Analysis of Sub-Band Allocation in Multi-Service Cognitive Radio Access Networks

Glaucio H. S. Carvalho, *Member, IEEE*, Isaac Woungang, *Senior Member, IEEE*, and
Alagan Anpalagan, *Senior Member, IEEE*

***Abstract*—This letter addresses the sub-band allocation problem in dynamic spectrum access (DSA) scheme in a multi-service cognitive radio network (CRN) built over a multi-service primary network (PRN). Both access networks support connection-level and packet-level services, where the Poisson process and the Markov modulated Poisson process are used to describe the connection-level and packet-level services, respectively. By means of a multi-dimensional continuous-time Markov chain model, the DSA scheme is assessed under various traffic load profiles. Numerical results show that the connection-level secondary users's quality of service may be significantly deteriorated by the packet-level primary user service.**

***Index Terms*—Multi-service cognitive radio network, bursty data traffic, dynamic spectrum access, Markov chain.**

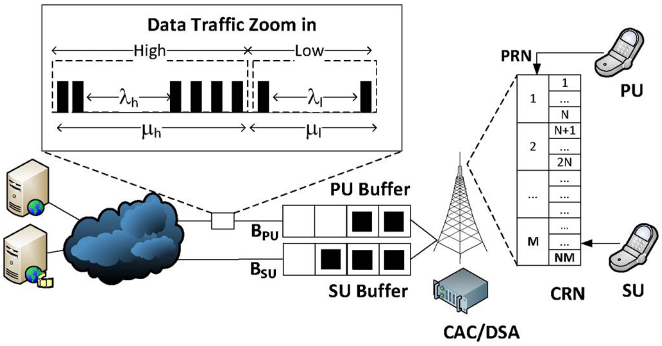

Fig. 1. Wireless network architecture.

I. INTRODUCTION

SUB-BAND allocation performed by dynamic spectrum access (DSA) schemes in cognitive radio networks (CRNs) has received a great attention in literature. In [1] and [2], a general model for channel allocation with channel reservation in CRN was proposed. Due to the channel reservation, a tradeoff between forced termination probability and blocking probability was analyzed. In [3], a DSA algorithm was proposed, which considers a buffer to accommodate the preempted secondary users (SUs) who are unable to perform the spectrum handoff. In [4], Wang *et al.* generalized most of the previous results by addressing the problems of channel reservation for spectrum handoff, buffers for SUs handoff and for SUs new calls, and impatient SUs. Distributed and centralized approaches, along with SU traffic prioritization in the DSA design were considered in [5]. In [6], a DSA algorithm was assessed under the conditions that the arrival processes follow a Markovian arrival process (MAP) and the channel holding times obey a phase-type (PH) distribution. In [1]–[6], the following conditions prevail: 1) the primary radio network (PRN) is a single-service network; 2) the PUs generate the traffic at the connection-level (CL) [1], [2], [4]–[6] or at the packet-level (PL) [3]; 3) the CRNs are single-service with PL traffic [3] or CL traffic [1], [2], [6], or they are multi-service with CL traffic [4], [5]; 4) most studies have been carried mainly for the PUs and SUs' Poissonian arrival at the CL, but in [3], ON/OFF traffic at the PL is assumed and in [6], a MAP arrival at the CL is assumed.

With the shift from voice-centric networks to multi-service ecosystems, both PRNs and CRNs have become multi-service networks that support both the CL and the PL traffic concurrently. Furthermore, with the prevalence of data-oriented systems, a proper characterization of data traffic, which encompasses its bursty and correlated nature [7], at PL becomes instrumental. In this respect, a channel learning algorithm that learns the bursty PU data traffic pattern and uses it to efficiently access the frequency channels is proposed in [8]. However, no sub-band allocation is investigated.

