{p}E[T_{ON}]$$

$$+ P\left(T_{OFF}^{(e)} < \Delta\right)\left(\frac{1}{p} - 1\right)E\left[T_{OFF} \mid T_{OFF} < \Delta\right]. \tag{18}$$

In Eq.(18), I and $T_{OFF}^{(e)(f)}$ are not independent. Hence,

$$E\left[IT_{OFF}^{(e)(f)}\right] = \int_0^\Delta x f_{OFF}^{(e)}(x)dx \tag{19}$$

For the probability of failure in the first attempt we have

$$P\left(T_{OFF}^{(e)} < \Delta\right) = \int_0^\Delta f_{OFF}^{(e)}(x)dx \tag{20}$$

Replacing Eq.(20) and Eq.(19) into Eq.(18), we obtain Eq.(15).

$\square$

The system described here has two different service times, depending on whether the arriving customer finds or not other customers in the system. As a result, we cannot simply use the Pollaczek-Khinchin (P-K) formula to derive the queueing delay for our system [11]. Nevertheless, we can still follow the "tagged-user" approach to find this delay.

Theorem 3. *Let an SU access a channel with generic PU activity, such that it experiences two service time distributions S' and S'' with known first and second moments. Then, the total system delay for SU packets is equal to*

$$E[T] = \frac{E[S']}{1 + \lambda\left(E[S'] - E[S'']\right)} + \frac{\lambda E[S''^2]}{2(1 - \lambda E[S''])} + \frac{\lambda\left(E[S'^2] - E[S''^2]\right)}{2 + 2\lambda\left(E[S'] - E[S'']\right)}$$

(21)

Proof. The system delay of an SU packet consists of its queueing delay $E[T_Q]$ and it's service time $E[S]$. We first consider the service time. An arriving packet will find the system busy with probability ρ, which is the utilization of the system, and idle with probability $1 - \rho$. So, the service time can be given as $E[S] = (1 - \rho)E[S'] + \rho E[S'']$. Applying Little's Law on the service part of the system, we get that $\rho = \lambda E[S]$. Substituting this above equation and solving for $E[S]$, we have

$$E[S] = \frac{E[S']}{1 + \lambda\left(E[S'] - E[S'']\right)}.$$

(22)

We now consider the queueing delay (incurred with probability ρ). A packet arriving in the queue finds a packet in service, and it will have to wait for the remaining (i.e. excess) service time S_e of that packet. Assume further that it finds additional N_Q packets in front of it in the queue. Then, the expected queueing time for that packet is

$$E[T_Q] = E[N_Q]E[S''] + E[S_e].$$

Using Little's law $E[N_Q] = \lambda E[T_Q]$ and rearranging we get

$$E[T_Q] = \frac{E[S_e]}{1 - \lambda E[S'']}.$$

(23)

The mean excess time $E[S_e]$ differs from that of an $M/G/1$. We will find it using renewal-reward theory [7]. A renewal starts whenever a packet arrives and finds the system empty. The busy period (B) ends when all the packets that were generated in the meantime get transmitted and the system becomes idle again. The period until a next busy cycle begins is the idle period (I). So, in our case a cycle consists of a busy and an idle period. From renewal-reward theory we know that the mean excess time is equal to the ratio between the mean reward during a cycle and the mean duration of the cycle $E[S_e] = \frac{E[R]}{E[X]}$. The reward is defined as the remaining service during an arrival, similarly to the M/G/1 case (this is illustrated in Fig. 2).

From Fig. 2 we can infer that the excess time is

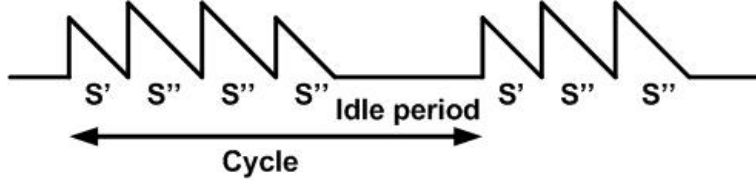$$E[S_e] = \frac{E[\frac{1}{2}S'^2] + ME[\frac{1}{2}S''^2]}{E[B] + E[I]}.$$

(24)

