Azarfar *et al.* EURASIP Journal on Wireless Communications and Networking 2014, **2014**:206
http://jwcn.eurasipjournals.com/content/2014/1/206

EURASIP Journal on Wireless Communications and Networking
a SpringerOpen Journal

RESEARCH Open Access

Priority queueing models for cognitive radio networks with traffic differentiation

Arash Azarfar*, Jean-François Frigon and Brunilde Sansò

Abstract

In this paper, we present a new queueing model providing the accurate average system time for packets transmitted over a cognitive radio (CR) link for multiple traffic classes with the preemptive and non-preemptive priority service disciplines. The analysis considers general packet service time, general distributions for the channel availability periods and service interruption periods, and a service-resume transmission. We further introduce and analyze two novel priority service disciplines for opportunistic spectrum access (OSA) networks which take advantage of interruptions to preempt low priority traffic at a low cost. Analytical results, in addition to simulation results to validate their accuracy, are also provided and used to illustrate the impact of different OSA network parameters on the average system time. We particularly show that, for the same average CR transmission link availability, the packet system time significantly increases in a semi-static network with long operating and interruption periods compared to an OSA network with fast alternating operating and interruption periods. We also present results indicating that, due to the presence of interruptions, priority queueing service disciplines provide a greater differentiated service in OSA networks than in traditional networks. The analytical tools presented in this paper are general and can be used to analyze the traffic metrics of most OSA networks carrying multiple classes of traffic with priority queueing service differentiation.

Keywords: Cognitive radio; Dynamic spectrum access; Queueing theory; Priority queueing; Differentiation

1 Introduction

Opportunistic spectrum access (OSA) is considered an important technology to address current and predicted exponential traffic growth in wireless networks [1-4]. Such growth is predominantly driven by multimedia traffic, such as video streaming [4]. Thus, it is expected that OSA networks will carry several traffic classes with different quality of service (QoS) requirements and importance.

The research objective of this paper is to obtain analytical tools to analyze traffic metrics, such as the packet system time, for differentiated services in opportunistic spectrum access networks. Such tools are required to evaluate the packet-level impact of OSA network parameters, novel medium access control (MAC) algorithms, channel sensing order strategies, etc. Moreover, those analytical tools can be used as a decision-making process for multimedia MAC algorithms [5], for OSA networks employing cognitive radio (CR) nodes, which possess learning and decision-making capabilities.

1.1 Related work

Queueing models are the preferred approach to derive analytical results to analyze traffic metrics, and in [6], the authors argue that queueing models can be the best choice to analyze delay in a cognitive radio network. Priority service disciplines, such as the preemptive and non-preemptive service disciplines, are the most common approaches to implement service differentiation in communication networks. Furthermore, in an OSA network, the CR users must stop transmitting on an operating channel if the channel's primary user (PU) is detected or if the channel quality is unacceptable due, for example, to deep fading or interference. In the queueing model, the operating channel is the server of the queue. To achieve our objective, we must therefore analyze queueing models with priority service disciplines in the presence of frequent queue server interruptions.

Queueing models with preemptive priority service discipline and interruptions have been previously studied

*Correspondence: arash.azarfar@polymtl.ca
Department of Electrical Engineering, École Polytechnique de Montréal, C.P. 6079, succ. Centre-Ville, Montréal, QC H3C 3A7, Canada

Springer

© 2014 Azarfar et al.; licensee Springer. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited.

[7-11]. Some of that work considered that interruption periods are server busy periods generated by higher priority classes of traffic. For this approach, the interruption periods are not generally distributed since they depend on the arrival rate and service rate of higher priority classes. Obtaining the interruption period distribution is therefore not always straightforward. Inversely, given an interruption period distribution, it is not easy to find the appropriate arrival and service processes whose busy period has this distribution. The other articles that studied generally distributed interruption periods only considered a single traffic class and only provided bounds.

Few have attempted to provide queueing models for opportunistic spectrum access networks. In [6,12-14], queueing models for an OSA network with a single class of traffic were derived using a similar approach as in [7-11], whereby the server interruption periods for the cognitive radio users are busy periods generated by the preemptive primary traffic. This approach has several major deficiencies. First, those models are limited to exponential operating period length. Also, as previously discussed, they cannot address arbitrary interruption period lengths before the transmission can resume. This is particularly important for OSA networks since the interruption period length depends on several factors such as the MAC policy, the number of available channels, the number of competing CR users, etc. Even for the simple case where CR users wait until the PU releases the channel, the interruption period is not necessarily distributed as the busy period of a PU Poisson traffic [15]. The approach of simply considering the interruption periods as a preemption from PU traffic is therefore not accurate and general enough to analyze OSA networks.

In [16], we addressed several of those problems in a new queueing model for a single class of CR traffic for general operating and interruption period lengths. However, this model was limited to constant service time. In [17], an optimal threshold for the queue length to decide whether a packet should join the queue or not is derived. However, the model is again not general and cannot be used to analyze traffic metrics.

To the best of our knowledge, [18] is one of the few papers discussing a queueing system with multiple classes of traffic in cognitive radio networks. The authors analyze a T-preemptive scheme and, similarly to the other work on opportunistic spectrum access networks, the queueing analysis does not consider general interruption lengths, and it is specific to the priority service disciplines considered. In [19], the authors consider a queueing model with multiple classes and a single priority service discipline for cognitive radios to address the problem of channel selection.

It is also important to note that there have been efforts to analyze a cognitive radio network from other perspectives using other performance metrics such as energy consumption [20] and throughput [21]. As the focus of this paper is limited to the cognitive radio network delay analysis, those other works are not discussed here.

1.2 Contributions

In this paper, we consider real server interruptions distinct from the service times for a high priority class as well as general service time. We thus present, to the best of our knowledge, the first queueing model providing the accurate average system time for a Poisson packet arrival process with general service time transmitted over a CR link with general interruption periods and exponentially distributed operating periods for both a single traffic class and for multiple traffic classes with the preemptive and non-preemptive priority service disciplines. We also derive an approximate analysis for general operating period distributions. We further introduce two novel priority service disciplines which are specific to OSA networks with service interruptions. In the first novel OSA service discipline that we name *exceptional non-preemptive*, the service is in general non-preemptive except for low priority arrivals in an empty queue during an interruption period, which can be preempted by high priority packet arrivals during the same interruption period. In the second novel OSA service discipline that we name *preemptive in case of failure*, the service is non-preemptive during the operating periods but high priority packets can preempt low priority packets at the end of an interruption period. We provide an accurate analysis for the first novel OSA priority service discipline while approximate results are provided for the second (exact results are derived for the preemptive in case of failure service discipline for exponential service times). To limit the analysis complexity and presentation, in the preemptive models, we only study a two-class traffic model; however, we provide guidelines on extension of those results to more general scenarios. Since no specific assumptions are made regarding the nature of the operating and interruption periods, the results and derivations presented in this paper can be used to analyze the traffic metrics of most OSA networks with different MAC protocols. The final contributions of this paper are new insights on OSA networks based on the average system time analysis. Particularly, we show that, for the same average CR transmission link availability (ratio between average channel availability period length and average interruption period), the packet system time significantly increases as the operating and interruption periods average length exceeds the packet service time. We also present results emphasizing the critical importance of minimizing the interruption period lengths to minimize the packet system time in OSA networks. Another conclusion that we present is that

priority queueing service disciplines provide a greater differentiated service in OSA networks than in traditional networks.

The reminder of the paper is organized as follows. In Section 2, we present the cognitive radio system and the queue model. In Section 3, an M/G/1 queue with interruptions and with a single class of CR traffic is discussed. The results are then used in Section 4 to analytically solve four priority queueing disciplines in the presence of interruptions. We also present in Section 5 an alternative approach to analyze the preemptive and non-preemptive disciplines for exponential operating periods. Analytical and simulation results are presented in Section 6, and finally, Section 7 concludes the paper with some remarks on future research directions. All the parameters and notations used throughout the paper are summarized in Table 1.

Table 1 Notations

Notation	Description
Y	Length of operating periods (RV)
R	Length of recovery (interruption) periods (RV)
C	$C = Y + R$
λ	Packet arrival rate
A	packet inter-arrival time (RV)
X	Completion time (RV)
X^*	Completion time in alternative model (RV) (Section 3.4)
Index $_b$	For packets entered a busy system
Index $_e$	For packets entered an empty system
Index $_a$	For packets entered an empty-available system
Index $_u$	For packets whose service started at the beginning of a Y
B	HP busy periods (RV)
B_b	Busy period started with X_b
B_Z	Busy period started with $Z + X_b$
T	Real service time (RV)
J	Remaining completion time of the packet in service
$\hat{Z}(s)$	LST of a continuous random variable Z
$f_Z(t)$	PDF of a random variable Z
$F_Z(t)$	CDF of a random variable Z
$m(t)$	Average number of renewals until time t
$m_{a\|b\|u}(t)$	$m(t)$ for packets of type a, b, or u
$m^2(t)$	Second moment of the number of renewals until time t
ρ	$\lambda E[X]$
ρ_b	$\lambda E[X_b]$
P_{ae}	Probability of system being available when empty (arrival point)
W	Waiting time in the queue
W^*	Waiting time for the alternative model
D	Total time spent in the system
$P_{ai\|C\|Y\|R}$	Probability of HP arrival in a cycle (C), Y, or R
K	Local parameter to count the number of an event
$i = 1, 2$	Subindex represents the traffic class
α	Exponentially distributed parameter when $F_Y = 1 - e^{-\alpha t}$
β	Exponentially distributed parameter when $F_R = 1 - e^{-\beta t}$
$\tilde{F}_Z()$.	Min. of random variable Z and an exponentially distributed (Eq. 4)
γ	Exponentially distributed parameter when $F_T = 1 - e^{-\gamma t}$
N_p	Number of priority classes
T_v	Service time of virtual packets which form the interruptions
$P_{A<C}$	Probability of arrival in C=$Pr(A < C)$
R_r	Remaining of R after an arrival in R

2 Cognitive radio queue model

The cognitive radio queue model can be summarized as follows. We consider a pair of cognitive radio users operating using opportunistic spectrum access over one or more wireless channels. During an operating period, packets that are in one of the CR nodes queue are transmitted to the other CR node according to a chosen service discipline. As illustrated in Figure 1, the CR nodes opportunistically operate over a channel for a random duration Y until the channel becomes unavailable. When the channel becomes unavailable, the packet transmission is interrupted for a random length R until an available operating channel can be used by the CR pair, at which time the packet transmission is resumed.

We now describe the details of the model. The OSA network assigns a channel to the pair of CR users according to its MAC protocol and channel assignment algorithm. Transmission over the assigned channel can be multiplexed with other CR users, the only assumption for our model is that during channel availability periods, the pair of CR users have access to a constant service rate over the assigned channel (the packet length defines the distribution of the real service time). If, as in IEEE 802.22, periodic quiet periods are required to sense the channel or perform other OSA network tasks, the channel service rate can be scaled accordingly. Channels are assumed to be homogeneous with the same service rate. The CR nodes communicate over the assigned channel for a random operating duration Y until the channel becomes unavailable and packet transmissions must be stopped. We denote the instant where the channel becomes unavailable for operation as a failure event [22]. To illustrate the generality of this model, we now give a few examples of failure events. A failure can be due to the appearance of the primary user, a false detection of the primary user, a link failure due to excessive transmission loss (fading, shadowing, or

Table 1 Notations *(Continued)*

C_r	Remaining of C after an arrival in C
$Z_1^{<Z_2}$	$Z_1 \mid (Z_1 < Z_2)$ (for two RVs)
S	Initial setup time (RV) (Section 3.5)
P_0	Probability of system being empty
$P_{L\mid NH}$	Probability of system being empty of HP but not LP

distance), or interference from other secondary users. A failure event can also be due to the OSA protocol. For example, CR users might have to release a channel after a fixed period of time, even if no primary user appears. Note that it is also implicitly assumed that no miss-detection may occur, so there will be no performance degradation for primary users to be analyzed. For the queueing model analysis, we are only looking at transmission operation from a marked user point of view, so the interactions of different users (cooperation or competition) is not explicitly considered in the model. In other words, we have no assumption on the interaction of the users. Indeed, the type of interaction affects the distributions of Y and R, and the same model that we are using will be applicable.

For the model and its analysis, only the distribution of Y is required and the exact reason for the failure event is irrelevant as long as it is independent of the packet transmission process (e.g., the pair of CR users are not reassigned to a new channel after each packet transmission or when the packet queue is empty). Note that when the CR users start using a channel, unless it is immediately after a channel unavailability period, they generally have no knowledge about how long this channel has been available. Therefore, Y is a function of the residual time of the availability period of the channel [16].