In this letter, we investigate the sub-band allocation in DSA schemes for CRN, where the PRN and CRN are multi-service and support multi-level services. In this scenario, the Poisson traffic and the Markovian modulated Poisson process (MMPP) are respectively used to represent the arrivals at the CT and the PL. The use of MMPP is justified by the fact that the long-term memory presented in the current Internet applications becomes meaningless beyond a certain time scale [9]. Furthermore, the MMPP captures very well the burstiness and correlated nature of the network traffic [10] while making the analysis more tractable. With the aim to gain insight into the SU's performance under such PRN, a multi-dimensional continuous-time Markov chain (CTMC) model is proposed. Due to the severe SUs' QoS degradation, numerical results and analysis unfold the need to rethink the DSA design in the next generation wireless networks where data services will be prevalent.

II. NETWORK ARCHITECTURE

A. System Model

Fig. 1 summarizes the system under analysis. As shown, a network controller performs the call admission control (CAC)

Manuscript received June 26, 2014; accepted July 18, 2014. Date of publication July 31, 2014; date of current version December 17, 2014. This work was supported in part by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) of Brazil and by the National Science and Engineering Research Council of Canada (NSERC) under Grant REF RGPIN/293233-2011. The associate editor coordinating the review of this paper and approving it for publication was T. Q. S. Quek.

G. H. S. Carvalho and I. Woungang are with the Department of Computer Science, Ryerson University, Toronto, ON M5B 2K3, Canada (e-mail: glauciohscarvalho@gmail.com; iwoungan@scs.ryerson.ca).

A. Anpalagan is with the Department of Electrical and Computer Engineering, Ryerson University, Toronto, ON M5B 2K3, Canada (e-mail: alagan@ee.ryerson.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LWC.2014.2344669

2162-2337 © 2014 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

and DSA roles to manage the PUs and SUs. Furthermore, two FIFO buffers, PU buffer and SU buffer, accommodate the data packets that come from the application servers and that will be delivered to the PUs or SUs. The licensed spectrum consists of M primary bands, each of which is divided into N sub-bands [1], [2] to form the CRN. The system supports the connections at two levels: the connection-level (CL) and the packet-level (PL). Therefore, there are four service classes into the system: connection-level PU (CLPU), packet-level PU (PLPU), connection-level SU (CLSU), and packet-level SU (PLSU). To use the wireless spectrum, the following hierarchy prevails: CLPU $>$ PLPU $>$ CLSU $>$ PLSU.

For the channel allocation, a CLPU (reap. CLSU) requires one band (reps. sub-band) to accommodate its service whereas a PLPU (reap. PLSU) uses all available primary bands (reps. sub-bands) for data transmission. When a PU arrives and claims a band, the preempted SU performs the spectrum handoff to resume its service as long as there is enough room to host it. Otherwise, the service will be abruptly forced to terminate. To mitigate the forced termination occurrences, a total of $L < MN$ sub-bands are reserved for spectrum handoff. Finally, perfect spectrum sensing is assumed as in [5], [6].

B. Traffic Assumptions

1) CL Traffic: The PUs and the SUs arrive into the system according to two independent Poisson processes with rates λ_{pcl} and λ_{scl} respectively. Their service times are assumed to be exponentially distributed with rates μ_{pcl} and μ_{scl} respectively [1], [2], [4], [5].

2) PL Traffic: The data traffic is represented by the MMPP that is a doubly stochastic Poisson process whose rate varies according to a Markov process. A MMPP with two states, namely, low and high, is used to represent the low and high data traffic burstiness patterns highlighted in the zoom-in shot illustrated in Fig. 1. This two-state MMPP is parameterized [10] by the infinitesimal generator $\mathbf{Q_s} = \begin{bmatrix} -\mu_l & \mu_l \\ \mu_h & -\mu_h \end{bmatrix}$ of the underlying Markov chain and the rate matrix $\mathbf{\Lambda_s} = \begin{bmatrix} \lambda_l & 0 \\ 0 & \lambda_h \end{bmatrix}$, where μ_l, μ_h, λ_l, and λ_h are respectively the transition rate from the state low to the state high, the transition rate from the state high to the state low, the data packet rate in the state low, and the data packet rate in the state high. This process is also characterized by the average data packet rate $\lambda_a = \lambda_l\mu_h + \lambda_h\mu_l/\mu_h + \mu_l$ and the degree of burstiness of the data source $\aleph = \lambda_h/\lambda_a$.

