A Discrete-Time Multi-server Model for Opportunistic Spectrum Access Systems

Islam A. Abdul Maksoud(✉) and Sherif I. Rabia

Department of Engineering Mathematics and Physics, Faculty of Engineering, Alexandria University, Alexandria, Egypt
{islammax,sherif.rabia}@alexu.edu.eg

Abstract. In opportunistic spectrum access communication systems, secondary users (SUs) exploit the spectrum holes not used by the primary users (PUs) and cease their transmissions whenever primary users reuse their spectrum bands. To study the mean time an SU spends in the system we propose a discrete-time multi-server access model. Since periodic sensing is commonly used to protect the PU, discrete-time models are more convenient to analyze the performance of the SU system. Additionally, a multi-server access model is assumed in order to give the SU the capability to access a channel that is not occupied by a PU or any other SUs. We derive the probability generating function of the number of connections in the system. Then we derive a formula for the mean response time of an SU. In the numerical results we show the relationship between the mean response time and the SU traffic intensity. In addition we show the effect of changing the number of channels in the system and the PU traffic intensity on the mean response time of an SU.

Keywords: Discrete-time queueing · Multi-server · Cognitive radio · Opportunistic spectrum access

1 Introduction

The conventional static spectrum access policy had caused a paradoxical spectrum scarcity problem. Although some bands are rarely accessed by their licensed holders, some other bands (e.g. ISM band) are overloaded with traffic [1]. In order to properly manage the problem, another newly dynamic policy is proposed. One of the most appealing models of the new dynamic policy is the opportunistic spectrum access (OSA) model [1]. In the OSA model, unlicensed users (called secondary users or SUs) access the frequency bands owned by the licensed users (called primary users or PUs) only when the PUs do not use their frequency bands and cease their connections when the PUs reuse their channels. In order to realize the OSA model with both acceptable PU protection and SU performance, the SU should be able to perform the following main spectrum management functionalities: spectrum sensing, spectrum decision, spectrum sharing, and spectrum mobility [1],[2].

In literature, there exists a rich body of research works analyzing the performance of OSA systems employing queueing. These works can be classified into continuous-time models and discrete-time models. In continuous-time models, e.g. [3]-[7], opportunities

© Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2015
M. Weichold et al. (Eds.): CROWNCOM 2015, LNICST 156, pp. 381–388, 2015.
DOI: 10.1007/978-3-319-24540-9_31

can be exploited whenever they appear in the time axis. PUs and SUs can start their connection at any time and depart the system at any time. However, in discrete-time models, e.g. [8]-[13], the time axis is divided into equal time slots. Primary or secondary users are allowed to access the system (i.e., start their connections or transmissions) only at the beginning of a time slot. Although continuous-time models are easier in analysis, discrete-time models are more realistic for OSA systems. This is because an SU usually has one transceiver to either detect the PU or transmit its data. As a result, the most common protocol is to perform periodic sensing to detect the PU, then transmit the data in between the sensing periods if a spectrum hole is available.

The models analyzing the performance of OSA systems can also be classified into single-server access and multi-server access models. In single-server-access models, e.g. [12],[14]-[17], SUs access a certain channel based on a probability profile or a deterministic profile. If there is no PU accessing the channel, then all SUs accessing the same channel either contend on the channel using conventional MAC protocols or queue until they have the right to start their transmissions. However, in multi-server-access models, SUs access only the channels that have no PUs or any other SUs accessing them at the moment (i.e., completely idle). If there is a buffer in multi-server access models it will be a single global buffer for all the channels instead of a local buffer at each channel as in single-server access models.

Since there is an abundance of single-server discrete-time queueing literature, most papers in OSA working with discrete-time models are restricted to single-server access, e.g. [9],[10],[12],[13]. To the best of the authors' knowledge, the only discrete-time model applying multi-server access is [11]. However, [11] is with no buffer. That meant higher forced termination probability for the ongoing SU connections. In this paper, we present a discrete-time model with multi-server access but with an infinite buffer. The infinite buffer vanishes the forced termination probability but with the cost of increasing the delay of some connections. One of the main objectives of the paper is to study the mean time in system of an SU connection and its relation with different parameters, namely, the overall SU traffic, the number of channels in the system and the PU traffic.