The recovery or interruption period denotes the period of time R during which the CR users cannot transmit and try to recover the transmission [22]. The length of R depends on the OSA network model, but only its distribution is relevant for the queue analysis. We will use a few examples of recovery periods to demonstrate the generality of the proposed model. For OSA MAC protocols in which the CR users buffer the packets until the operating channel becomes available again [23], the distribution of R is identical to the channel unavailability period distribution. For network with a channel switching policy in which when the channel becomes unavailable, the CR users enter a competition with other CR users to be granted access to a new channel [23], the distribution of R will depend on the MAC competition protocol (e.g., slotted Aloha), the number of users, the number of available channels, etc. Even if the user is blocked due to other users transmitting on all channels, the total time of blocking until a successful channel reservation is included in the recovery time. For OSA networks where a channel is granted by a spectrum server, the length of R can be a fixed period of time (query and service time, radio switching time, etc.). A methodology to find the recovery period distribution for two baseline multichannel opportunistic spectrum access MAC protocols is provided in [24].

To summarize, determining the distribution of Y and R according to the OSA network model under study is outside the scope of this paper. But once the distributions are known, the queue model that we are presenting can be used to find the traffic metrics for the OSA network CR users.

As illustrated in Figure 2, we consider a CR system with N traffic classes where each class i, $i = 1, \ldots, N$, has an independent Poisson packet arrival process with rate λ_i and the total arrival rate is $\lambda = \sum \lambda_i$. We also denote by A_i the inter-arrival time between packets of class i, $i = 1, \ldots, N$, and define $\mathbb{A} = \min\{A_1, ..., A_N\}$ as the inter-arrival time between packets in the system. Throughout the paper, for any random variable Z, $f_Z(.)$ and $F_Z(.)$ respectively represent the probability density/mass function (PDF or PMF) and the cumulative distribution function (CDF) of the random variable Z. Moreover, $\widehat{Z}(s)$ represents the Laplaceñ-Stieltjes transform (LST) of the distribution $F_Z(.)$ of the random variable Z.

Lower index classes have higher priority. In the special case of two traffic classes (e.g., voice and data), we designate the index 1 traffic as high priority (HP) and the index 2 traffic as low priority (LP). Packets from traffic class i have a random *real* service time T_i. The real service time is the total transmission time of the packet and excludes the time spent during interruption periods during the service of a packet. From the queueing point of view, the user's operation can thus be modeled as an M/G/1 queue with random service interruptions.

We must also introduce the notion of *completion time* X, which represents the whole time in service for a packet including the real service time T and the interruptions that may occur during its service. We

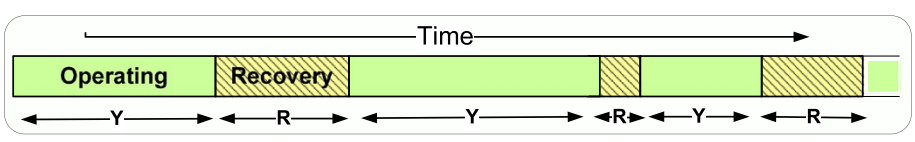

Figure 1 Operation model for a cognitive radio link alternating between operating and recovery (interruption) periods. Identical instances of Y and R are illustrated.

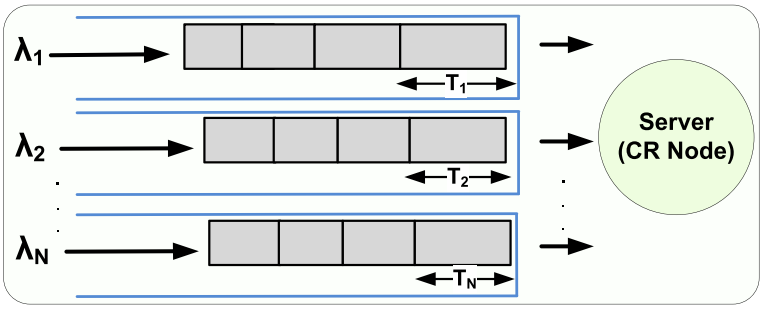

Figure 2 Queue model with a multiple-class cognitive radio traffic.

assume a service-resume model which means that after a packet service interruption, only the remaining part of the packet needs to be transmitted. This implies that the completion time of a packet is formed by alternating instances of Y and R named $Y_1, Y_2, \ldots$ and $R_1, R_2, \ldots$ respectively. It is worth noting that in this work, the term *packet* designates a general block of data to be transmitted, which in itself can be composed of smaller disjoint parts. For instance, the block of data could be a MAC frame composed of several higher OSI layers' data blocks or a jumbo frame composed of smaller wireless or Ethernet frames, as in IEEE 802.11n with MAC service data unit (MSDU) or MAC protocol data unit (MPDU) aggregation [25]. The service-resume model is thus realistic for those scenarios where the packet of data to be transmitted consists of several smaller frames because, in case of an interruption, there is no need to retransmit the parts or frames which have already successfully been transmitted. The queue size is assumed infinite, so packet loss and blocking are irrelevant and the main performance metrics are the total time spent in the queue (waiting time) W and in the system (system time or sojourn time) $D = X + W$.

We consider four different service disciplines: the classical *non-preemptive* and *preemptive-resume* schemes [26], and two novel disciplines we propose in this paper. As illustrated in Figure 3, if during a recovery period a low priority packet arrives in an empty system followed by a high priority (HP) packet, in a non-preemptive scheme the LP packet will be transmitted first. In other words, the LP packet cannot be preempted even if its real service has not started yet.

In the new scheme that we call *exceptional non-preemptive*, an HP packet can preempt a lower priority packet only if its real service has not started yet. As can be seen in Figure 3, the difference between a non-preemptive and an exceptional non-preemptive scheme is only for the LP packets which arrive to an empty unavailable system. We also propose a *preemption in case of failure* discipline where HP packets cannot preempt an LP packet in service until the LP service is finished or if an interruption occurs. In other words, at the end of a recovery period, the priority is always given to HP packets. Meanwhile, in the classical preemptive scheme HP packets can preempt LP packets at any time. The two proposed schemes are defined based on the existence of interruptions and can specifically be used in OSA networks with service interruptions.

3 Single traffic class analysis

In this section, we analyze the queue model with a single traffic class and obtain results which will be extended in the next section to the analysis of multiple classes of

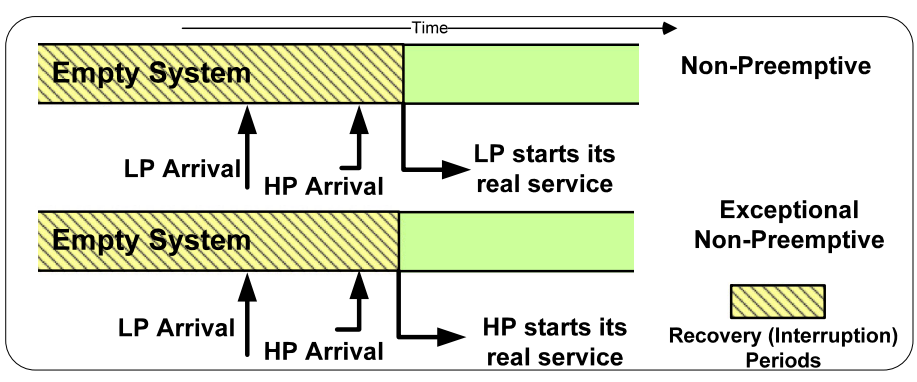

Figure 3 Comparison between non-preemptive and exceptional non-preemptive schemes.

traffic. The results presented here are an extension of the work presented in [16] where only constant service time was considered. In this section, the analysis is done for the general case where the packet length follows an arbitrary random distribution.

As can be seen in Figures 4 and 5, we can distinguish three types of packet in the CR queueing model.

There are packets which enter an empty available system, and their real service starts immediately (see Figure 4(1)). The subscript 'a' is used to designate this case. Considering the completion time of the packets of type 'a', we can see that the distribution of the first operating period is different from the following ones because it represents the residual part of Y. We thus use the notation of Y_{1a} to designate the first operating period for the case 'a' packets.

Packets that enter an empty system during an interruption period (empty unavailable system) must wait until the end of the recovery period before starting their real service (see Figure 4(2)). The subscript 'u' is used to designate this case. For those packets, the distribution of the first operating period is the same as the operating period distribution and is simply denoted by Y_1. On the other hand, the distribution of the first recovery period is different from the following recovery periods because it represents the residual part of R. R_r is used to denote the remaining part of the recovery period in which the arrival has occurred. Note that we consider the arrival time as the start of the completion time in case 'u' (i.e., R_r) is not accounted as waiting time but as completion time.

Finally, there are packets that enter a busy system and are queued (see Figure 5). Their service starts immediately after the completion time of the previous packets, and the subscript 'b' is used to designate this case. Note that in this case, the completion time of the packet is always started within an operating period (similar to the case 'a'). Their first operating period is thus called Y_{1b} because its distribution is different from general Y.

In the following, we will find the first two moments of the completion time X and then analyze the average waiting time and other metrics. We also provide simplifications for the special case of exponential operating periods.

3.1 Completion time

Suppose X_a represents the completion time of the packets of type 'a' (similarly X_b and X_u for cases 'b' and 'u'). Let also X_e define the completion time of a packet which arrives to an empty system (cases 'a' and 'u' together). The first two moments of X_e are given as follows:

$$E[X_e] = P_{ae}E[X_a] + (1 - P_{ae})E[X_u], \quad (1)$$

$$E\left[X_e^2\right] = P_{ae}E[X_a^2] + (1 - P_{ae})E\left[X_u^2\right], \quad (2)$$

where P_{ae} is the average probability that the server is available when the system is empty. From [7], P_{ae} can be found equal to the following:

$$P_{ae} = 1 - \frac{(1 - \widetilde{F}_Y(\lambda))(1 - \widetilde{F}_R(\lambda))}{\lambda E[Y]\,(1 - \widetilde{F}_Y(\lambda)\widetilde{F}_R(\lambda))}, \quad (3)$$

where $\widetilde{F}_Z(.)$, for an arbitrary distribution function $F_Z(.)$, is given as follows:

$$\widetilde{F}_Z(\lambda) = \int_0^\infty e^{-\lambda t} dF_Z(t). \quad (4)$$

For $Z = Y$ or R, this function gives the probability that the length of the operating or recovery period, respectively, is less than a packet inter-arrival time. Note that P_{ae} is a conditional probability, conditioned on the fact that the system is empty. In general, P_{ae} is only a function of the moments of Y and R. We therefore assume in the following that there is no correlation between P_{ae} and Y and R.

The first moment of X can be obtained by solving the following equation:

$$E[X] = \rho E[X_b] + (1 - \rho)E[X_e], \quad (5)$$

where $\rho = \lambda E[X]$ is the probability of the system being not empty. The second moment of X can be found equal to:

$$E[X^2] = \rho E\left[X_b^2\right] + (1 - \rho)E\left[X_e^2\right]. \quad (6)$$

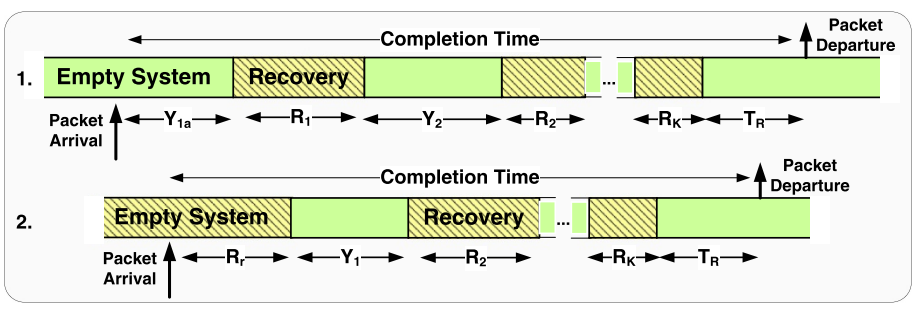

Figure 4 Completion time for the case 'a' (1.) and case 'u' (2.).

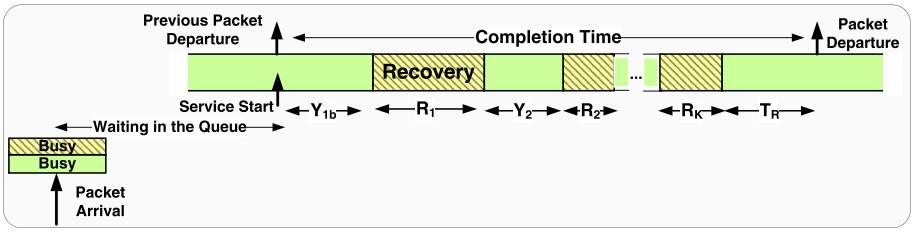

Figure 5 Completion time for the third case when the packet enters a busy system and is queued (case '*b*').