9

Figure 2: Renewal cycle

In Eq.(24), M is the average number of arrivals finding the server busy during a renewal cycle, and is

$$M = \lambda E\left[B\right]. \tag{25}$$

In a long run, the utilization of the system is $\rho = \frac{E[B]}{E[B]+E[I]}$. The average idle period is $E[I] = \frac{1}{\lambda}$. Then, for the average busy period we have

$$E\left[B\right] = \frac{E\left[S'\right]}{1 - \lambda E\left[S''\right]}. \tag{26}$$

Replacing Eq.(26) into Eq.(25) and Eq.(24), as well as Eq.(25) into Eq.(24), we have

$$E[S_e] = \frac{\lambda E[S''^2]}{2} + \frac{\lambda \left(E[S'^2] - E[S''^2]\right)}{2 + 2\lambda \left(E[S'] - E[S'']\right)} \left(1 - \lambda E[S'']\right). \tag{27}$$

Now, replacing Eq.(27) into Eq.(23) gives us

$$E[T_Q] = \frac{\lambda E[S''^2]}{2(1 - \lambda E[S''])} + \frac{\lambda \left(E[S'^2] - E[S''^2]\right)}{2 + 2\lambda \left(E[S'] - E[S'']\right)} \tag{28}$$

Finally, by replacing Eq.(22) and Eq.(28) into $E[T] = E[S] + E[T_Q]$, we obtain Eq.(21). $\qquad\square$

We also need the 2nd moments for S' and S'' in Eq.(21). We will derive them in the following Lemmas.

Lemma 4. *The second moment of non-queued packets $E[S'^2]$ is given by*

$$
\begin{aligned}
E[S'^2] \quad &= \Delta^2 + \frac{P}{1 - p_1 p_2} \left(\Delta^2 p_2 + 2\Delta e^{\lambda\Delta} \left(E[T_1] - \frac{1}{\lambda} \right) + \left(e^{\lambda\Delta} - 1 \right) \left(\left(E[T_1] - \frac{1}{\lambda} \right)^2 + \frac{1}{\lambda^2} + \Delta^2 \right) \right) \\
&+ \quad \frac{P}{1 - p_1 p_2} \left(E[T_{ON}^2] + 2E[T_{ON}] \left(E[T_2] - \frac{1}{\lambda} \right) + \left(\left(E[T_2] - \frac{1}{\lambda} \right)^2 + \frac{1}{\lambda^2} \right) (1 - p_2) \right)
\end{aligned}
\tag{29}
$$

Proof. The proof for the second moment of the non-queued packets follows similar steps as the proof in Lemma 1. First, we express the second moment as

$$S'^2 = I_{OFF}^1 S_{OFF}'^2 + I_{ON}^1 S_{ON}'^2 + I_{OFF}^2 S_{OFF}'^2 + I_{ON}^2 S_{ON}'^2 + \dots, \tag{30}$$

10

where the indicator random variables I_{ON}^i and I_{OFF}^i are identical as in the proof of Lemma 1. We separate the sum in Eq.(30) into two terms

$$Y_{OFF}^2 = \sum_{i=1}^{\infty} I_{OFF}^{(i)} S_{OFF}'^2 \text{ and } Y_{ON}^2 = \sum_{i=1}^{\infty} I_{ON}^{(i)} S_{ON}'^2. \tag{31}$$

The expectation of terms in Y_{OFF}^2 is

$$E[I_{OFF}^{(i)} S_{OFF}'] = \int_{A_i}^{B_i} \Delta^2 \lambda e^{-\lambda \cdot x} dx$$
$$+ \int_{B_i - \Delta}^{B_i} (B_i - x + T_1)^2 \lambda e^{-\lambda \cdot x} dx, \tag{32}$$

where $A_1 = 0$, $A_2 = t_1^{OFF,e} + t_1^{ON}$, $A_3 = t_1^{OFF,e} + t_1^{ON} + t_2^{OFF} + t_2^{ON}, \ldots$ and $B_1 = t_1^{OFF,e}$, $B_2 = t_1^{OFF,e} + t_1^{ON} + t_2^{OFF}, \ldots$, as depicted in Fig. 1. Calculating the integrals in Eq.(32) we get

$$E[I_{OFF}^{(i)} S_{OFF}'] = \Delta^2 e^{-\lambda A_i} + 2\Delta e^{-\lambda(B_i - \Delta)} \left(E[T_1] - \frac{1}{\lambda} \right)$$
$$+ e^{-\lambda B_i} \left(e^{\lambda \Delta} - 1 \right) \left(\left(E[T_1] - \frac{1}{\lambda} \right)^2 + \frac{1}{\lambda^2} + \Delta^2 \right) \tag{33}$$