The data packet transmission times over the PRN and the CRN are assumed to be exponentially distributed random variables with means $1/\mu_p^{pu}$ and $1/\mu_p^{su}$ respectively. These random variables depend on the physical layer parameters such as modulation and coding schemes as well as the packet size S_d. Considering the current wireless systems (such as LTE and WiMAX) in which the physical layer data rates are in the order of Mbps and $S_d = 480$ bytes [7], we have $1/\mu_p^{pu} = S_d/1$ (Mbps) $\approx 10^{-3}$ s. Based on the fragmentation of the primary band in N sub-bands as illustrated in Fig. 1, the SU data packet transmission time over the CRN is given by $1/\mu_p^{su} = N/\mu_p^{pu}$. For simplicity, it is considered that no preemptive priority is applied over an ongoing PLSU transmission to drop it when the CAC/DSA authorizes its transmission. This assumption is rooted on the fact that the physical layer data rates follow an upward trend. Therefore, for modern systems, the data packet transmission time will be very short when compared against the CL service duration.

III. PERFORMANCE ANALYSIS

A. Markov Model

Let the tuple $(\mathcal{S}, \mathbf{Q})$ denote the proposed multi-dimensional CTMC model, where $\mathcal{S}$ is the state space and $\mathbf{Q}$ the infinitesimal generator. A state $i \in \mathcal{S}$ is defined as $(n_p, d_p, b_p, n_s, d_s, b_s)$, where $n_p \in \{0, 1, \ldots, M\}$, $d_p \in \{0, 1\}$, $b_p \in \{0, 1, \ldots, B_{pu}\}$, $n_s \in \{0, 1, \ldots, MN - L\}$, $d_s \in \{0, 1\}$, and $b_s \in \{0, 1, \ldots, B_{su}\}$ denote respectively the number of ongoing CLPU, the state of the PU data packet source (0 for low and 1 for high), the number of data packets in the PU buffer, the number of ongoing CLSUs, the state of the SU data packet source, and the number of data packets in the SU buffer. Let x_a^b be the x_a^{th} state variable of the state b, where a denotes either p or s. The transition rates in $\mathbf{Q}$ among a state i and its successors $j \in \mathcal{S}$ for $i \neq j$ are triggered by the following random events:

1) Arrival of a CLPU Service: If $n_p^i < M$, then $n_p^j = n_p^i + 1$, with rate λ_{pcl}. Moreover, n_s^j will be equal to $(M - n_p^i - 1)N$ if $n_s^i + n_p^i N > (M - 1)N$ or n_s^i otherwise [1], [2].

2) Departure of a CLPU Service: If $n_p^i > 0$, then $n_p^j = n_p^i - 1$ with rate $n_p^i \mu_{pcl}$.

3) Arrival of More Bursty and Less Bursty PLPU (resp. PLSU) Service: If $d_p^i = 0$ (resp. $d_s^i = 0$), the PU (resp. SU) data source will get more bursty with rate μ_l and $d_p^j = 1$ (resp. $d_s^j = 1$). On the other hand, if $d_p^i = 1$ (resp. $d_s^i = 1$), the PU (resp. SU) data source will get less bursty with rate μ_h and $d_p^j = 0$ (resp. $d_s^j = 0$).

4) Generation of a PU (resp. SU) Data Packet in Low State and High State: If $b_p^i < B_{pu}$ (resp. $b_s^i < B_{su}$) and $d_p^i = 0$ (resp. $d_s^i = 0$), then $b_p^j = b_p^i + 1$ (resp. $b_s^j = b_s^i + 1$) with rate λ_l. However, if $d_p^i = 1$ (resp. $d_s^i = 1$), then the PU (resp. SU) buffer occupancy will grow with rate λ_h.