We will now find the first two moments of X_a, X_u, and X_b.

3.1.1 Arrival to an empty-available system

For packets of case '*a*', as illustrated in Figure 4(1), X_a can be given as follows:

$$X_a = \begin{cases} T, & Y_{1a} \geq T \\ Y_{1a} + R_1 + Y_2 + R_2 + \ldots + Y_{K_a} + R_{K_a} + T_R, & \text{Otherwise}, \end{cases} \quad (7)$$

where Y_{1a} is the random remaining time of the first operating period until the next interruption, T_R is the transmission time of the last part of the packet, and K_a is the number of operating periods required to transmit the entire packet. For the second case where $Y_{1a} < T$, we can thus write the following:

$$T = Y_{1a} + \ldots + Y_{K_a} + T_R. \quad (8)$$

If we consider the operating periods $\{Y_{1a}, Y_2, \ldots, Y_{K_a}\}$ as a renewal process, K_a is the number of renewals of Y during the real transmission time of a packet and its distribution can be found from the renewal theory results [27,28]. The first two moments of the number of renewals during $(0, t]$ composed of instances of Y_{1a} and Y are given by the following:

$$m_a(t) = \mathcal{L}^{-1}\left\{\frac{\widehat{Y}_{1a}(s)}{s(1-\widehat{Y}(s))}\right\}, \quad (9)$$

$$m_a^2(t) = \mathcal{L}^{-1}\left\{\frac{\widehat{Y}_{1a}(s)(1+\widehat{Y}(s))}{s\left(1-\widehat{Y}(s)\right)^2}\right\}. \quad (10)$$

The moments of K_a are then given by the following:

$$E[K_a] = \int_0^\infty E[K_a|T=t]f_T(t)dt = \int_0^\infty m_a(t)f_T(t)dt. \quad (11)$$

We can rewrite (7) as follows:

$$X_a = T + \sum_{k=1}^{K_a} R_k. \quad (12)$$

We then obtain the first moment of X_a as follows:

$$E[X_a] = E[T] + E[K_a]\,E[R]\,. \quad (13)$$

For the second moment, we use the facts that K_a is independent of the recovery process, but not the service time, and that the variance of the random sum $\sum_{k=1}^{K_a} R_k$ is equal to $E[K_a]\,Var(R) + (E[R])^2 Var(K_a)$ to obtain the following:

$$E\left[X_a^2\right] = E\left[T^2\right] + 2E\left[TK_a\right]E[R] + E\left[K_a\right]\left(E\left[R^2\right] - (E[R])^2\right) + (E[R])^2 E\left[K_a^2\right], \quad (14)$$

where

$$E\left[TK_a\right] = \int_0^\infty E\left[TK_a|(T=t)\right]f_T(t)dt = \int_0^\infty t m_a(t) f_T(t)dt. \quad (15)$$

3.1.2 Arrival to a busy system

Based on the distribution of the first operating period Y_{1b} (see Figure 5), we can find the first two moments of K_b, the number of renewals for the case '*b*', and X_b , as we did for K_a and X_a in the previous results.

3.1.3 Arrival to an empty-unavailable system

For the case '*u*', as illustrated in Figure 4(2), the completion time of the user started within a recovery period and we have the following:

$$X_u = R_r + Y_1 + R_2 + \ldots + Y_K + R_{K+1} + T_R = R_r + X_u^*. \quad (16)$$

The moments of X_u^* can be found as for X_a by replacing Y_{1a} with Y_1. R_r[a], the remaining time in the first recovery period, can be written as follows:

$$R_r = R - \mathbb{A}|(R > \mathbb{A}), \quad (17)$$

where $\mathbb{A}|(R > \mathbb{A})$ is the inter-arrival time conditioned on the fact that it should be less than R. The moments of R_r based on R and $\mathbb{A}$ are found as follows.

As mentioned above, we encounter several times throughout the paper the random variables $Z = V|(V < U)$ and $Q = V - U|(V > U)$ for any two arbitrary random variables V and U. We derive here the statistics of these two random variables. We have the following:

$$f_Z(t) = \frac{Pr(U > t)f_V(t)}{Pr(U > V)} = \frac{(1 - F_U(t))f_V(t)}{Pr(U > V)}. \quad (18)$$

When U is exponentially distributed with parameter α, we then have:

$$Pr(V < U) = \int_0^\infty e^{-\alpha t} f_V(t) = \widehat{V}(\alpha), \tag{19}$$

$$f_Z(t) = \frac{e^{-\alpha t} f_V(t)}{\widehat{V}(\alpha)}. \tag{20}$$

In this case, $E[Z]$ and $E[Z^2]$ can respectively be given by:

$$E[Z] = \frac{-d/d\alpha \widehat{V}(\alpha)}{\widehat{V}(\alpha)}, \tag{21}$$

$$E[Z^2] = \frac{d^2/d\alpha^2 \widehat{V}(\alpha)}{\widehat{V}(\alpha)}. \tag{22}$$

For the second random variable, Q, we still assume that U is exponentially distributed with parameter α. Then, after some algebra manipulations (details can be found in ([7] Lemma2) or in [9]), we obtain:

$$E[\,Q] = \frac{E[V]}{1 - \widehat{V}(\alpha)} - \frac{1}{\alpha}, \tag{23}$$

and

$$E[Q^2] = \frac{E[V^2] - 2\frac{E[V]}{\alpha}}{1 - \widehat{V}(\alpha)} + \frac{2}{\alpha^2}. \tag{24}$$

Based on the moments of $R_r = R - \mathbb{A}|(R > \mathbb{A})$, the completion time can thus be found as:

$$\begin{aligned} E[X_u] &= E[R_r] + E[X_u^*], \\ E[X_u^2] &= E[R_r^2] + E[X_u^{*2}] + 2E[R_r]\,E[X_u^*]. \end{aligned} \tag{25}$$

3.2 Queue performance metrics

To find the waiting time for our system, we can use the same approach as the one used to derive the Pollaczek–Khinchine formula for M/G/1 queues [26]. When a packet arrives, it waits for the remaining completion time of the packet in service (if any), which, from renewal theory results, is equal to $\frac{E[X^2]}{2}$, and then the completion time of all packets in the queue. For the packets which are in the queue, the completion time is always distributed with X_b (they are queued, so they have not arrived to an empty system). However, for the packet which is initially in service, the general completion time should be used because no knowledge is available to know whether this packet has been of case 'a', 'b', or 'u'. We thus have the following:

$$E[W] = \frac{\lambda E[X^2]}{2(1 - \lambda E[X_b]\,)}. \tag{26}$$

The average system time is given by $E[D] = E[W] + E[X]$.

3.3 Busy periods

Similarly to an M/G/1 queue without interruption [9], we can find the busy periods' distribution for our queue with interruption. This result will be useful to analyze the priority disciplines. We know that the first completion time in a busy period is an instance of X_e. However, for other busy periods which are initiated during X_e, the busy period is started with an instance of X_b because the packets enter a non-empty system. Therefore, we can find the LST of the busy periods as follows:

$$\widehat{B}(s) = \widehat{X}_e(s + \lambda - \lambda \widehat{B}_b(s)), \tag{27}$$

where $\widehat{B}_b(s)$ is the LST of the busy periods which are initiated during X_e with an instance of X_b. $\widehat{B}_b(s)$ itself can be found from the following equation:

$$\widehat{B}_b(s) = \widehat{X}_b(s + \lambda - \lambda \widehat{B}_b(s)). \tag{28}$$

From the equation above, we can find the first and the second moments of $B_b(t)$ as follows:

$$E[B_b] = \frac{E[X_b]}{1 - \lambda E[X_b]} \text{ and } E\left[B_b^2\right] = \frac{E\left[X_b^2\right]}{(1 - \lambda E[X_b]\,)^3}. \tag{29}$$

The first and the second moments of the general busy periods are then given by the following:

$$E[B] = \frac{E[X_e]}{1 - \lambda E[X_b]}, \tag{30}$$

$$E[B^2] = \lambda E\left[B_b^2\right] E[\,X_e] + (1 + \lambda E[B_b]\,)^2\, E\left[X_e{}^2\right]. \tag{31}$$

3.4 Alternative model

An alternative model is to consider the start of the real service as the start of the completion time. We introduce this model since it will be useful to analyze some of the priority schemes. This alternative model does not affect the completion time for arrivals in empty available and busy systems (cases 'a' and 'b'), but for an arrival in an empty unavailable system, the remaining time of the recovery period R_r is considered as waiting time. The completion time is given by X_u^* (see (25)). The first two moments of X^*, the overall completion time for this alternative model, can then be found using the same approach as for X. Since the system time for both models must be the same, we then have that the average waiting time for this model is given by:

$$E\left[W^*\right] = E[W] + E[X] - E\left[X^*\right]. \tag{32}$$

3.5 Approximate and exponential operating periods

As discussed in [7], in general, it is very complex to find the exact distribution of Y_{1a} and Y_{1b}, since they depend on the time when a packet arrives or a packet service has terminated. An approximation for Y_{1a} and Y_{1b} is to assume that they may be started uniformly during an operating period [16], which is sometimes called *random modification* of Y [8] or *equilibrium excess distribution* [7].

For the special, yet important, case that the operating periods are distributed with an exponential distribution Y with parameter α (i.e., $F_Y = 1 - e^{-\alpha t}$), from the memoryless property we have that $Y_{1a} = Y_{1b} = Y$ and $K_a = K_u = K_b$. Using (9) and (10), we have that $m(t) = \alpha t$ and $m^2(t) = \alpha^2 t^2 + \alpha t$. It is then straightforward to derive the moments of the completion time as:

$$E\left[X_{a,b,u}\right] = E[T]\,(1 + \alpha E[R]\,)\,, \tag{33}$$

$$E[TK_a] = \alpha E[\,T^2]\,, \tag{34}$$

$$E\left[X_b^2\right] = E\left[T^2\right](1 + \alpha E[\,R]\,)^2 + \alpha E[T]\,E[R^2]\,, \tag{35}$$

$$E\left[X_u^2\right] = E\left[X_a^2\right] + E\left[R_r^2\right] + 2E[X_a]\,E[R_r]\,. \tag{36}$$

For exponentially distributed operating periods, we can further model our queue as a queue with an initial setup time [9] to find the waiting time. The initial setup time S for a packet which initiates the busy period is R_r with probability $(1 - P_{ae})$ and zero otherwise. Thus, we can find the moments of S based on the moments of Y and R: $E[S] = (1 - P_{ae})E[R_r]$ and $E[S^2] = (1 - P_{ae})E\left[R_r^2\right]$. From ([9] 2.44a), we then have:

$$\begin{aligned} E[D] &= E[X_b] + \frac{\lambda E\left[X_b^2\right]}{2(1 - \lambda E[X_b]\,)} + \frac{2E[S] + \lambda E[S^2]}{2(1 + \lambda E[S]\,)} \\ &= E[X_b] + \frac{\lambda E\left[X_b^2\right]}{2(1 - \lambda E[X_b]\,)} + \frac{E[R^2]}{2(E[Y] + E[R]\,)}. \end{aligned} \tag{37}$$

The steady-state probability of the system being empty, P_0, can be given by:

$$P_0 = \frac{E[I]}{E[I] + E[B_s]} = \frac{1 - \lambda E[X_b]}{1 + \lambda E[S]}, \tag{38}$$

where $E[I] = \frac{1}{\lambda}$ is the average of idle periods (no packet in the system), and $E[B_s]$ is the average of busy periods initiated by $S + X_b$ which can be found from Section 3.3.

3.6 Case study: comparison between switching and buffering OSA strategies

We now present a case study to validate the theoretical analysis and to discuss how it can be used to gain insight on the performance of OSA networks. In this case study, we compare two common OSA strategies which, following the detection of primary users activity on the operating channel, either *switch* to a new channel or *buffer* packets while waiting for the primary users to release the channel [23,29].

It is assumed that there is a large set of similar channels with exponentially distributed availability (I) and unavailability periods (U). For both OSA policies, we have $Y = I$ and for the buffering policy $R = U$ [16]. For the OSA switching policy, we use the common random sensing model in which the channels are sensed successively in a random order until an available channel is found. The interruption time R is thus geometrically distributed with a success probability $\frac{E[I]}{E[I]+E[U]}$ for each time slot of τ (τ is the amount of time required to switch to and sense a channel). In the theoretical model, R is approximated by an exponential distribution with an average length $E[R] = \frac{\tau(E[I]+E[U])}{E[I]}$.