Summing over all terms of Y_{OFF}^2 in Eq.(31) we have

$$E[Y_{OFF}^2 \mid t_s] = \Delta^2 \sum_i e^{-\lambda A_i} + 2\Delta \sum_i e^{-\lambda(B_i - \Delta)} \left(E[T_1] - \frac{1}{\lambda} \right)$$
$$+ \sum_i e^{-\lambda B_i} \left(e^{\lambda \Delta} - 1 \right) \left(\left(E[T_1] - \frac{1}{\lambda} \right)^2 + \frac{1}{\lambda^2} + \Delta^2 \right) \tag{34}$$

Eq.(34) is the expectation of Y_{OFF}^2 conditioned on the given realization t_s of the sample space. We find Y_{OFF}^2 by taking the expectation of the all possible sample paths

$$E\left[Y_{OFF}^2 \right] = \int_0^{\infty} E\left[Y_{OFF}^2 \mid t_s \right] f_{t_s} \left(x_1^{OFF,e}, x_1^{ON}, \ldots \right) dx_1^{OFF,e} dx_1^{ON} \ldots \tag{35}$$

As before, this integral is then splitted into a product of expectations and after some calculus we get

$$E[Y_{OFF}^2] = \Delta^2$$
$$+ \frac{P}{1 - p_1 p_2} \left(\Delta^2 p_2 + 2\Delta e^{\lambda \Delta} \left(E[T_1] - \frac{1}{\lambda} \right) + \left(e^{\lambda \Delta} - 1 \right) \left(\left(E[T_1] - \frac{1}{\lambda} \right)^2 + \frac{1}{\lambda^2} + \Delta^2 \right) \right) \tag{36}$$

The term Y_{ON}^2 that is related to the packets arriving during an ON period, is derived as follows. The expectation of terms of Y_{ON}^2 in Eq.(31) are

$$E[I_{ON}^{(i)} S_{ON}'^2] = \int_{B_i}^{C_i} (C_i - x + T_2)^2 \lambda e^{-\lambda \cdot x} dx, \tag{37}$$

11

After solving the integral in Eq.(37) we obtain

$$
\begin{aligned}
E[I_{ON}^{(i)} S_{ON}'^2] \quad &= (C_i - B_i)^2 \, e^{-\lambda B_i} + 2 \left(E[T_2] - \frac{1}{\lambda} \right) (C_i - B_i) \, e^{-\lambda B_i} \\
&+ \quad \left(\left(E[T_2] - \frac{1}{\lambda} \right)^2 + \frac{1}{\lambda^2} \right) \left(e^{-\lambda B_i} - e^{-\lambda C_i} \right)
\end{aligned}
\tag{38}
$$

where $C_i \in \left\{ t_1^{ON} + t_1^{OFF,e}, t_1^{ON} + t_1^{OFF,e} + t_2^{OFF} + t_2^{ON}, \ldots \right\}$. C_i are also shown in Fig. 1. Summing over the all Y_{ON}^2 terms in Eq.(31) yields

$$
\begin{aligned}
E[Y_{ON}^2 \mid t_s] \quad &= \sum_i (C_i - B_i)^2 \, e^{-\lambda B_i} + 2 \left(E[T_2] - \frac{1}{\lambda} \right) \sum_i (C_i - B_i) \, e^{-\lambda B_i} \\
&+ \quad \left(\left(E[T_2] - \frac{1}{\lambda} \right)^2 + \frac{1}{\lambda^2} \right) \sum_i \left(e^{-\lambda B_i} - e^{-\lambda C_i} \right)
\end{aligned}
\tag{39}
$$

The expectation over all the sample paths is

$$
E\left[Y_{ON}^2 \right] = \int_0^\infty E\left[Y_{ON}^2 \mid t_s \right] f_{t_s} \left(x_1^{OFF,e}, x_1^{ON}, \ldots \right) dx_1^{OFF,e} dx_1^{ON} \ldots
\tag{40}
$$