In [18],[19], Bruneel and Laevens investigated the discrete-time queueing analysis of an infinite buffer multi-server system with the number of available servers changing randomly over time. The queueing analysis was in terms of the probability generating function (PGF) of the system contents [18] and the PGF of the delay [19]. They assumed general i.i.d bulk arrivals for the number of arrivals during a single time slot. A single arrival is a packet, where a packet constitutes the amount of data transmitted in a single time slot. In this paper, we utilize the model already available in [18],[19], however, we extend the model to be applied to users whose connection is constituted of a geometric number of packets instead of the deterministic service time assumed in [18],[19].

2 System Model

2.1 Primary Network

We assume a primary network with multiple PUs and a PU base station (PU-BS). The PUs are working on a frequency band that is equally divided into M homogeneous channels. The network is time-slotted with all PUs synchronized to the same time slot structure with the help of the PU-BS. That is, any PU can only access a channel at the beginning of a time slot.

It is assumed that the occupation of the channels by PUs is independent and identically distributed (i.i.d). Among adjacent time slots, the PU activity is assumed to be independent, i.e., the state of the primary activity in a certain time slot is independent of the state of the primary activity on that channel in previous time slots. We assume that the probability to have channel i occupied by a PU at time slot n is a constant probability p.

2.2 Secondary Network

We assume a centralized network of SUs which have capability to exploit the time slots not used by PUs in the M-channel band. SUs must first synchronize to the time slot structure of the PU network. With the SU network thoroughly synchronized to the PU network, SUs can sense the PU activity at the channels only at the beginning of the time slot. In this paper, we assume that, at each time slot, all SUs see the same spectrum opportunities and that the sensing results are perfect. In addition, any sensing overhead or delay is neglected.

We assume that SUs connection requests arrive to the SU base station (SU-BS) as a Poisson process with an arrival rate λ. If there is a channel available to the incoming request, it immediately uses that channel. If no channel is available, the SU request is saved in an infinite buffer at the SU-BS. Whenever one channel or more become available, the SU-BS assigns available channels to the requests waiting in the buffer in a first come first serve (FCFS) discipline. The length of the SU connection is assumed to be a random variable with a geometric distribution for the number of packets. A packet is the amount of data transmitted in a single time slot. The average length of the SU connection is assumed to be $1/s$. We assume that each SU is equipped with one transceiver, thus, each SU connection is assigned to at most one channel at a time.

If an SU is using a channel at a certain time slot and a PU occupies the channel on the next time slot, then if the SU connection has still some remaining packets not transmitted, the SU immediately switches to another available channel. This is called the handoff process. If no other channel is available, this indicates the failure of the handoff process. However, instead of dropping the failed-handoff connection, we propose suspending the SU connection while keeping a request at the infinite buffer in the SU-BS to complete the connection. The position of that request in the buffer is based on the timing the SU sent its request to the SU-BS as a new arrival (i.e., when it first entered the system). Priority is always given to the request with the earliest arrival time as a new connection. By doing so, priority is given to interrupted connections over newly coming ones. This is often desirable in communication systems, as more delay in an ongoing connection is more annoying for the user than the delay encountered before the connection is setup.

In this paper, we also make the assumption that arrivals at a certain time slot cannot be transmitted before the next time slot.

3 System Analysis

In this paper, we use the same analysis in [18] to model an OSA network with respect to the performance of SUs. However, instead of assuming the user's connection as a single packet [18],[19], by a single and simple tweak we extend the model to be applied to users whose connection is constituted of a geometric number of packets.

The discrete-time system they investigated can be described in detail as follows. They assumed general i.i.d bulk arrivals with a PGF $A(z)$ for the number of packets arriving during a single time slot. In addition, they assumed a general i.i.d distribution for the number of available servers during a single time slot with a PGF $C(z)$, where transitions in the number of available servers can only occur at the boundaries between consecutive slots. Additionally, they put the assumption that packets arriving in a particular time slot cannot be transmitted during this same slot.

Given that the mean bulk size $\bar{a}$ is less than the mean number of available servers $\bar{c}$, the analysis yields the following expression of the steady state PGF of the system contents,

$$V(z) = \frac{(\bar{c}-\bar{a})A(z)(z-1)}{z^M - z^M C(1/z)A(z)} \prod_{i=1}^{M-1} \frac{z-z_i}{1-z_i} ,$$

where M is the maximum number of available servers during a time slot, and z_i's are the zeros of the denominator excluding $z_0 = 1$, i.e., z_i's are all the roots (except $z_0 = 1$) of the complex equation

$$z^M - z^M C(1/z)A(z) = 0 .$$

Since we assume Poisson arrival bulks, we have

$$A(z) = e^{-\lambda_s(1-z)} ,$$

where $\lambda_s = \lambda \Delta T$ is the mean number of secondary arrivals during a single slot and ΔT is the time slot duration. Additionally, since we assume i.i.d Bernoulli PU occupation of the channels with probability of occupation p in a time slot, the PGF of the number of available channels during a time slot, denoted $\acute{C}(z)$, can be expressed as

$$\acute{C}(z) = [(1-q) + qz]^M ,$$

where $q = 1 - p$.