5) Transmission of a PU Data Packet: If $b_p^i > 0$ and $n_p^i < M$, then $b_p^j = b_p^i - 1$ with rate $(M - n_p^i)\mu_p^{pu}$. Since the transmission of a data packet through the PRN implies the preemption of all in-progress CLSUs, we get $n_s^j = 0$.

6) Arrival of a CLSU Service: If $n_s^i + n_p^i N < MN - L$ and $b_p^i = 0$, then $n_s^j = n_s^i + 1$ with rate λ_{scl}. It should be noted that different from previous works, here, the CLSU admission is subject to the PRN and PU buffer occupancies.

7) Departure of a CLPU Service: If $n_s^i > 0$, then $n_s^j = n_s^i - 1$ with rate $n_s^i \mu_{scl}$.

8) Transmission of a SU Data Packet: If $b_s^i > 0$, $b_p^i = 0$, and $n_s^i + n_p^i N < MN$, then $b_s^j = b_s^i - 1$ with rate $(MN - n_s^i - n_p^i N)\mu_p^{su}$. It should be noted that considering the transmission hierarchy, a SU data packet in the SU buffer will only be transmitted provided that there is at least an available sub-band and no data packet in the PU buffer.

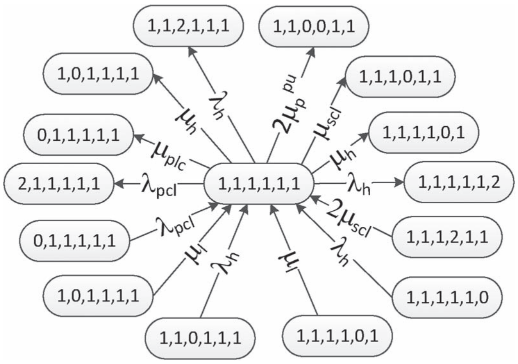

Fig. 2. Transitions from and to state $(1, 1, 1, 1, 1, 1) \in \mathcal{S}$ considering the setting described in Section IV.

Due to the complexity of the proposed multi-dimensional CTMC model, it is impractical to graphically represent a complete state transition diagram even for a small-scale system. Therefore, an example of a particular state $i = (n_p = 1, d_p = 1, b_p = 1, n_s = 1, d_s = 1, b_s = 1) \in \mathcal{S}$ and all possible transitions from and to state i are shown in Fig. 2. As it can be observed, the system moves from the state i to state (2, 1, 1, 1, 1, 1) or (0, 1, 1, 1, 1, 1) upon an arrival or a departure of a CLPU service with rate λ_{pcl} or μ_{pcl} respectively. In the state i, both MMPP data sources are in state high. Thus, with rate μ_h, the PU data source becomes less bursty and triggers a transition from state i to state (1, 0, 1, 1, 1, 1). While in state high, a PU data packet is generated with rate λ_h and joins the PU data buffer, moving the system to state (1, 1, 2, 1, 1, 1). When a PU data packet is transmitted with rate $2\mu_p^{pu}$, it preempts all in-progress CLSUs. In this case, the system transits to state (1, 1, 0, 0, 1, 1). By using the same approach, the remaining state transitions can be obtained similarly. Because the PU buffer is not empty in the state i, an arriving CLSU service request is rejected and the transmission of a SU data packet is not allowed.