As before, after performing some simple calculus steps we get

$$
E[Y_{ON}^2] = \frac{P}{1 - p_1 p_2} \left(E[T_{ON}^2] + 2E[T_{ON}] \left(E[T_2] - \frac{1}{\lambda} \right) + \left(\left(E[T_2] - \frac{1}{\lambda} \right)^2 + \frac{1}{\lambda^2} \right) (1 - p_2) \right)
\tag{41}
$$

Finally by summing Eq.(41) with Eq.(36) we obtain Eq.(29). $\qquad\square$

Lemma 5. *The second moment of queued packets $E[S''^2]$ is given by*

$$
E[S''^2] = \Omega_1 + \Omega_2,
\tag{42}
$$

where

$$
\Omega_1 = \Delta^2 + 2\Delta \int_0^\Delta x f_{OFF}^{(e)}(x) dx + 2\Delta \int_0^\Delta f_{OFF}^{(e)}(x) dx \left(\frac{1}{p} E[T_{ON}] + \left(\frac{1}{p} - 1 \right) E\left[T_{OFF} \mid T_{OFF} < \Delta \right] \right),
$$

and

$$
\begin{aligned}
\Omega_2 \quad &= \int_0^\Delta x^2 f_{OFF}^{(e)}(x) dx + 2 \int_0^\Delta x f_{OFF}^{(e)}(x) dx \left(\frac{1}{p} E[T_{ON}] + \left(\frac{1}{p} - 1 \right) E[T_{OFF} \mid T_{OFF} < \Delta] \right) \\
&+ \quad \int_0^\Delta f_{OFF}^{(e)}(x) dx \left(\frac{2(1-p)}{p^2} E[T_{ON}]^2 + \frac{1}{p} E[T_{ON}^2] + \frac{2}{p} E[T_{ON}] \left(\frac{1}{p} - 1 \right) E[T_{OFF} \mid T_{OFF} < \Delta] \right) \\
&+ \quad \int_0^\Delta f_{OFF}^{(e)}(x) dx \left(\frac{(1-p)(2-p)}{p^2} E[T_{OFF} \mid T_{OFF} < \Delta]^2 + \left(\frac{1}{p} - 1 \right) Var\left(T_{OFF} \mid T_{OFF} < \Delta \right) \right)
\end{aligned}
$$

Proof. The random variable S'', as seen before, is

$$
S'' = \Delta + I \left(T_{OFF}^{(e)(f)} + \sum_{i=1}^{N} T_{ON} + \sum_{i=1}^{N-1} T_{OFF}^{(f)} \right).
\tag{43}
$$

Squaring both sides of previous equation we have

$$S''^2 = \Delta^2 + 2\Delta I T_{OFF}^{(e)(f)} + 2\Delta I \sum_{i=1}^{N} T_{ON} + 2\Delta I \sum_{i=1}^{N-1} T_{OFF}^{(f)}$$
$$+ \quad I \left(T_{OFF}^{(e)(f)} + \sum_{i=1}^{N} T_{ON} + \sum_{i=1}^{N-1} T_{OFF}^{(f)} \right)^2. \tag{44}$$

Taking the expectation of both sides of Eq.(44) we have

$$E\left[S''^2 \right] = \Omega_1 + \Omega_2. \tag{45}$$

In Eq.(45)

$$\Omega_1 = E\left[\Delta^2 + 2\Delta I T_{OFF}^{(e)(f)} + 2\Delta I \sum_{i=1}^{N} T_{ON} + 2\Delta I \sum_{i=1}^{N-1} T_{OFF}^{(f)} \right],$$

which is equivalent to

$$\Omega_1 = \Delta^2 + 2\Delta E[I T_{OFF}^{(e)(f)}] + 2\Delta E\left[I \sum_{i=1}^{N} T_{ON} \right] + 2\Delta E\left[I \sum_{i=1}^{N-1} T_{OFF}^{(f)} \right]$$