Figure 6 compares the packet sojourn time ($E[D]$), for the system parameters indicated in the figure, of these two models obtained with exact Monte-Carlo simulations and using the theoretical result (37). First, the presented results confirm the accuracy of the theoretical model and its applicability to different OSA strategies. We can also observe that, as can be expected, the threshold point for the average channel unavailability length $E[U]$ where the switching policy becomes preferable over the buffering policy increases from 5.2 to 11.4 units of time when τ increases from 5 to 10 units of time. Note that this threshold is not simply given by the value $E[R]$ where the average interruption time for both policies are equal, but is obtained by finding the value of $E[R]$ where (37) is the same for both OSA policies. Based on the knowledge of the CR users sensing and switching time, and the estimated values of the average channel availability and interruption period lengths [30], the OSA network can therefore use (37) to optimally decide between the switching and buffering policies to minimize the packets sojourn time. In the remainder of the paper, we will derive similar relationships that can be used to analyze and optimally control on OSA network with multiple classes of traffic with priority queueing differentiated services.

4 Priority queueing

We can now tackle the analysis of the four priority queueing disciplines for the general queueing model with N_p classes of CR traffic. Let $\rho_b = \sum_{j=1}^{N} \rho_{b,j} = \sum_{j=1}^{N} \lambda_j E[X_{b,j}]$ and $\mathbb{A} = \min\{A_1, \ldots, A_N\} \rightarrow \mathbb{A} \sim EXP(\lambda)$. We also use the notation $P_{ae}(\lambda)$ and $R_r(\mathbb{A})$ to highlight that P_{ae} and R_r in (3) and (17), respectively, should be calculated with combined λ and $\mathbb{A}$. For the non-preemptive and preemptive priority queueing disciplines presented in Section 4.1 and 4.3, respectively, results for a general distribution for the operating periods are presented. For the exceptional non-preemptive and preemption in case of failure service disciplines, introduced in this paper and presented in Section 4.2 and 4.4 respectively, we only analyze the case of exponential operating periods due to the analytical complexity of those schemes without the assumption of memoryless operating periods. We also present in Section 5 an alternative approach to analyze the preemptive and non-preemptive disciplines for exponential operating periods.

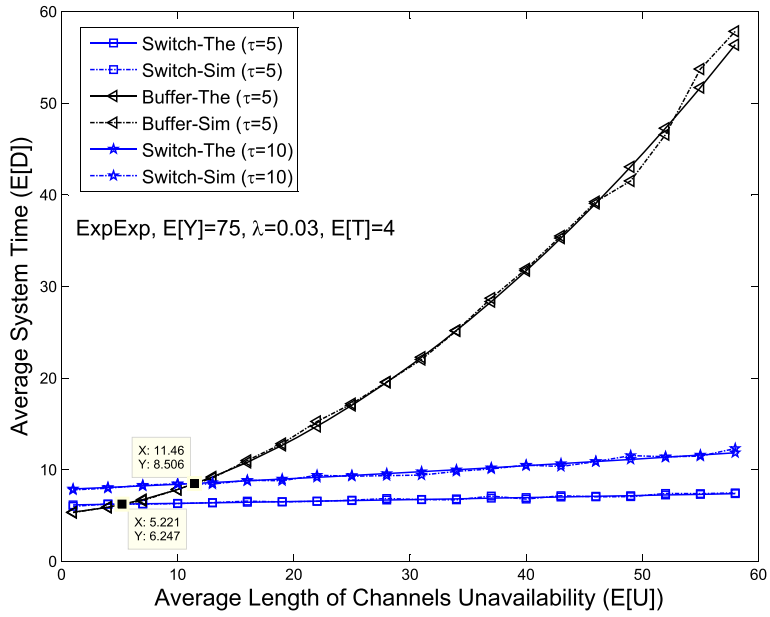

Figure 6 Decision on employing a buffering or switching policy fulfilled with analytical queueing results.

4.1 Non-preemptive

Since the packet service cannot be preempted in this scheme, the completion time of any packet for the three cases ('a', 'b' and 'u') will be the same as for the single traffic queue. The moments of the general completion time X_i for class of traffic i, $i = 1, \ldots, N$, can then be found by solving the system of N equations and N unknowns obtained from (5) for the N classes of traffic where ρ is replaced by $\sum_{j=1}^{N} \lambda_j E[X_j]$. Then, similar to an M/G/1 queue [26], we have:

$$E[W_i] = \frac{E[J]}{\left(1 - \sum_{j=1}^{i} \rho_{b,j}\right)\left(1 - \sum_{j=1}^{i-1} \rho_{b,j}\right)}. \tag{39}$$

where J is the remaining completion time of the packet in service and is given by:

$$E[J] = \sum_{j=1}^{N} \frac{\lambda_j}{2} E\left[X_j^2\right]. \tag{40}$$

Note that since no knowledge is available about the packet in service, the general completion time is used. However, the denominator represents the completion time of the queued packets which is $X_{b,j}$ for class j.

When Y is exponentially distributed, we have that the completion time for the three cases ('a', 'b' and 'u') has the same distribution in the alternative model presented in Section 3.4. Using the same approach as for (37), a closed-form relation can be obtained for the system time by using a queue model with an exceptional completion time X_e for the first packet which initiates a busy period [9]. X_e is given by:

$$\begin{aligned} X_e &= \sum_{i=1}^{N} \frac{\lambda_i}{\lambda} X_{e,i} \\ &= \sum_{i=1}^{N} \frac{\lambda_i}{\lambda} \left[P_{ae}(\lambda) X_{b,i} + (1 - P_{ae}(\lambda))(X_{b,i} + R_r(A))\right]. \end{aligned} \tag{41}$$

It is straightforward to find the first two moments of $X_{e,i}$ and X_e. We then obtain [9]:

$$\begin{aligned} E[D_i] = &\frac{(1-\rho_b)\left(E\left[X_{e,i}\right] + E\left[X_{b,i}\right]\right) + \lambda E\left[X_{e,i}\right] E\left[X_{b,i}\right]}{1 + \lambda E\left[X_e\right] - \rho_b} \\ &+ \frac{\lambda\left[(1-\rho_b) E\left[X_e^2\right] + E[X_e]\left(\sum_{j=1}^{N} \lambda_j E[(X_{b,j})^2]\right)\right]}{2\,(1+\lambda E[X_e]-\rho_b)\left(1-\sum_{j=1}^{i} \rho_{b,j}\right)\left(1-\sum_{j=1}^{i-1} \rho_{b,j}\right)}. \end{aligned} \tag{42}$$

The first term represents the average completion time and the second term, the average waiting time. Similar to (38), P_0, the steady-state probability of the system being empty, is given for this queue by:

$$P_0 = \frac{1 - \rho_b}{1 - \rho_b + \lambda E[X_e]}. \tag{43}$$

4.2 Exceptional non-preemptive

In this scheme, a packet which arrives in an empty unavailable system (case 'u') can be preempted at the end of the arrival recovery period by a higher priority packet

which also arrives in the same recovery period. However, this is in fact a non-preemptive discipline for the alternative model presented in Section 3.4 since in this model a packet which arrives in an empty unavailable system does not start the service, but is queued waiting to obtain the server which will be given to the queued packet with the highest priority. We thus have a non-preemptive queue with initial setup time [9]. Using the same approach as in Section 3.5, we obtain that:

$$E[D_i] = E[X_{b,i}] + \frac{\sum_{j=1}^{N} \lambda_j E\left[\left(X_{b,j}\right)^2\right]}{2\left(1-\sum_{j=1}^{i}\rho_{b,j}\right)\left(1-\sum_{j=1}^{i-1}\rho_{b,j}\right)} + \frac{(1-\rho_b)\left(\lambda E[S^2]+2E[S]\right)}{2\,(1+\lambda E[S]-\rho)\left(1-\sum_{j=1}^{i}\rho_{b,j}\right)\left(1-\sum_{j=1}^{i-1}\rho_{b,j}\right)}. \tag{44}$$

The probability of the system being empty is equal to:

$$P_0 = \frac{1-\rho_b}{1+\lambda E[S]}. \tag{45}$$

4.3 Preemptive

In this scheme, the highest class is not affected by the other classes of traffic. Its completion time and system time can thus directly be found using the results presented in Section 3. Let us now analyze the performance of the low priority class for a two priority class system.

To solve this system, let us find the distribution of Y_2 and R_2, respectively the operating and interruption periods, from the perspective of the low priority (LP) packets[b]. The system is unavailable for LP traffic both due to the activity of high priority (HP) users and due to channel interruptions. As illustrated in Figure 7a, Y_2 is the minimum between the time to the next interruption and the arrival of an HP packet: any one which arrives sooner initiates an interruption period for LP packets. We thus have:

$$Y_2 = \min(Y, A_1) \rightarrow 1-F_{Y_2}(t) = (1-F_Y(t))\left(1-F_{A_1}(t)\right). \tag{46}$$

When Y is exponentially distributed with parameter α, the distribution of Y_2 can be given by:

$$F_{Y_2}(t) = 1 - e^{-(\lambda_1+\alpha)t}. \tag{47}$$

To calculate R_2, we have to distinguish between the events that caused the period of interruption. If, as illustrated in Figure 7a, a high priority (HP) packet arrived and preempted the low priority (LP) traffic, the length of R_2 is equal to one busy period of HP packets which is distributed according to $B_{b,1}$. On the other hand, if the channel interruption caused the unavailability, the two cases shown in Figure 7b and 7c may happen. First, if no HP packet arrives during R, the length of R_2 is equal to $R|R < A_1$. If an HP packet arrives during R, R_2 will be $A_1|(R > A_1)$, the interruption period until the HP packet arrival, extended with an HP busy period B_{R_r}. B_{R_r} can be found from (27), replacing X_e by $X_{b,1} + R_r$ where $R_r = R - A_1|(R > A_1)$ is the remaining time of the server interruption after the HP packet arrival and X_b by $X_{b,1}$. We then have:

$$R_2 = \begin{cases} B_{b,1} & Pr(A_1 < Y), \\ R|(R \le A_1) & Pr(Y \le A_1 \& R \le A_1), \\ A_1|(R > A_1) + B_{R_r} & Pr(Y \le A_1 \& R > A_1). \end{cases} \tag{48}$$

Please be aware that in order to simplify the notation, from now on, we use the notation $V^{<Z}$ to denote $V|(V < Z)$ for any two random variables V and Z. The probability of an HP arrival during R can be given by:

$$P_{A_1<R} = \int_0^\infty (1-e^{-\lambda_1 r})dF_R(r). \tag{49}$$

For exponential Y, we have:

$$Pr(Y \le A_1) = \frac{\alpha}{\alpha+\lambda_1}, \tag{50}$$

and then

$$E[R_2] = \frac{\lambda_1}{\alpha+\lambda_1}E[B_b] + \frac{\alpha}{\alpha+\lambda_1} \times \left[(1-P_{A_1<R})\frac{\frac{-d}{d\lambda_1}\widehat{R}(\lambda_1)}{\widehat{R}(\lambda_1)} + P_{A_1<R}(E\left[A_1^{<R}\right]+E[B_{R_r}])\right]. \tag{51}$$

$\widehat{R}(\lambda_1)$ is the Laplace–Stieltjes transform (LST) of R, replacing s with λ_1, and the derivation is with respect to λ_1. Then, using (23) for $E[R_r]$ and (30) for busy periods, we have:

$$E[B_{R_r}] = \frac{E[X_{b,1}]+E[R_r]}{1-\lambda_1 E[X_{b,1}]} = \frac{E[X_{b,1}] + \frac{E[R]}{1-\widehat{R}(\lambda_1)} - \frac{1}{\lambda_1}}{1-\lambda_1 E[X_{b,1}]}. \tag{52}$$

The second moment of R_2 can be computed similarly, where the second moment of the busy periods can be found from (31) and the second moment of $A_1^{<R}$ and $R^{<A_1}$ can be derived from (22) in Section 3.1.3. It should be taken into account that B_{R_r} and $A_1^{<R}$ are correlated, so $E\left[B_{R_r}A_1^{<R}\right]$ should be calculated separately, using, for instance, the same approach as in (15).

From the equations above, one can find the moments of the operating and interruption periods from the perspective of LP packets (Y_2 and R_2). Then, we return to the original M/G/1 queue with interruptions and replace Y and R in (37) with Y_2 and R_2, respectively, to find the performance metrics of the LP packets.