B. SU Performance Metrics

Let $\pi_{i=(n_p,d_p,b_p,n_s,d_s,b_s)}$ be the steady state probability of the proposed multi-dimensional CTMC model. The CLSU blocking probability $P_B(L)$ for the system under analysis is defined in (1). This definition extends the one presented in [2] by considering the contributions to the $P_B(L)$ due to the PLPU service as specified in the summation in the 2nd term of (1)

$$P_B(L) = \sum_{i\in\mathcal{S}} \delta(n_s + n_p N \geq MN - L)\pi_i + \sum_{i\in\mathcal{S}} \delta(b_p > 0)\pi_i \tag{1}$$

where $\delta(\cdot)$ is the Dirac delta function. Similarly, the CLSU forced termination probability $P_F(L)$ given in (2), extends the result presented in [2]. It is worth mentioning that: 1) the summation in the 1st term of (2) expresses only the conditions under which the forced termination takes place. Notwithstanding, as specified in [2], the condition over the value of η in these states, i.e., $\delta(n_s - (M - n_p - 1)N - \eta)$ still prevails; 2) following the same reasoning, the summation in the 2nd term of (2) shows that all in-progress CLSUs are terminated when a PU data packet is transmitted over the PRN with rate $(M - n_p)\mu_p^{pu}$

$$P_F(L) = \frac{\sum_{i\in\mathcal{S}}\sum_{\eta=1}^{N} \eta\lambda_{pcl}\delta(n_s + n_p N > (M-1)N)\pi_i}{\lambda_{scl}(1 - P_B(L))} + \frac{\sum_{i\in\mathcal{S}} n_s(M - n_p)\mu_p^{pu}\delta(b_p > 0)\delta(n_s > 0)\pi_i}{\lambda_{scl}(1 - P_B(L))}. \tag{2}$$

Using (1) and (2), the CLSU throughput $T(L)$ (measured in connections per second) can be obtained [1], [2] as

$$T(L) = \lambda_{scl}\left(1 - P_B(L)\right)\left(1 - P_F(L)\right). \tag{3}$$

In terms of PLSU performance, (4) and (5) show the SU data packet blocking probability $P_{BD}(L)$ and the delay $D_D(L)$ respectively. It should be noted that the condition $\delta_1 = \delta(b_s > 0)\delta(b_p = 0)\delta(n_s + n_p N < MN)$ in (5) specifies that for a SU data packet to be transmitted, there must be a SU data packet in the SU buffer, no data packet in the PU buffer, and available sub-bands in the CRN

$$P_{BD}(L) = \sum_{i\in\mathcal{S}} \delta(b_s = B_{su})\pi_i \tag{4}$$

$$D_D(L) = \frac{\sum\limits_{i\in\mathcal{S}} b_s\delta(b_s > 0)\pi_i}{\sum_{i\in\mathcal{S}}(MN - n_s - n_p N)\mu_p^{su}\delta_1\pi_i}. \tag{5}$$

C. Analysis of the MMPP Data Source

To evaluate the system performance based on the MMPP parameters, it is assumed that $\mu_h = \Delta\mu_l$. By substituting it in λ_a, we obtain

$$\lambda_a = \frac{\lambda_h + \lambda_l\Delta}{1 + \Delta}. \tag{6}$$

Now, by re-writing $\aleph$ as a function of (6), we relate λ_h and λ_l in terms of the degree of burstiness as shown in (6), which facilitates the analytical assessment of the MMPP data source features on the system performance

$$\lambda_h = \frac{\Delta\aleph}{1 + \Delta - \aleph}\lambda_l. \tag{7}$$

IV. NUMERICAL RESULTS

For numerical computation, the same parameters used in [1] are considered: $\lambda_{pcl} = 0.1$, $\lambda_{scl} = 0.68$, $\mu_{pcl} = 0.06$, $\mu_{scl} = 0.82$, $M = 3$, and $N = 6$. For this setting, the correct values of $P_B(L = 2) = 0.1619$ and $P_F(L = 2) = 0.0367$ are given in [2]. These values are used in this work as baseline scenario to assess the impact of the PL traffic on the multi-service CRN. It should be noted that the baseline scenario takes into account the CLPUs and CLSUs only. The remaining parameters used in the experiments are $\Delta = 10$, $\mu_l = 10^{-3}$, and $B_{pu} = B_{su} = 20$.