Beside the random variables I and $T_{OFF}^{(e)(f)}$ which are not independent, all the other random variables are independent. Hence,

$$\Omega_1 = \Delta^2 + 2\Delta E\left[I T_{OFF}^{(e)(f)} \right] + 2\Delta E[I] E\left[\sum_{i=1}^{N} T_{ON} \right] + 2\Delta E[I] \left[\sum_{i=1}^{N-1} T_{OFF}^{(f)} \right] \tag{46}$$

In the last equation $E\left[I T_{OFF}^{(e)(f)} \right] = \int_0^{\Delta} x f_{OFF}^{(e)}(x) dx$. The expectation of the random variable I is $E[I] = P\left(T_{OFF}^{(e)} < \Delta \right) = \int_0^{\Delta} f_{OFF}^{(e)}(x) dx$. By using the two previous expressions and the Wald's equality in Eq.(46) we get

$$\Omega_1 = \Delta^2 + 2\Delta \int_0^{\Delta} x f_{OFF}^{(e)}(x) dx + 2\Delta \int_0^{\Delta} f_{OFF}^{(e)}(x) dx \frac{1}{p} E[T_{ON}]$$
$$+ \quad 2\Delta \int_0^{\Delta} f_{OFF}^{(e)}(x) dx \cdot \left(\frac{1}{p} - 1 \right) E\left[T_{OFF} \mid T_{OFF} < \Delta \right]. \tag{47}$$

The other part of Eq.(45) is given by

$$\Omega_2 = E\left[I \left(T_{OFF}^{(e)(f)} + \sum_{i=1}^{N} T_{ON} + \sum_{i=1}^{N-1} \right)^2 \right]. \tag{48}$$

13

After taking the square of the right-side of Eq.(48) we have

$$
\begin{aligned}
\Omega_2 \;&= E\left[IT_{OFF}^{(e)(f)2}\right] + E[I]E\left[\left(\sum_{i=1}^{N} T_{ON}\right)^2\right] + E[I]E\left[\left(\sum_{i=1}^{N-1} T_{OFF}^{(f)}\right)^2\right]\\
&+\; 2\left[IT_{OFF}^{(e)(f)}\right] E\left[\sum_{i=1}^{N-1} T_{OFF}^{(f)}\right] + 2E[I]E\left[\sum_{i=1}^{N} T_{ON}\right] E\left[\sum_{i=1}^{N-1} T_{OFF}^{(f)}\right]
\end{aligned}
\tag{49}
$$

In Eq.(49)

$$
E\left[IT_{OFF}^{(e)(f)2}\right] = \int_0^{\Delta} x^2 f_{OFF}^{(e)}(x)dx
$$

Using Wald's equation in Eq.(49), similarly as in Eq.(46) we have

$$
\begin{aligned}
\Omega_2 \;&= \int_0^{\Delta} x^2 f_{OFF}^{(e)}(x)dx + 2\int_0^{\Delta} x f_{OFF}^{(e)}(x)dx \left(\frac{1}{p}E[T_{ON}] + \left(\frac{1}{p}-1\right)E[T_{OFF} \mid T_{OFF} < \Delta]\right)\\
&+\; \int_0^{\Delta} f_{OFF}^{(e)}(x)dx \left(\frac{2(1-p)}{p^2}E[T_{ON}]^2 + \frac{1}{p}E[T_{ON}^2] + \frac{2}{p}E[T_{ON}]\left(\frac{1}{p}-1\right)E[T_{OFF} \mid T_{OFF} < \Delta]\right)\\
&+\; \int_0^{\Delta} f_{OFF}^{(e)}(x)dx \left(\frac{(1-p)(2-p)}{p^2}E[T_{OFF} \mid T_{OFF} < \Delta]^2 + \left(\frac{1}{p}-1\right)Var\left(T_{OFF} \mid T_{OFF} < \Delta\right)\right)
\end{aligned}
$$

Finally, summing Eq.(50) and Eq.(47), we get Eq.(42). $\qquad\square$

Our analytical results suggest that the exact SU performance has an intricate dependence on PU characteristics that goes beyond channel utilization. At the same time, the key additional statistics needed are the second (and in congested cases) the third moments of PU active and idle periods (the success probabilities could be approximated using second moments through Chebyshev's inequality). This implies that by collecting such statistics for different channels, an SU can use our result to evaluate each channel's predicted performance, and choose according to application needs. In the next section, we validate our analysis, and also explore the effect of active and idle period variability on performance.