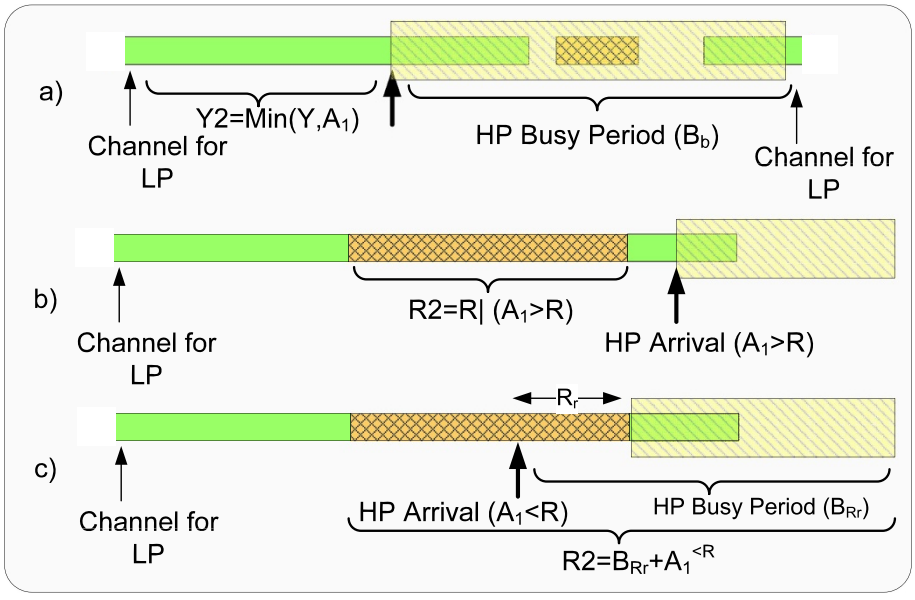

Figure 7 Operating and interruption periods (Y_2 and R_2) from the perspective of LP packets. (a) R_2 is equal to one busy period of HP packets, **(b)** R_2 is equal to $R|R < A_1$, and **(c)** R_2 is equal to $A_1|(R > A_1) + B_{R_r}$.

Note that it is not easy to extend the proposed approach for more than two classes of CR traffic and it is one of the limitations of this work; however, the proposed model can be used to find a bound for the performance of aggregated low priority traffic (combination of all low priority classes). That is, after the first round of analysis, we re-analyze the LP traffic in a recursive way with the highest classes as the aggregated new HP class and all others aggregated as the new LP class. This approach however results in approximations and has a high complexity.

4.4 Preemption in case of failure

In this priority queueing model, high priority packets can only preempt the service from a low priority packet if an interruption occurs. When the service is resumed after an interruption, the priority is given first to high priority packets (HP). In other words, if an HP packet arrives while a low priority (LP) packet is in service, the HP service is started either after the end of the LP service or after an interruption, any one which occurs sooner. As expected, for any class of traffic the performance metrics for this scheme lies between the non-preemptive and preemptive schemes ([31] Vol. 2). The completion time of HP packets is not affected by this service discipline and can be found from the original single traffic queue. However, the waiting time of the HP packet is affected since it must wait until the end of the LP packet transmission or an interruption before starting its service. In the following, we first analyze the completing time of LP packets with this service discipline and then study the HP and LP waiting time. Finally, we discuss the special case where the service time of low priority packets is exponentially distributed.

4.4.1 *Completion time of the LP packets*

To find the completion time of the LP packets, we follow a similar approach as the one used for the preemptive scheme. Due to the memoryless operating periods distribution, from the LP packets perspective, the operating period distribution is not affected by the preemptive in case of failure service discipline and we have $Y_2 = Y$. On the other hand,

R_2, the length of the interruption period from the LP users perspective, is a function of the remaining service time of the LP packet at the HP packet arrival time. That is, the longer the remaining service time until the next interruption, the more HP packets can arrive and thus their busy period will get longer. However, unless the service time is memoryless (this special case is discussed in Section 4.4.3), the remaining service time is not the same for each interruption R_2. Therefore, the completion time cannot be modeled as a renewal process because the instances of R_2 are not identical. We will thus provide approximations for the moments of X_2 (or X_2^*) for two extreme cases: when the operating periods are much larger than the service time of type-2 (LP) packets ($Y >> T_2$) (*large scenarios*) and when it is smaller ($Y < T_2$) (*small scenarios*).

For $Y >> T_2$, it can be assumed that the service of an LP packet is finished in at most two operating periods. This assumption is a trade-off between accuracy and complexity and it is equivalent to assuming at most one instance

of R_2 interruption during the completion time. The completion time will be found based on the alternate model (Section 3.4). The expectation of X_2^* is then given by:

$$E[X_2^*] \approx \begin{cases} E[T_2] & Y \geq T_2 \\ E[T_2] + E[R_2] & Y < T_2. \end{cases} \tag{53}$$

Note that even with the assumption of having a large scenario, there is still a positive probability to have $T_2 > Y$ since both T_2 and Y are random variables. The length of R_2 depends on the arrival of an HP packet and its arrival time. R_2 can be given by:

$$R_2 = \begin{cases} R^{<A_1} & \text{No HP arrival,} \\ B_{C_r} - \left(Y^{<T_2} - A_1 | \left(A_1 < Y^{<T_2}\right)\right) & \text{HP arr. in } Y^{<T_2}, \\ A_1^{<R} + B_{R_r} & \text{HP arr. in } R. \end{cases} \tag{54}$$

where C_r is the remaining time of the cycle (a cycle consists of an operating period Y followed by a recovery period R) after the arrival of an HP packet, and B_{C_r} represents the HP busy period which is initiated with $C_r + X_{b,1}$. However, $Y^{<T_2} - A_1$, the remaining time of the operating period $Y^{<T_2}$ after the HP packet arrival, should be excluded from R_2 since the HP packet does not immediately preempt the LP packet. In the third case, the HP arrival occurs in R. The busy period of HP packets thus starts with $R_r + X_{b,1}$ and the length of the total interruption is A_1 in addition to the HP busy period. $E[R_2]$ can thus be given by:

$$\begin{aligned} E[R_2] = &\left(1 - P_{A_1 < C^{<T_2}}\right) \frac{\frac{-d}{d\lambda_1}\widehat{R}(\lambda_1)}{\widehat{R}(\lambda_1)} + P_{A_1 < C^{<T_2}} \Bigg[P_{ae}(\lambda_1) \\ &\times \left(\frac{E[R] + E\left[Y^{<T_2} - A_1 | A_1 < Y^{<T_2}\right] + E[X_{b,1}]}{(1 - \lambda_1 E[X_{b,1}])} \right. \\ &\left. - E\left[Y^{<T_2} - A_1 | A_1 < Y^{<T_2}\right] \right) + (1 - P_{ae}(\lambda_1)) \\ &\times \left(E\left[A_1^{<R}\right] + \frac{E\left[X_{b,1}\right] + E\left[R_r\right]}{\left(1 - \lambda_1 E\left[X_{b,1}\right]\right)} \right) \Bigg], \end{aligned} \tag{55}$$

where $P_{A_1 < C^{<T_2}}$ is the probability of an arrival in $C^{<T_2} = Y^{<T_2} + R$, and P_{ae} is calculated for HP packets. The second moment of R_2 can be found similarly using the second moment of the busy periods and the relations provided in Section 3.1.3. However, the correlation of random variables B_{R_r} and $A_1^{<R}$ should be taken into account.

For the case where $Y < T_2$, we assume that the duration of HP busy periods is independent of the activity of LP packets. Therefore, the interruption periods from the perspective of LP packets have the same distribution and a renewal process can be considered to analyze the completion time. When a LP packet starts its service, as illustrated in Figure 8, it holds the channel (available or unavailable) for a cycle and if there is an HP arrival during the cycle, it releases the channel to HP packets at the end of the cycle. So, the probability of releasing the channel to HP traffic at the end of the cycle is given by:

$$P_{A_1 < C} = Pr(\text{HP arrival in } C) = \int_0^\infty \left(1 - e^{-\lambda_1 c}\right) f_C(c) dc. \tag{56}$$

For R_2, (54) is still valid except that we eliminate the condition that $Y < T2$.

We then obtain:

$$\begin{aligned} E[R_2] = &\left(1 - P_{A_1 < C}\right) \frac{\frac{-d}{d\lambda_1}\widehat{R}(\lambda_1)}{\widehat{R}(\lambda_1)} \\ &+ P_{A_1 < C} \left[P_{ae}(\lambda_1) \left(\frac{E[R] + E[Y]}{\left(1 - \lambda_1 E\left[X_{b,1}\right]\right)} - E[Y] \right) \right. \\ &\left. + (1 - P_{ae}(\lambda_1)) \left(E\left[A_1^{<R}\right] + \frac{E[X_{b,1}] + E[R_r]}{(1 - \lambda_1 E[X_{b,1}])} \right) \right]. \end{aligned} \tag{57}$$

The second moment can be found similarly. The moments of R_2 can then be substituted into the results for the single traffic queue to find the moments of the completion time of LP packets $E[X_2]$ and $E[X_2^2]$.

4.4.2 Waiting time

For the high priority (HP) traffic, the waiting time of HP arrivals during Y in a system empty of HP packets is affected by the presence of LP packets. For those packets, the new waiting time is zero if the LP queue is empty or, if the LP queue is not empty, the minimum of the remaining time of the arrival cycle and the remaining service time of the LP packet in service. The difficulty to compute the waiting time is thus the dependency of both types of traffic on each other. That is, while the waiting time of the HP packets is affected by the lower class, both the waiting time and the completion time of LP packets are affected by HP traffic. This obliges us to use approximations and bounds to find the waiting time of LP and HP packets.

The waiting time of HP packets is upper bounded by an M/G/1 queue with vacation in (37) if we neglect the unknown part of the service that the LP has received so far and assume that the remaining LP service time is still T_2. That is, the initial setup time S can be approximated as follows:

$$S = \begin{cases} R_r(\lambda_1) & 1 - P_{ae}(\lambda_1), \\ Y^{<T_2} + R & Pr(Y \leq T_2)(1 - P_0)P_{ae}(\lambda_1), \\ T_2^{<Y} & Pr(Y > T_2)(1 - P_0)P_{ae}(\lambda_1), \\ 0 & P_0 P_{ae}(\lambda_1), \end{cases} \tag{58}$$

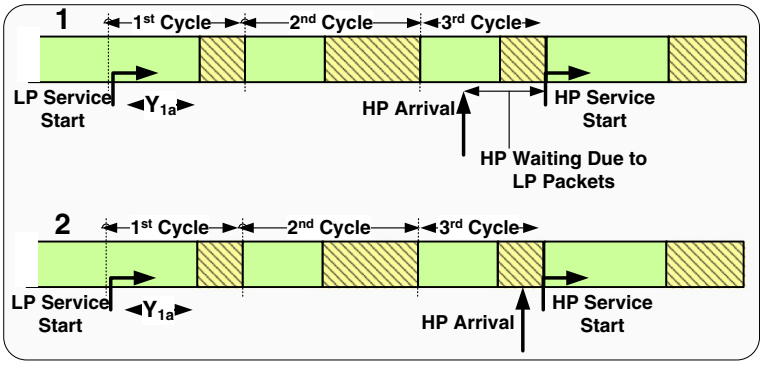

Figure 8 Cycles and holding periods for the LP packets in the discipline of preemptive in case of failure (FP). The LP packet can hold the channel for three cycles. However in a preemptive scheme (Pr) (not shown here), it can keep the channel for two complete cycles and releases the channel in the middle of the third cycle if there is an HP arrival (section 1).

where $P_{ae}(\lambda_1)$ and $R_r(\lambda_1)$ only take HP packets into account, and P_0 is the probability of system being empty of any type of packet. Note that the correct probability to be used here instead of $1 - P_0$ is $P_{L|NH}$, which is the probability that there are LP packets in the system given that it is empty of HP packets. However, this probability cannot be found without any assumption on T_2's distribution; therefore, we used P_0 as an approximation. This approximation results in a setup time, S, larger than its real value, which provides an upper bound in (59) that can be sometimes looser than expected. P_0 is the same for all priority queueing models [9] and is given in (43). An obvious lower bound on the waiting time is given by assuming that HP packets always preempt LP packets (preemptive discipline). We can thus write:

$$\frac{\lambda E\left[X_{b,1}^2\right]}{2(1-\lambda E[X_{b,1}])} + \frac{E[R^2]}{2(E[Y]+E[R])} \leq E[W_1^*] < \frac{\lambda E\left[X_{b,1}^2\right]}{2(1-\lambda E[X_{b,1}])} + \frac{2E[S]+\lambda_1 E[S^2]}{2(1+\lambda E[S])}. \tag{59}$$

We can then use those bounds on the HP waiting time to find corresponding bounds on the LP waiting time through the conservation law (CL) in a queue with multiple classes of traffic [31] which indicates that the quantities

$$\kappa = \lambda_1 E[T_1]\,E[W_1] + \lambda_2 E[T_2]\,E[W_2], \tag{60}$$

and

$$\kappa^* = \lambda_1 E[T_1]\,E[W_1^*] + \lambda_2 E[T_2]\,E[W_2^*], \tag{61}$$

for alternative model, are constant for all priority service disciplines. κ and κ^* can thus be computed with the waiting time of LP and HP packets found for one of the previous priority disciplines.