Since the service time of a packet in [18],[19] was a single time slot, the PGF of the number of available servers during a single time slot $C(z)$ therein played the role of describing the number of actual departures at the end of each time slot. However, in this paper, we aim at extending the model to geometric service time. That means that the number of available servers only describes the number of potential departures in a time slot. In order to describe the number of actual departures during a time slot in the new model we must include the parameter s of the geometric distribution of number of packets in a connection (i.e., the service time). Thus, for the case of geometric number of packets of each connection, the PGF of the number of departures during a time slot can be described as

$$C(z) = [(1 - qs) + qsz]^M .$$

Accordingly, by applying the same analysis in [18], the PGF of the steady state number of users in the system can be described as

$$V(z) = \frac{(Mqs-\lambda_s)(z-1)}{z^M e^{\lambda_s(1-z)} - [(1-qs)z+qs]^M} \prod_{i=1}^{M-1} \frac{z-z_i}{1-z_i},$$

provided that $Mqs > \lambda_s$, where z_i's are all the roots (except $z_0 = 1$) of the complex equation

$$z^M e^{\lambda_s(1-z)} - [(1-qs)z + qs]^M = 0 . \tag{1}$$

4 Performance Analysis

In this section, we compute the mean response time of an SU connection in the system. The response time is defined as the total time an SU connection spends in the system. Because we assume that an arriving connection cannot depart in the same time slot it came in, the calculation of the response time is divided into two parts. Starting from the time instant the SU arrived to the system, the first part of the response time is the time until the next time slot. Then, from the beginning of that time slot until the departure of the SU is the second part. The mean value of the first part is $\Delta T/2$. For the second part, we calculate it by using Little's theorem as follows. The mean number of connections in the system can be computed from the PGF by calculating $V'(1)$. The expression for $V'(1)$ is found in [19] to be,

$$V'(1) = \sum_{i=1}^{M-1} \frac{1}{1-z_i} - M + A'(1) + \frac{A''(1)+C''(1)+2C'(1)\left(1-A'(1)\right)}{2\left(C'(1)-A'(1)\right)} .$$

For our model, we have $A'(1) = \lambda_s$, $A''(1) = {\lambda_s}^2$, $C'(1) = Mqs$, and $C''(1) = M(M-1)(qs)^2$. Then, by using Little's result, the second part of the mean response time can be expressed as $[V'(1)/A'(1)]\,\Delta T$. Hence, the mean SU response time can be expressed as

$$\text{mean SU response time} = \left(\frac{1}{2} + \frac{V'(1)}{A'(1)}\right) \Delta T .$$

5 Numerical Results

To illustrate the above analysis, we present the following numerical example to study the mean SU response time vs. the effective SU traffic intensity $I = \bar{a}/\bar{c}$. We consider two scenarios. In the first scenario, the PU traffic is fixed while the maximum number of available channels is varied. In the second scenario, the maximum number of available channels is fixed while PU traffic is changed. We note that in order to find all the roots of (1), we replaced the right hand side with the 50th degree Taylor polynomial, then solved the resulting polynomial equation.

5.1 Scenario 1

Fig.1 shows the relation between the SU mean response time and the SU effective traffic intensity with a fixed PU probability of time-slot occupation $p = 0.5$. The mean number of packets in an SU connection is assumed to be 4 packets. The relation is drawn for different number of channels, namely, 4,8,12 and 16 channels. The figure obviously shows that the mean SU response time is nearly stable until SU traffic intensity ≈ 0.7. However, afterwards the mean response time increases rapidly as we approach the instability region (i.e.,$\bar{a} \geq \bar{c}$). As expected, increasing the number of available channels decreases the mean SU response time. Additionally, increasing the number of available channels widens the interval for which the mean SU response time is stable and makes the curve more acute.