3 Performance analysis

The traditional metric for characterizing the PU activity is the duty cycle. It is defined as $\frac{E[T_{ON}]}{E[T_{ON}]+E[T_{OFF}]}$. Unless otherwise stated, in all the scenarios below the packet size is taken to be 0.25, although other values of Δ lead to the same conclusions. To validate our theory against simulations we take combinations of exponentially (Exp) and Bounded Pareto (BP) distributed ON-OFF periods (as an example of "heavy-tailed" distributions). For the Bounded Pareto distribution, we take the lower bound $L = 0.215$, upper bound $H = 400$, and the shape parameter $\alpha = 1.2$.

Fig. 3 shows the average packet delay in a cognitive network for two different primary user activity scenarios for exponentially distributed OFF periods. The arrival rate is $\lambda = 0.1$. For the exp-exp distributions a low primary user activity (duty cycle of 0.2) gives a utilization of 0.05. When the duty cycle is 0.8, the utilization is 0.29. For the BP ON periods, lower relative primary user activity of
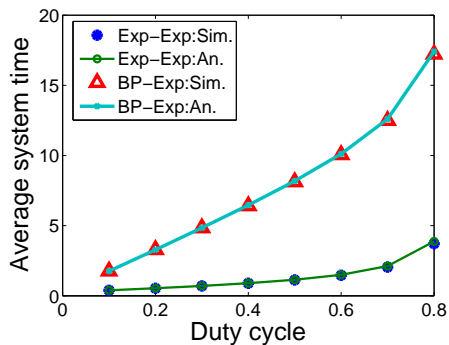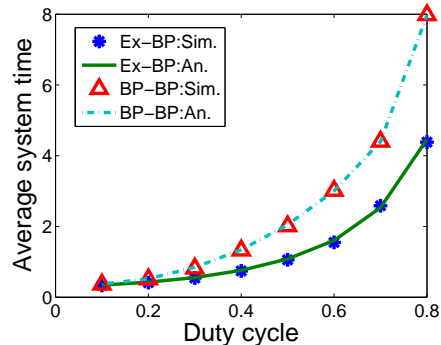
Figure 3: System time for exp-OFF   Figure 4: System time for BP-OFF

0.2 corresponds to a utilization of 0.09, and higher relative licensed user activity of 0.8 to a server utilization of 0.41. Different values of duty cycle give different levels of utilization, since the mean service time depends on the values of $E[T_{ON}]$ and $E[T_{OFF}]$. i.e. of duty cycle. The first thing to observe is a good match between theory and simulations. Furthermore, we can also observe that higher duty cycle implies higher delays, which is expected because the higher the duty cycle is, the primary user is more active, and there is less time for the cognitive user to operate. We can also see that for the same average ON and OFF durations (the same duty cycle), the delays are higher when the primary user has busy periods with higher variablility. This is the first interesting conclusion that comes out of our model. Despite the two channels looking similar, from the point of view of average PU activity, variability can further affect delays. This is reminiscent of the inspection paradox [7], albeit the dynamics of Equations (1) and (15) are in fact more complex.

Fig. 4 shows the packet delays for Bounded Pareto distributed OFF periods. The arrival rate is low (0.01). This arrival rate corresponds to sparse traffic. As we can observe from Fig. 4, there is a good match between theory and simulations for the generic distributed OFF periods, also.

So far, we have considered some standard distributions for the general ON and OFF periods, to see the effect of high or low variance. We are also interested to see how our model can predict performance under more "realistic" PU patterns. To this end, we consider two recently proposed models for PU activity, one for cellular channels [5] and one for WiFi channels [6]. We have tried to implement the proposed models according to the respective descriptions, although some details are not specified there.

Fig. 5 shows the packet delay incurred by a WLAN network. The distributions for this simulation are taken from [6]. The ON periods are deterministic, while OFF periods have bimodal distributions. The arrival rate is 0.01. As we can see our theory provides a very good match with simulations.