An alternative approach is to directly find the LP waiting time and then use the conservation law to obtain the HP waiting time. But, similarly to the HP waiting time, it is difficult to find an exact expression for the LP waiting time due to the strong interdependence between both types of traffic. We thus propose to compute bounds as follows.

Using the approximations for the first two moments of X_2 or X_2^* found previously, the minimum waiting time of LP packets can be found using the Pollaczek–Khinchine relation [26]. An upper bound for the LP packets' waiting time is naturally given by the waiting time in the preemptive discipline model.

We thus have two upper and lower bounds for both the HP and LP waiting time. The tighter bounds can then be selected as the final lower and upper bound for both traffic categories.

4.4.3 Exponentially distributed LP service time

The queue model for the HP traffic is an M/G/1 queue with vacations. Since for a memoryless exponentially distributed service time, the remaining parts of the service are identically distributed, it is possible to exactly express S, the initial setup time for the HP traffic, as follows:

$$S = \begin{cases} R_r(\lambda_1) & 1 - P_{ae}(\lambda_1), \\ Y^{<T_2} + R & Pr(Y \leq T_2) P_{L|NH} P_{ae}(\lambda_1), \\ T_2^{<Y} & Pr(Y > T_2) P_{L|NH} P_{ae}(\lambda_1), \\ 0 & Otherwise, \end{cases} \tag{62}$$

where $P_{ae}(\lambda_1)$ and $R_r(\lambda_1)$ only take HP packets into account. $E[S]$ is then given by:

$$E[S] = (1 - P_{ae}(\lambda_1))\,E[R_r] + P_{L|NH} P_{ae}(\lambda_1) \times \left[\left(\frac{1}{\gamma_2+\alpha} + \frac{\alpha}{\alpha+\gamma_2} E[R]\right)\right]. \tag{63}$$

α is the exponential parameter for Y and γ_2 is the exponential parameter for T_2. The unknown in the preceding

equation is $P_{L|NH}$, which is the probability that there are LP packets in the system given that it is empty of HP packets, and can be expressed as:

$$P_{L|NH} = 1 - \frac{P_0}{P_{0,1}}, \tag{64}$$

where $P_{0,1}$, the probability that the system is empty of HP packets, can be found from the original queue by substituting $E[S]$ with the one that was calculated in (63), and, as indicated previously, P_0 is the probability that the system is empty of all packets and is given in (43). We thus finally have two equations ((63) and (64)) which can be used to find the two unknowns $E[S]$ and $P_{L|NH}$. The second moment of S can then be found and the HP waiting time is given by (37).

The LP packets waiting time can be found easily from the conservation law and the previous result for the HP packets waiting time. To find the system time, we will now find an exact expression for the first two moments of the completion time of LP packets. The interruptions, from the LP packets point of view, are identically distributed due to the exponential LP service time distribution. The same renewal process approach which was used as an approximation for $Y < T_2$ duration in (57) can thus be used exactly for this case. The only change is that Y should be replaced with $Y^{<T_2} = Y|(Y < T_2)$. We then have

$$\begin{aligned} E[R] = &\left(1 - P_{A_1<Y|Y<T_2} - P_{A_1<R}\right)E\left[R^{<A_1}\right] \\ &+ {}_{A_1<Y|Y<T_2}\left(E[B_{C_r}] - E\left[Y_r^{<T}\right]\right) \\ &+ P_{A_1<R}\left(E[B_{R_r}] + E\left[A_1^{<R}\right]\right), \end{aligned} \tag{65}$$

where

$$\begin{aligned} P_{A_1<Y|Y<T_2} &= \frac{\lambda_1}{\lambda_1 + \alpha + \gamma}, \\ P_{A_1<R} &= \left(1 - P_{A_1<Y|Y<T_2}\right) Pr(R < A_1), \\ E[B_{C_r}] &= \frac{E[R] + \frac{1}{\gamma_2+\alpha} + E[X_{b,1}]}{1 - \lambda_1 E[X_{b,1}]}, \\ E[B_{R_r}] &= \frac{E[R_r] + E[X_{b,1}]}{1 - \lambda_1 E[X_{b,1}]}. \end{aligned} \tag{66}$$

Similarly, $E\left[R_2^2\right]$ and consequently $E[X_2]$ and $E\left[X_2^2\right]$ can be found.

It is worth noting that only the assumption of exponential service time for LP packets is required to find the results above. The service time of HP packets can be general.

5 Alternative approach for priority queueing analysis

In this section, we propose an alternative approach to analyze the preemptive and non-preemptive priority queueing service disciplines in OSA networks. As interruptions have a preemptive behavior, we can model the interruptions as the highest priority type of traffic whose packet inter-arrival time is distributed with a random variable Y and whose *busy periods* are distributed with a random variable R. We can then have an estimate for the service time of these virtual highest priority packets whose busy period models the interruptions. Closed-form relations cannot be derived in general since for any distribution of R and Y, a different formula for the busy periods and consequently for the service time exists. If Y is exponentially distributed with parameter α, we can assume an M/G/1 queue for the interruptions. Note that, as discussed in the introduction, this *alternative approach provides an approximation* because we know that the real distribution of the busy periods in an M/G/1 queue is a complicated function built on the Bessel function [9], which cannot exactly be matched to R. We discussed this alternative approach in this paper for completeness and as an extension of the analysis provided in [6,12,13] for multiple classes of traffic with preemptive and non-preemptive service disciplines.

Using this alternative approach, it is possible to find approximate results for the preemptive and non-preemptive service disciplines as follows. We first find the first two moments of the service time of the virtual packets which form the interruptions from the distribution of the busy periods of a regular M/G/1 queue [9]:

$$\widehat{R}(s) = \widehat{T}_v(s + \alpha - \alpha\widehat{R}(s)), \tag{67}$$

$$E[T_v] = \frac{E[R]}{1 + \alpha E[R]}, E\left[T_v^2\right] = E[R^2]\,(1 - \alpha E[T_v])^3. \tag{68}$$

where T_v stands for the service time of the virtual packets which form the interruptions. As interruptions have a preemptive behavior, we can model the preemptive service discipline as a preemptive-resume queue with $N + 1$ classes of traffic, the highest priority packets being the virtual packets. The following extensions of the P-K formula for preemptive-resume schemes [9] can then be used to find the average waiting time of other classes of traffic:

$$E[W_i] = \frac{E[J_i]}{(1 - \alpha E[T_v] - \ldots - \lambda_{i-1}E[T_{i-1}])(1 - \alpha E[T_v] - \ldots - \lambda_i E[T_{i-1}])}, \tag{69}$$

where $E[J_i]$ can be given by:

$$E[J_i] = \frac{1}{2}\alpha E\left[T_v^2\right] + \sum_{j=1}^{i} \frac{1}{2}\lambda_j \left(E\left[T_j^2\right]\right), \tag{70}$$

and the moments of T_v can be found from (68). The system time is then given by:

$$E[D_i] = \frac{E[T_i]}{1 - \rho_{i-1} - \cdots - \rho_1 - \rho_v} + E[W_i] \tag{71}$$

For the non-preemptive service discipline, the queue can be modeled as a priority queue where the highest class of traffic (virtual) behaves preemptively, but other classes behave non-preemptively. The extensions of the P-K formula given in (69) can be used to find the average waiting time of other classes of traffic where $E[J_i] = E[J]$ is the same for all priority classes and is equal to:

$$E[J] = \frac{1}{2}\alpha E\left[T_v^2\right] + \sum_{j=1}^{N} \frac{1}{2}\lambda_j \left(E\left[T_j^2\right]\right). \tag{72}$$

The system time can be written as:

$$E[D_i] = \frac{E[T_i]}{1 - \alpha E[T_v]} + E[W_i]\,. \tag{73}$$

6 Simulation results

In this section, we validate the analytical results of the OSA networks priority queueing disciplines presented in this paper by comparing with system accurate Monte-Carlo simulation results. The presented results also give several insights on the performance of OSA networks with mixed traffic. We consider in the numerical evaluation a system with two classes of traffic: high priority (HP) and low priority (LP), also denoted as type-1 and type-2 packets, respectively. We assumed exponentially distributed operating periods and considered the cases of exponentially and constantly distributed interruption periods. The service time (packet length) is also assumed to have either an exponential or a constant distribution. Due to space limitations, we only present the results for the exponential and constant (deterministic) distributions.

Unless mentioned otherwise, the HP arrival rate is assumed equal to 0.03 and the average real service times are $E[T_1] = 3$ and $E[T_2] = 5$ (the unit of time is irrelevant). As an example, if the unit of time is *millisecond* (ms) and the channel data rate is 4 Mbps, those service times represent 1, 500 bytes and 2, 500 bytes packets, respectively.

The duration of operating and interruption periods are selected to model two different scenarios. The first scenario is for an almost quasi-static cognitive radio network where $E[Y] \gg E[T]$ and used $E[Y] = 75$ and $E[R] = 15$. The second scenario is for a highly dynamic cognitive radio network [32] where $E[Y] < E[T]$ and used $E[Y] = 1$ and $E[R] = 0.2$. Note that the average server availability is the same for both scenarios, only the dynamics are different.

In the figures, the different distribution cases are denoted by 'ExpExp', 'ExpDet', 'DetExp' and 'DetDet', respectively for the distributions of T and R, as summarized in Table 2. The four service disciplines are denoted in the figures as 'Non' (non-preemptive), 'ENo' (exceptional non-preemptive), 'Pr' (preemptive) and 'FP' (preemption in case of failure). The suffix 'Sim' indicates the simulation results, 'The' corresponds to the analytical evaluation of the theoretical results presented in Section 4, and 'Alt' indicates the analytical evaluation of the theoretical results for the alternative approach presented in Section 5.

Table 2 Service time and recovery time distribution cases

Service time (*T*)	**Exponential**	**Exponential**	**Constant**	**Constant**
Recovery time (*R*)	Exponential	Constant	Exponential	Constant
Notation	ExpExp	ExpDet	DetExp	DetDet

6.1 Exponential recovery and service time

The recovery time for a random channel selection recovery algorithm (i.e., the user senses a list of channels one by one until finding an available channel) or a slotted-Aloha competition with other CR users can be accurately modeled with an exponential distribution.

Figure 9 shows the HP and LP average system time respectively for this case. As expected, and can also be observed for all the presented results, the HP system time increases from the preemptive, preemptive in case of failure, exceptional non-preemptive and non-preemptive service disciplines, and the LP system time increases in the inverse order of service disciplines. The results also show that the simulation and theoretical analysis results closely match, which validates the priority queueing analysis.

Results are also presented for static and dynamic operating period scenarios. We can observe that for the same server availability ratio $E[Y]/E[R]$, the system time is worst for both classes of traffic and all service disciplines for the static scenario where $E[R]$ and $E[Y]$ are much larger than the service time. This is due to the fact that the long recovery periods in the static scenario have a severe impact on the OSA queue performance metrics for all traffic classes. To further investigate this important finding, in Figure 10 we present the system time as a function of $E[T_1]/E[R]$ for a server availability ratio $E[Y]/E[R]$ fixed to five. This figure clearly shows that the OSA network system time performance deteriorates as the system dynamic decreases (i.e., when $E[T_1]/E[R]$ decreases) with an inflexion point when the service time is approximately equal to the average interruption length. Furthermore, both HP and LP packets are similarly affected. That is, queueing disciplines cannot protect HP traffic against long interruptions. This is expected since interruptions indeed preempt the server. Note that for a traditional OSA throughput analysis based on a saturated-traffic model, no major performance changes will be observed as a function of the OSA network dynamic since the main factor is the server availability ratio $E[Y]/E[R]$. Only the complete queueing analysis presented in this paper can give an insight on the important impact of system dynamics on the OSA performance.