The experiments were conducted considering that: both the CLPU and the CLSU service arrival rates are constant

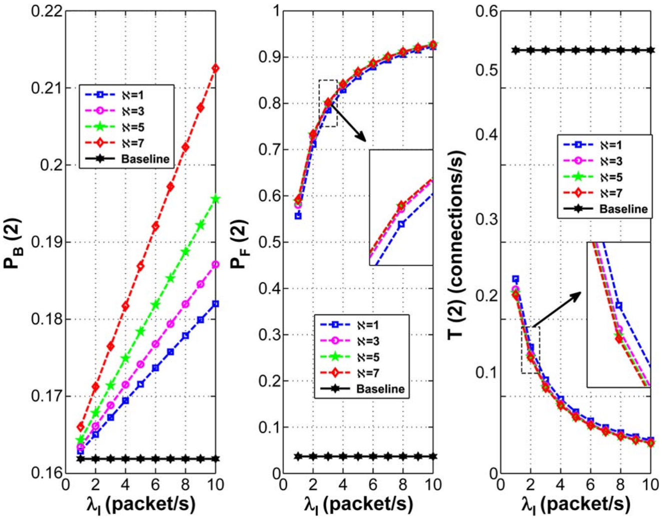

Fig. 3. Left: CLSU blocking probability, Middle: CLSU forced termination probability, Right: CLSU throughput versus λ_l for different ℵ degrees.

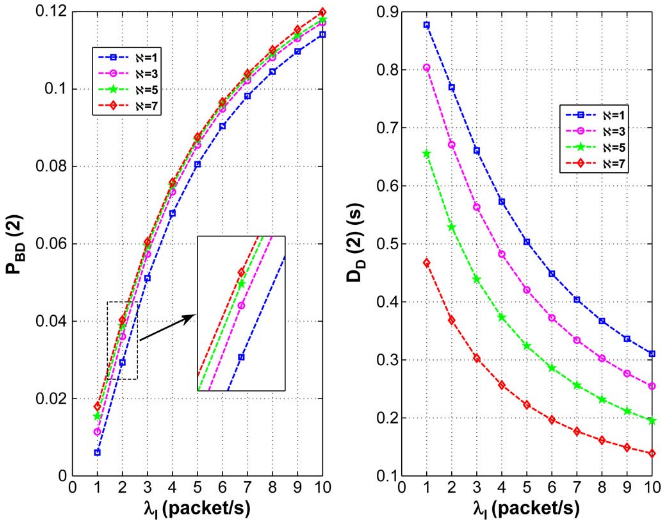

Fig. 4. Left: PLSU blocking probability and Right: PLSU delay versus λ_l for different ℵ degrees.

(according to the baseline scenario), the λ_l values for PLPUs and PLSUs vary in {1, 10}, and their respective λ_h values are given by (7). Under this configuration, Fig. 3 depicts how the data traffic deteriorates the CLSU QoS. First, it can be observed that there is a noticeable CLSU QoS degradation when the PLPU service is taken into account when compared with the baseline scenario. In a nutshell, it means that the single-service PRN assumption clearly overestimates the CLSU performance. Second, the degree of burstiness ℵ greatly compromises the $P_B(2)$. This is attributed to the fact that $P_B(L)$ is not only subject to the PRN conditions as in the baseline scenario, but is also heavily determined by the PU buffer occupancy. This way, the more bursty the data traffic gets (higher ℵ values), the higher the PU buffer occupancy becomes and consequently the worse the $P_B(2)$ will be. On the other hand, the effect of ℵ in $P_F(2)$ seems to be marginal. This is due to the fact that when a PU data packet is transmitted, it preempts all in-progress CLSUs. Consequently, the rate at which the packets come into the PU buffer is more important than the amount at which they join it. Therefore, $P_F(L)$ is more sensitive to the PLPU arrival rates than to the degree of burstiness ℵ. Thus, in terms of traffic burstiness, the Poisson assumption ($\aleph = 1$) is more critical for $P_B(L)$ than for $P_F(L)$. Third, an analysis of $T(2)$ reveals that there is a considerable reduction in the SUs' throughput caused by the high levels of $P_B(L)$ and $P_F(L)$. Ultimately, the assumption of multi-service PRN supporting the CL and PL services exposes the need of a much more careful DSA design to ensure the CLSUs' QoS provisioning. Fig. 4 depicts the PLSU's QoS against an increase of both PLPU and PLSU arrival rates and degree of burstiness. This way, as expected, $P_{BD}(2)$ also increases. In contrast, the delay perceived by the SU data packets is shortened despite the increase in the PLSU traffic profile. This occurs because the higher the PU's λ_l and ℵ are, the fewer the number of CLSUs kept into the system is (as shown previously). Therefore, there are more available sub-bands to be used by the PLSU traffic.