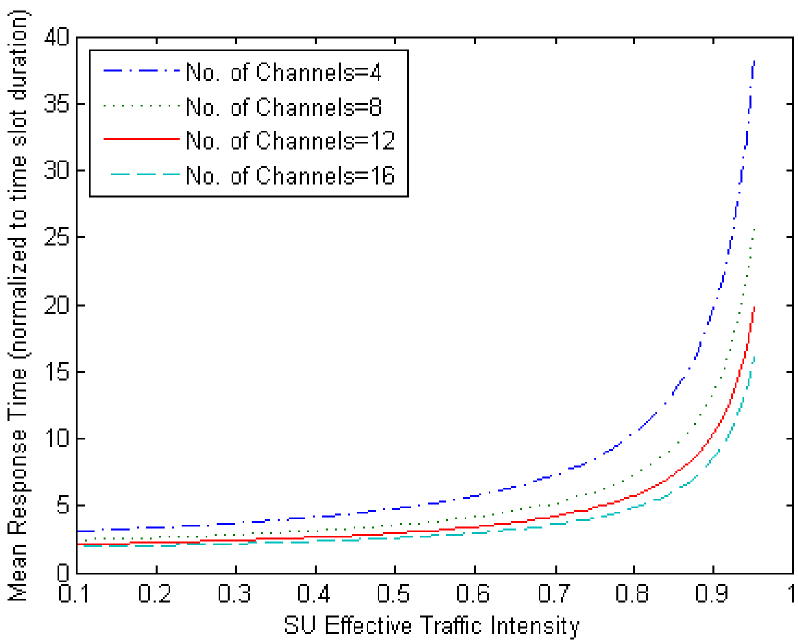

Fig. 1. SU mean response time vs. SU effective traffic intensity for different number of channels in the system.

5.2 Scenario 2

Fig.2 shows the relation between the SU mean response time and the SU effective traffic intensity with a fixed maximum number of available channels $M = 12$. As in scenario 1, the mean number of packets in an SU connection is assumed to be 4 packets. The relation is drawn for different PU traffic activity. Namely, the probability of PU time-slot occupation p is taking the values 0.2, 0.4, 0.6 and 0.8. Expectedly, the figure shows an increase of the mean SU response time as the primary traffic increases.

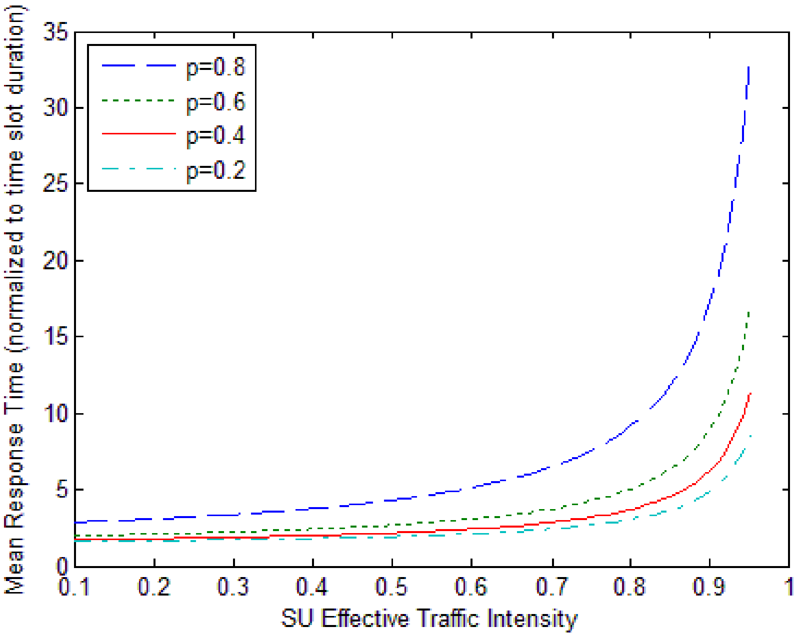

Fig. 2. SU mean response time vs. SU effective traffic intensity for different PU probability of time slot occupation.

6 Conclusion and Future Work

In this paper we have assumed a discrete-time multi-server model to analyze the performance of the SUs with geometric service time. We first derived the PGF of the number of SU connections in the system. Then, we derived a formula for the mean response time of an SU. In the numerical results we have presented the relation between the mean SU response time and the SU traffic intensity in different scenarios of either changing the number of channels in the system or changing the PU traffic intensity. It is shown that the mean SU response time keeps nearly stable for moderate SU traffic intensities, however, the mean response time increases rapidly for large values of SU traffic intensities as the instability region is approached. As expected, the mean SU response time increases when the number of channels in the system decreases or when the primary traffic increases. In the future work, we aim at deriving the PGF of the delay of an SU connection.