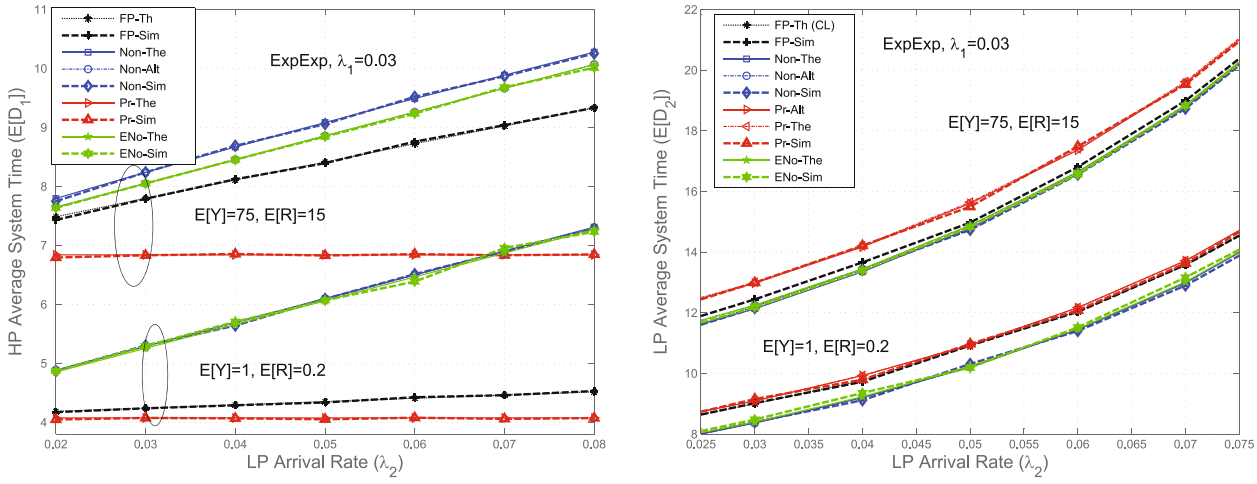

Figure 9 System time of high priority (HP) and low priority (LP) packets vs. LP arrival rate.

The other interesting point to observe is that for dynamic scenarios, due to frequent short interruptions, the preemption in case of failure service discipline enables the quick preemption of LP packets by HP packets. This service discipline performance is thus close to the preemptive scheme for dynamic scenarios. On the other hand, for static scenarios the preemption in case of failure service discipline performance gets closer to the non-preemptive discipline due to the lack of opportunities for HP packets to preempt LP packet service. Meanwhile, the system time for the exceptional non-preemptive discipline is very close to the non-preemptive scheme performance in dynamic scenarios because the probability of an HP arrival in an empty system in the same recovery period as an LP arrival is very low. In large scenarios, their performances start to differ. However, the performance gain remains small. Those results show that the novel priority discipline of preemption in case of failure for OSA networks can significantly improve the system time of HP packets in several deployment scenarios while having a lower implementation complexity than a full preemptive service discipline.

In Figure 11, we study the CR traffic system time as a function of the operating and interruption period length.

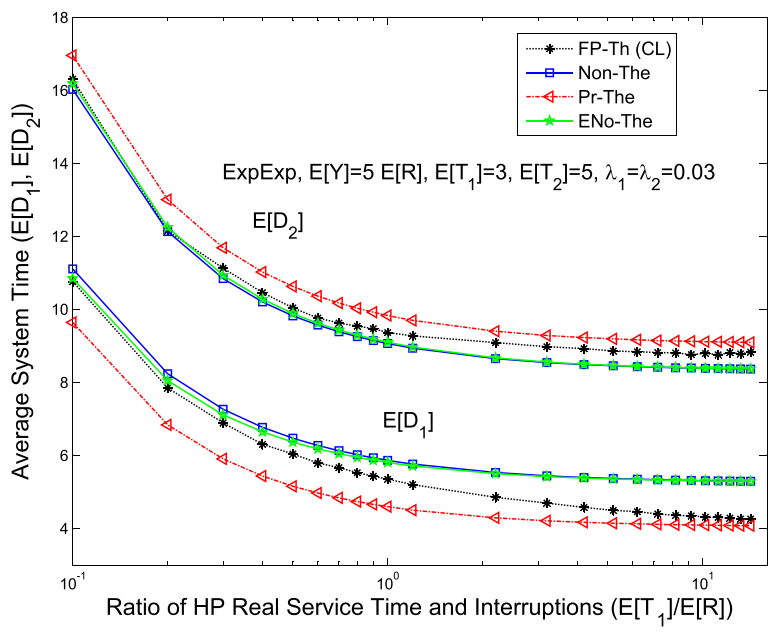

Figure 10 System time of LP and HP packets vs. the variations of *E*[*R*] and *E*[*Y*] when their ratio is fixed.

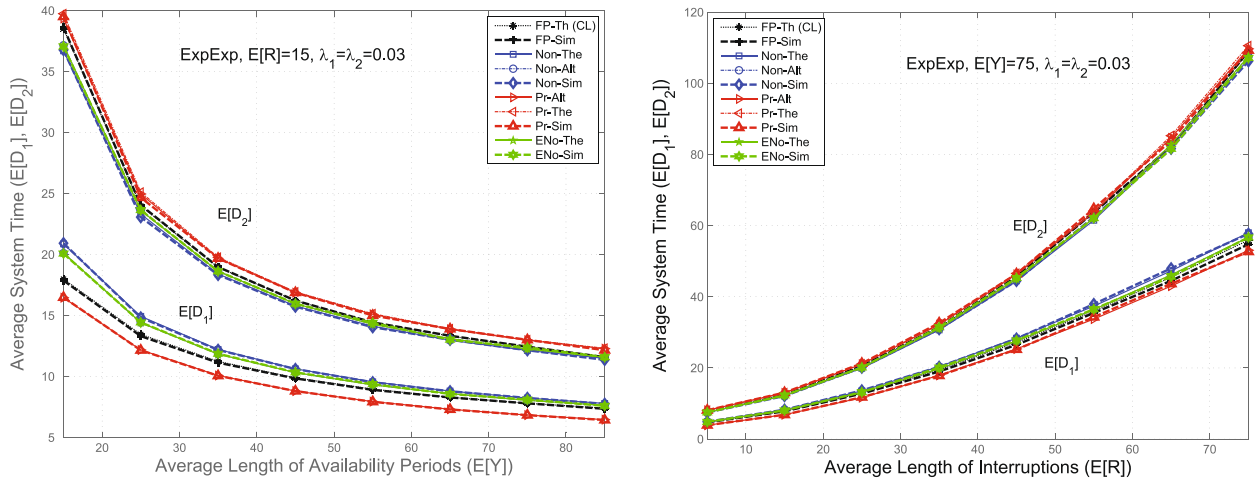

Figure 11 System time of HP and LP packets vs. operating and interruption period duration.

Those results illustrate the validity of the queue analysis for a wide range of operating and interruption periods. Note that since $E[R]$ or $E[Y]$ is fixed, the server availability increases as a function of $E[Y]$ in the former case and decreases as a function of $E[R]$ in the later case. We can observe that as the server availability decreases, either due to shorter availability periods or longer interruption periods, the system time significantly increases, with LP traffic being more affected than HP traffic due to the priority service disciplines. It is interesting to again note that the system time increases much faster when the interruption period increases than when the availability period decreases. For example, starting from the point where $E[Y] = 75$ and $E[R] = 15$ to the point where $E[Y] = E[R]$, the HP system time approximately increases by a factor of three when the operating period length decreases and by a factor of eight when the interruption period length increases. Those results underline the critical importance of minimizing the interruption period length in OSA networks.

The results presented in Figure 12 further motivate the importance of the theoretical analysis provided in this paper to correctly analyze the performance of OSA networks. A straightforward tempting simplification that could be used to analyze the CR queue system time is to use the standard M/G/1 formulas without interruption and increase the packets real service time T_i by the ratio $\frac{E[Y]+E[R]}{E[Y]}$ to compensate for the average throughput loss due to interruptions. It can be shown for both traffic classes that the queues will saturate at the same traffic load for both the simplified analysis and the correct analysis. However, as can be seen in Figure 12, the M/G/1 simplification (referred as 'NoInt') significantly underestimates by almost an order of magnitude the real performance of the queue for both preemptive and non-preemptive priority disciplines (this simplification does not allow the analysis of the two other OSA service disciplines due to the absence of interruptions). This error is due to the fact that the simplified modeling is in fact equivalent to assuming that the interruption periods approach a length of zero. But, as we have discussed previously, interruption periods have a major impact on the OSA queue performance. The accurate modeling of the interruption periods, as we provided in this paper, is thus critical to obtain a valid OSA queue analysis.

Figure 12 also presents the system time when both traffic classes are mixed without a priority service discipline (i.e., both packet types are queued together and are served in a first-in-first-serve scheme). It is interesting to observe that for a standard system without interruption, the system time increase for LP packets when a priority service discipline is used is almost the same as the system time decrease for HP packets (i.e., the mixed traffic service time is almost exactly in the middle between the LP and HP service times with both priority service discipline). However, this is not the case for the OSA network where the HP packet system time decreases significantly more than the system time increase for LP packets. This is due to the fact that the interruption periods preempt both classes of traffic. Those results indicate that differentiated service with priority queueing has a bigger impact in OSA networks than in conventional networks and should thus be actively used when they carry multiple classes of traffic with different QoS requirements.

6.2 Exponential recovery time and constant service time

In Figure 13, the HP and LP packet lengths are both constant ('DetExp' scenario). Thus, the lower bound for the waiting time of LP packets for the preemptive in case of failure service discipline is presented. The results

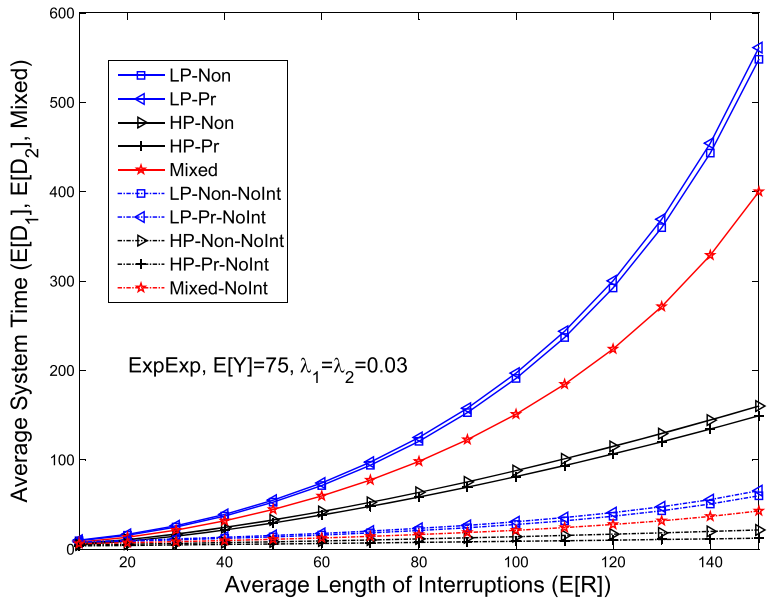

Figure 12 **Performance comparison of accurate queueing models with interruption with a modified M/G/1 with longer packets but no interruption.**

are generated for two different values of HP arrival rate. To better observe the accuracy of the completion time approximations for the preemptive in case of failure service discipline, simulation results and approximations for the moments of the LP packets completion time, X_2 and X_2^*, are compared in the upper part of Table 3. Note that the completion time of LP packets is independent of their arrival rate. The results show the accuracy of the analysis for constant service time and the slight deterioration due to the approximate analysis for the preemptive in case of failure service discipline. On the other hand, we can observe the accuracy of the theoretical analysis with deterministic packet service time for the three other service disciplines.

6.3 Constant recovery time

A constant recovery time occurs, for instance, in scenarios where the information concerning the channels'

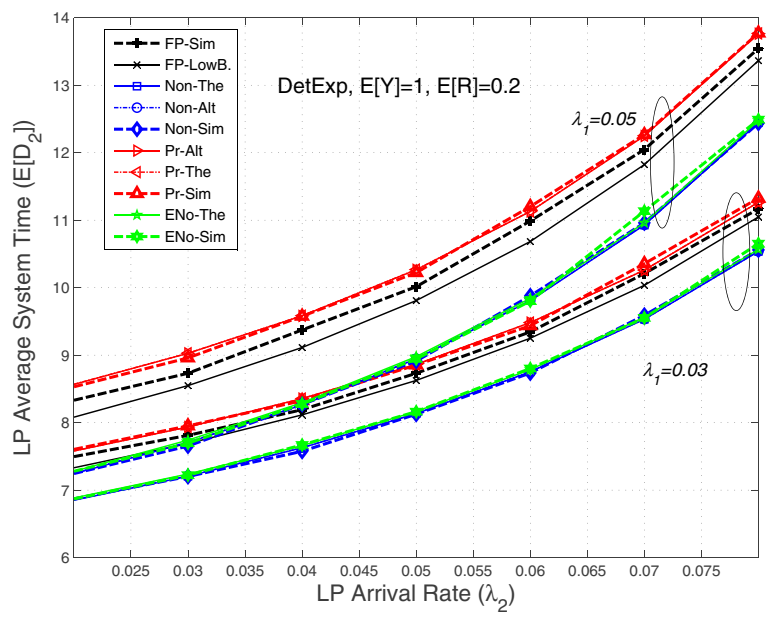

Figure 13 **System time of LP packets vs. LP arrival rate for two values of HP arrival rate (small scenario).**

Table 3 Moments of the LP completion time with preemption in case of failure (FP) service discipline for DetExp and DetDet scenarios (S: simulation, A: approximation)

Scenario	λ_1	$E[X_2]$-S	$E[X_2]$-A	$E[X_2^2]$-S	$E[X_2^2]$-A
Small,DE	0.03	6.62	6.72	47.01	49.51
Small,DE	0.05	7.10	7.32	57.05	62.33
Large,DE	0.03	6.13	6.12	77.53	74.96
Large,DE	0.05	6.26	6.25	88.37	85.62
Small,DD	0.03	6.61	6.72	46.76	49.23
Small,DD	0.05	7.10	7.31	56.69	61.98
Large,DD	0.03	6.14	6.12	58.40	56.45
Large,DD	0.05	6.26	6.25	65.00	63.64

occupancy is provided in advance; therefore, no random sensing is required and the recovery time only represents a constant time for negotiation and radio alignment. In order to compare the results with the previous scenario, we assume the same average values.