Fig. 6 shows the packet delay in a cellular network. The model description from
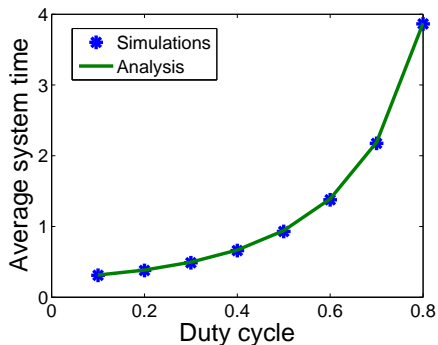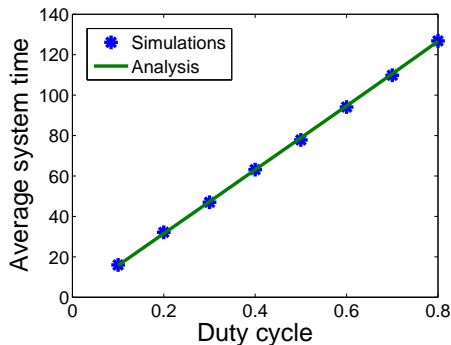
15

Figure 5: WLAN primary user

Figure 6: Cellular primary user

[5] is being used. The ON periods underly a multimodal distribution, while the OFF periods are exponentially distributed. The packet size is 0.01. The arrival rate is $\lambda = 0.1$. For a duty cycle of 0.8, this arrival rate corresponds to an (maximum) utilization of 0.74. More details about the two models can be found in [5] and [6].

We have established so far that (i) our analytical model correctly predicts performance in all generic PU channels considered, and under various levels of congestion (we note here that we have performed a large number of other scenarios, with similar conclusions), and (ii) that even if two channels have similar *average* PU activity, variability (e.g. PU traffic burstiness) can further degrade performance. We now go a step further and consider a channel A with high average PU activity (duty cycle of 0.6) and exponentially distributed ON (activity) durations with variance equal to 1. We put it "against" a channel B with much lower PU activity (duty cycle 0.3), but lognormally distributed ON periods (which have a heavier tail than exponential). Keeping the mean of the ON period unchanged, and increasing the variance of the lognormal distribution gives us an interesting insight into the effect of both PU average activity and variability on cognitive user performance.

Table 2 displays the ratio of $\dfrac{\text{SU delay on channel B}}{\text{SU delay on channel A}}$.

Table 2: The ratio of delays for two different channels

Variance for CH B	1	10	20	30	40	50	60
Ratio of Delays	0.5	1.2	1.8	2.3	2.6	3	3.2

We can observe from Table 2 that for similar variance, channel B which is less "busy" is better. However, by increasing the variance, the ratio keeps growing and exceeds 3 for a variance higher than 50. In fact, in theory, this difference can become arbitrarily high (i.e. for real heavy-tailed distributions, like Pareto with parameter < 2). We can thus conclude that, contrary to common belief that tends to consider an underutilized part of the spectrum as "good spectrum" for cognitive users, the impact of variability of the primary user actitivty is much

more important than the duty cycle itself. This could be key for example in delay-sensitive, but low throughput applications like M2M.

4   Related work

Some interesting works also model the PU user with a stochastic ON-OFF process, but assume a 2-state Markov Chain for it [12, 13]. Service times are derived from their 2-state Markov chain, and system times using an $M/G^Y/1$ system with bulk departures. While the exponential assumption is more convenient for analysis, it turns out to be inaccurate for both cellular [14], and WiFi systems [6]. As our analysis suggests, it can also lead to (arbitrarily) inaccurate predictions. In a recent work [15], a simple approximate model is used for delay prediction of an SU packet, assuming generic OFF periods, but this model also suffers from large inaccuracies, when the OFF periods are not exponential.

In order to depart from the strong exponential assumption, some recent works [8, 16, 17] have capitalized on the measurement-based study of [13], in which the Poisson approximation seems to be decent for call arrivals, but call duration is generically distributed. These works model SUs together with PUs, as an M/G/1 system with priorities and preemption. M/G/1 systems with priorities have been long analyzed (see e.g. [18]). Nevertheless, there are some important caveats in the above models. *First*, the system is preemptive-resume, that is, SU packets when preempted by a PU transmission can resume transmission from the point they stopped. In practice, SU packets will "collide" when a PU is detected, and have to *restart* in the next available cycle (possibly colliding again). Hence, this model is only approximately accurate for small packets (or long OFF periods). *Second*, while PU call arrivals might be approximated by a Poisson distribution, this *does not* mean that OFF periods are exponentially distributed, nor that the SU can directly infer the ON and OFF duration distributions. These depend on a number of system specifics and protocol interactions across the stack.