V. CONCLUSION

A multi-service PRN and CRN supporting the CL and PL services is studied in this letter. Under the predefined conditions, the numerical results have shown that PLPU may severely deteriorate the CLSUs' QoS mainly for high degree of burstiness. However, to ensure the attractiveness of CRN, it is crucial to ensure a minimum QoS level for the CLSUs. In this regard, for future works on DSA design, we are considering investigating some cross-layer design techniques and multi-cell cooperation in heterogeneous wireless networks as possible alternative ways to achieve that objective.

REFERENCES

[1] X. Zhu, L. Shen, and T.-S. P. Yum, "Analysis of cognitive radio spectrum access with optimal channel reservation," *IEEE Commun. Lett.*, vol. 11, no. 4, pp. 304–306, Apr. 2007.

[2] W. Ahmed, J. Gao, H. Suraweera, and M. Faulkner, "Comments on "Analysis of cognitive radio spectrum access with optimal channel reservation"," *IEEE Trans. Wireless Commun.*, vol. 8, no. 9, pp. 4488–4491, Sep. 2009.

[3] E. W. M. Wong and C. H. Foh, "Analysis of cognitive radio spectrum access with finite user population," *IEEE Commun. Lett.*, vol. 13, no. 5, pp. 294–296, May 2009.

[4] J. Wang, A. Huang, W. Wang, and T. Q. S. Quek, "Admission control in cognitive radio networks with finite queue and user impatience," *IEEE Wireless Commun. Lett.*, vol. 2, no. 2, pp. 175–178, Apr. 2013.

[5] V. K. Tumuluru, P. Wang, D. Niyato, and W. Song, "Performance analysis of cognitive radio spectrum access with prioritized traffic," *IEEE Trans. Veh. Technol.*, vol. 61, no. 4, pp. 1895–1906, May 2012.

[6] S. Tang and B. L. Mark, "Analysis of opportunistic spectrum sharing with Markovian arrivals and phase-type service," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 3142–3150, Jun. 2009.

[7] S.-R. Yang, S. -Y. Yan, and H.-N. Hung, "Modeling UMTS power saving with bursty packet data traffic," *IEEE Trans. Mobile Comput.*, vol. 6, no. 12, pp. 1398–1409, Dec. 2007.

[8] K. W. Choi and E. Hossain, "Opportunistic access to spectrum holes between packet bursts: A learning-based approach," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2497–2509, Aug. 2011.

[9] L. Muscariello, M. Mellia, M. Meo, M. A. Marsan, and R. L. Cigno, "Markov models of Internet traffic and a new hierarchical MMPP model," *Comput. Commun.*, vol. 28, no. 16, pp. 1835–1851, Oct. 2005.

[10] Y. Wu, G. Min, and L. T. Yang, "Performance analysis of hybrid wireless networks under bursty and correlated traffic," *IEEE Trans. Veh. Technol.*, vol. 62, no. 1, pp. 449–454, Jan. 2013.