Figure 14 illustrates the system time of HP and LP packets versus their arrival rate for the cases of exponential and constant service times.

It should be noted that for the selected values in the simulation, the upper and lower bounds proposed in Section 4.4.2 (e.g., Eq. (59)) are loose compared to the natural bounds of preemptive and non-preemptive disciplines, so they are not illustrated in the figures to enhance their clarity. As expected, it can be seen in both figures that the performance, when the distribution of real service time (packet length) is exponential, is worse compared to the case where the packet length is constant (with the same average). The results also validate the theoretical analysis presented in Section 4. However, it can be observed that the system time for the non-preemptive service discipline obtained with the alternative approach presented in Section 5 is not accurate. This shows, as discussed in the introduction, the limitations of alternative approaches which were previously proposed in the literature when the recovery period is not exponentially distributed. Furthermore, this alternative approach cannot be used to analyze more sophisticated service disciplines such as the exceptional non-preemptive and the preemptive in case of failure disciplines which provide interesting performance gains for OSA networks. Simulation and analytical results for the completion time of LP packets are compared in the lower part of Table 3 for two different values of HP arrival rate.

7 Conclusions

Priority queueing is a classical approach to implement traffic differentiation in communication links. To analyze priority queueing schemes for opportunistic spectrum access networks, we derived in this paper a general queueing model with interruptions for the preemptive and non-preemptive classical priority disciplines. Two new cognitive radio disciplines were also introduced in this paper: exceptional non-preemptive and preemptive in case of failure. The theoretical analysis was validated with simulation results and we investigated the behavior of those disciplines for different sets of parameters and distributions for the packet service time and interruption periods. We also showed how the analysis can be used by an OSA controller to make critical decisions such as selecting the channel switching policy or the priority queueing discipline based on the estimated channel parameters. It was also observed that even though the ratio of operating and interruption periods plays an important role, a significant performance decrease is

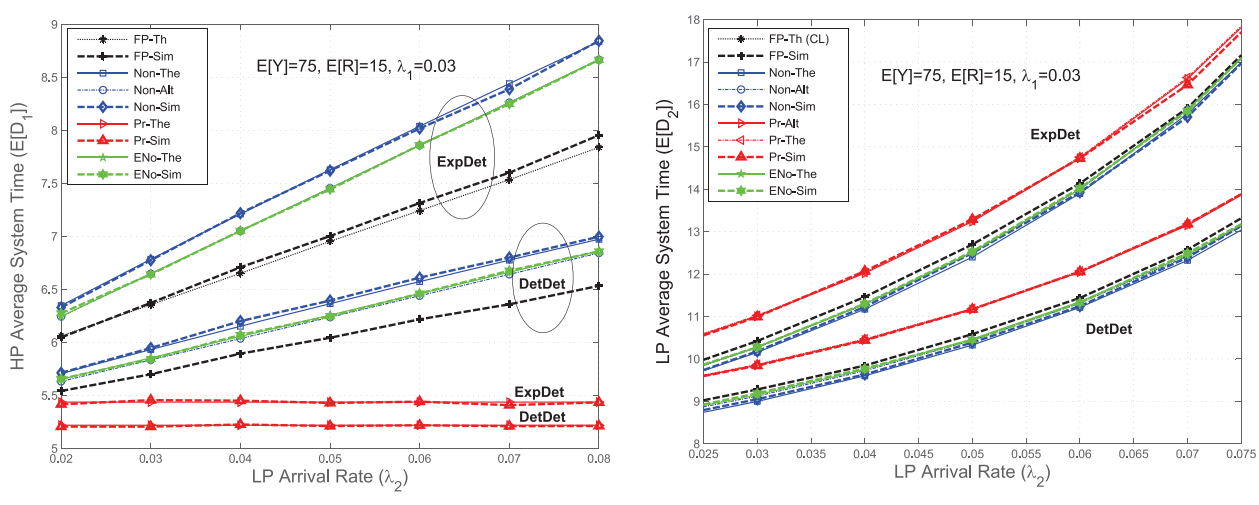

Figure 14 System time of high (HP) and low priority (LP) packets vs. LP arrival rate (DetDet and ExpDet, large scenario).

observed in a semi-static network with long operating and interruption periods compared to a fast-varying network with short periods and the same ratio. A simplified M/G/1 model with no interruption and compensated increased packet length cannot thus capture the queue metrics performance of OSA networks. We also presented results demonstrating the importance of priority queueing to provide differentiated service in the presence of frequent interruptions. As discussed in the introduction, an important area of future work is to use the results presented in this paper to further study and optimize MAC protocols and channel assignment policies in OSA networks based on not only a saturated mode throughput analysis but also on queue metrics. Another interesting area of research is to extend this work to the cases of queueing with service repeat after an interruption and for non-homogeneous channels with variable service rate for different operating periods.

Endnotes

[a]For notation simplicity, R_r is used instead of R_{1u}.

[b]They are here two new random variables. Not to be mistaken with identical instances of Y and R in previous sections.

Abbreviations

AP: Access point; CDF: Cumulative distribution function; CL: Conservation law; CR: Cognitive radio; CRN: Cognitive radio network; HP: High priority; LP: Low priority; MAC: Medium access control; OSA: Opportunistic spectrum access; P-K: Pollaczek–Khinchine; PDF: Probability density function; PU: Primary user; SU: Secondary user.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This research project was supported by NSERC under Grant STPG365205.

Received: 2 April 2014 Accepted: 21 November 2014
Published: 1 December 2014

References

1. FK Jondral, Cognitive radio: A communications engineering view. IEEE Wireless Commun. **14**(4), 28–33 (2007)
2. J Mitola, Cognitive radio architecture evolution. Proc. IEEE. **97**(4), 626–641 (2009)
3. J Xiao, RQ Hu, Y Qian, L Gong, B Wang, Expanding LTE network spectrum with cognitive radios: from concept to implementation. IEEE Wireless Commun. **20**(2), 12–19 (2013)
4. Cisco, White paper: Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013–2018, 1–40. San Jose, CA, US, (2014) http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.pdf
5. A Azarfar, J-F Frigon, B Sansò, in *IEEE 76th Vehicular Technology Conference(VTC2012-Fall)*. Dynamic selection of priority queueing discipline in cognitive radio networks, (2012), pp. 1–5
6. L Wang, C Wang, K Feng, A queueing-theoretical framework for QoS-enhanced spectrum management in cognitive radio networks. IEEE Wireless Commun. **18**(6), 18–26 (2011)
7. A Federgruen, L Green, Queueing systems with service interruptions. Oper. Res. **34**(5), 752–768 (1986)
8. B Avi-Itzhak, Preemptive repeat priority queues as a special case of the multipurpose server problem -II. Oper. Res. **11**(4), 610–619 (1963)
9. H Takagi, *Queueing Analysis: a Foundation of Performance Evaluation, Volume 1 Vacation and Priority Systems*, (North-Holland, Netherlands, 1991)
10. JDP Gaver, A waiting line with interrupted service, including priorities. J. R. Stat. Soc. Ser. B (Methodological). **24**(1), 73–90 (1962)
11. D Fiems, T Maertens, H Bruneel, Queueing systems with different types of server interruptions. Eur. J. Oper. Res. **188**(3), 838–845 (2008)
12. MM Rashid, J Hossain, E Hossain, VK Bhargava, in *Proc. of IEEE GLOBECOM '07*. Opportunistic spectrum access in cognitive radio networks: a queueing analytic model and admission controller design, (2007), pp. 4647–4652
13. A Laourine, S Chen, L Tong, in *Proc. of IEEE INFOCOM 2010*. Queuing analysis in multichannel cognitive spectrum access: a large deviation approach, (2010), pp. 1–9
14. S Wang, J Zhang, L Tong, in *Proc. of IEEE INFOCOM 2010*. Delay analysis for cognitive radio networks with random access: a fluid queue view, (2010), pp. 1–9
15. D Chen, S Yin, Q Zhang, M Liu, S Li, in *Proc. of MobiCom '09*. Mining spectrum usage data: a large-scale spectrum measurement study, (2009), pp. 13–24
16. A Azarfar, J-F Frigon, B Sansò, Analysis of cognitive radio networks based on a queueing model with server interruptions. IEEE International Conference on Communications, ICC 2012, 1703–1708 (2012)
17. H Li, Z Han, Socially optimal queuing control in cognitive radio networks subject to service interruptions: to queue or not to queue? IEEE, Trans. Wireless Commun. **10**(5), 1656–1666 (2011)
18. K Kim, T-preemptive priority queue and its application to the analysis of an opportunistic spectrum access in cognitive radio networks. Comput. Oper. Res. **39**(7), 1394–1401 (2012)
19. H-P Shiang, M van der Schaar, Queuing-based dynamic channel selection for heterogeneous multimedia applications over cognitive radio networks. IEEE Trans. Multimedia. **10**(5), 896–909 (2008)
20. Y Li, Y Jiang, D Jin, L Su, L Zeng, DO Wu, Energy-efficient optimal opportunistic forwarding for delay-tolerant networks. IEEE Trans. Vehic. Technol. **59**(9), 4500–4512 (2010)
21. P Pawelczak, S Pollin, H So, ARS Bahai, RV Prasad, R Hekmat, Performance analysis of multichannel medium access control algorithms for opportunistic spectrum access. IEEE Trans. Vehic. Technol. **58**(6), 3014–3031 (2009)
22. A Azarfar, J-F Frigon, B Sansò, Improving the reliability of wireless networks using cognitive radios. IEEE, Commun. Surv. Tutor. **14**(2), 338–354 (2012)
23. J Park, P Paweczak, D Cabric, Performance of joint spectrum sensing and MAC algorithms for multichannel opportunistic spectrum access ad hoc networks. IEEE Trans. Mobile Comput. **10**(7), 1011–1027 (2011)
24. A Azarfar, J-F Frigon, B Sansò, Delay analysis of multichannel opportunistic spectrum access MAC protocols. Submitted to IEEE Transactions on Mobile Computing (2014). [Preprint available at] http://arxiv.org/abs/1403.7192
25. D Skordoulis, Q Ni, H-H Chen, A Stephens, C Liu, A Jamalipour, IEEE 802.11n MAC frame aggregation mechanisms for next-generation high-throughput WLANs. IEEE Wireless Commun. **15**(1), 40–47 (2008)
26. D Bertsekas, R Gallager, *Data Networks*, 2nd edn. (Prentice-Hall, US, 1992)
27. DR Cox, *Renewal Theory*. (Methuen; Wiley, London, New York, 1962)
28. SM Ross, *Introduction to Probability Models, Ninth Edition*. (Academic Press, Inc, Orlando, FL, USA, 2006)
29. L Lai, H El-Gamal, H Jiang, HV Poor, Cognitive medium access: exploration, exploitation, and competition. IEEE Trans. Mobile Comput. **10**(2), 239–253 (2011)
30. W Gabran, CH Liu, P Pawelczak, D Cabric, Primary user traffic estimation for dynamic spectrum access. IEEE J. Select. Areas Commun. **31**(3), 544–558 (2013)
31. L Kleinrock, *Queuing Systems*. (Wiley, NJ, US, 1975)
32. H Khalife, N Malouch, S Fdida, Multihop cognitive radio networks: to route or not to route. IEEE Netw. **23**(4), 20–25 (2009)

doi:10.1186/1687-1499-2014-206
Cite this article as: Azarfar *et al.*: **Priority queueing models for cognitive radio networks with traffic differentiation.** *EURASIP Journal on Wireless Communications and Networking* 2014 **2014**:206.