Consequently, we believe our model allows one to capture a much larger number of channels and PU activity patterns, and the SU to be able to measure and evaluate the predicted performance of a channel, without the need to know anything about the type and number of PUs multiplexed on a channel, protocols used, and PU job statisitics.

5   Conclusion

In this paper, we have proposed a queueing analytical model for the performance of cognitive users under generic ON-OFF primary channel models, and we have validated it against both synthetic and realistic PU channel models. We have shown that variability of primary user activity is very important, and often more important than the utilization itself. This is the key for the protocol design. The

actual delay is a complex interplay between secondary traffic characteristics (intensity and packet sizes) and channel characteristics (1st and 2nd moments of idle and available durations of PU). In future work, we intend to extend our model to multihop networks, and also use our results to design better spectrum management, resource allocation and scheduling algorithms.

References

[1] B. Wang and K. Liu, "Advances in cognitive radio networks: A survey," *IEEE J. Sel. Topics Signal Process.*, 2011.

[2] "AT&T: Improving 3G network," http://gigaom.com/2008/06/08/3g-network-iphone/.

[3] P. Bahl, R. Chandra, T. Moscibroda, R. Murty, and M. Welsh, "White space networking with Wi-Fi like connectivity," in *Proc. of ACM SIGCOMM*, 2009.

[4] R. Lu, X. Li, X. Liang, X. Shen, and X. Lin, "GRS: The green, reliability, and security of emerging machine to machine communications," *IEEE Comm. Mag*, vol. 49, no. 4, pp. 28–35, April 2011.

[5] D. Willkomm, S. Machiraju, J. Bolot, and A. Wolisz, "Primary user behavior in cellular networks and implications for dynamic spectrum access," *IEEE Comm. Mag.*, vol. 47, no. 3, pp. 88–95, mar 2009.

[6] S. Geirhofer and L. Tong, "Dynamic spectrum access in the time domain: Modeling and exploiting white space," *IEEE Comm. Mag.*, vol. 45, pp. 66–72, 2007.

[7] S. M. Ross, *Stochastic Processes*, 2nd ed. John Wiley & Sons, 1996.

[8] X. W. C. Zhang and J. Li, "Cooperative cognitive radio with priority queueing analysis," in *Proc. of IEEE International Conference on Communications*, Dresden, Germany, 2009, pp. 1–5.

[9] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks:A POMDP framework," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, 2007.

[10] H. Kim and K. Shin, "Fast discovery of spectrum opportunities in cognitive radio networks," in *Proc. of IEEE DySPAN*, Oct 2008.

[11] L. Kleinrock, *Queueing theory, Volume I: Theory*. John Wiley & Sons, 1975.

[12] H. Su and X. Zhang, "Cross-layer based opportunistic MAC protocols for QoS pro-visionings over cognitive radio wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 118–129, Jan. 2008.

[13] L. T. S. Geirhofer and B. Sadler, "Cognitive medium access: Constraining interference based on experimental models," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 95–105, Jan. 2008.

[14] J. B. D. Willkomm, S. Machiraju and A. Wolisz, "Primary users in cellular networks: A large-scale measurement study," in *Proc. of IEEE DySPAN*, 2008, pp. 1–11.

[15] A. Lertsinsrubtavee, N. Malouch, and S. Fdida, "Controlling spectrum hand-off with a delay requirement in cognitive radio networks," in *Proc. of ICCCN*, Aug. 2012.

[16] I. Suliman and J. Lehtomaki, "Queueing analysis of opportunistic access in cognitive radios," in *Proc. of CogART*, 2009.

[17] T. Q. D. T. Hung and H.-J. Zepernick, "Average waiting time of packets with different priorities in cognitive radio networks," in *Proc. of IEEE ISWPC*, Modena, Italy, May 2009.

[18] L. Kleinrock, *Queueing theory, Volume II: Computer Applications*. John Wiley & and Sons, 1976.