References

1. Akyildiz, I.F., Lee, W.-Y., Vuran, M.C., Mohanty, S.: NeXt generation dynamic spectrum access cognitive radio wireless networks: a survey. Computer Networks **50**(13), 2127–2159 (2006)

2. Akyildiz, I.F., Lee, W.-Y., Vuran, M.C., Mohanty, S.: A survey on spectrum management in cognitive radio networks. IEEE Communications Magazine **46**(4), 40–48 (2008)
3. Zhu, X., Shen, L., Yum, T.S.P.: Analysis of cognitive radio spectrum access with optimal channel reservation. IEEE Communications Letters **11**(4), 304–306 (2007)
4. Ahmed, W., Gao, J., Suraweera, H., Faulkner, M.: Comments on Analysis of cognitive radio spectrum access with optimal channel reservation. IEEE Transactions on Wireless Communications **8**(9), 4488–4491 (2009)
5. Zhang, Y.: Dynamic spectrum access in cognitive radio wireless networks. In: IEEE International Conference on Communications ICC 2008, pp. 4927–4932 (2008)
6. Hong, C.P.T., Lee, Y., Koo, I.: Spectrum sharing with buffering in cognitive radio networks. In: Nguyen, N.T., Le, M.T., Świątek, J. (eds.) ACIIDS 2010. LNCS, vol. 5990, pp. 261–270. Springer, Heidelberg (2010)
7. Tang, S., Mark, B.L.: Performance analysis of a wireless network with opportunistic spectrum sharing. In: Global Telecommunications Conference, GLOBECOM 2007, 26-30, pp. 4636–4640 (2007)
8. Rashid, M., Hossain, M., Hossain, E., Bhargava, V.: Opportunistic spectrum scheduling for multiuser cognitive radio: a queueing analysis. IEEE Transactions on Wireless Communications **8**(10), 5259–5269 (2009)
9. Wu, Q., Xu, Y., Wang, J.: A discrete-time model for multichannel opportunistic spectrum access based on preempted queuing. In: International Conference on Wireless Communications and Signal Processing (WCSP), pp. 1–6 (2010)
10. Wu, Q., Wang, J., Xu, Y., Gao, Z.: Discrete-time Queuing Analysis of Opportunistic Spectrum Access: Single User Case. Frequenz **65**(11–12) (2011)
11. Gelabert, X., Sallent, O., Pérez-Romero, J., Agustí, R.: Spectrum sharing in cognitive radio networks with imperfect sensing: A discrete-time Markov model. Computer Networks **54**(14), 2519–2536 (2010)
12. Suliman, I., Lehtomaki, J.: Queueing analysis of opportunistic access in cognitive radios. In: Second International Workshop on Cognitive Radio and Advanced Spectrum Management, CogART 2009, pp. 153–157 (2009)
13. Sheikholeslami, F., Nasiri-Kenari, M., Ashtiani, F.: Optimal Probabilistic Initial and Target Channel Selection for Spectrum Handoff in Cognitive Radio Networks. IEEE Transactions on Wireless Communications **14**(1), 570–584 (2015)
14. Shiang, H.P., van der Schaar, M.: Queuing-based dynamic channel selection for heterogeneous multimedia applications over cognitive radio networks. IEEE Transactions on Multimedia **10**(5), 896–909 (2008)
15. Wang, C.-W., Wang, L.-C., Feng, K.-T.: A queueing-theoretical framework for QoS-enhanced spectrum management in cognitive radio networks. IEEE Wireless Communications **18**(6), 18–26 (2011)
16. Wang, C.-W., Wang, L.-C., Chang, C.-J.: Modeling and analysis for spectrum handoffs in cognitive radio networks. IEEE Transactions on Mobile Computing **11**(9), 1499–1513 (2012)
17. Wang, C.-W., Wang, L.-C.: Analysis of Reactive Spectrum Handoff in cognitive radio networks. IEEE Journal on Selected Areas in Communications **30**(10), 2016–2028 (2012)
18. Bruneel, H.: A general model for the behaviour of infinite buffers with periodic service opportunities. European Journal of Operational Research **16**(1), 98–106 (1984)
19. Laevens, K., Bruneel, H.: Delay analysis for discrete-time queueing systems with multiple randomly interrupted servers. European Journal of Operational Research **85**(1), 161–177 (1995)