The Impact of Stochastic Resource Availability on Cognitive Network Performance: Modeling and Analysis

Nadia Adem and Bechir Hamdaoui
School of Electrical Engineering and Computer Science
Oregon State University, Corvallis, Oregon 97331
Email:ademn,hamdaoui@eecs.oregonstate.edu

***Abstract*—Cognitive radio networks emerge as a promising solution for overcoming shortage and inefficient use of bandwidth resources by allowing secondary users (SUs) to access the primary users (PUs) channel so long as they do not interfere with them. The dynamical spectrum availability makes SU's packet average delay one of the most important performance measures of a cognitive network. It is important to understand the nature of delay, as well as its dependence on PU behaviors. In this paper, we analytically model and analyze the dynamics of the spectrum availability and their impact on the SU's packet delay. The cognitive network is modeled as a discrete-time queueing system. PU channel occupancy is modeled as a two-state Markov chain. Our contribution in this paper is defining and characterizing the properties of the random process that describes the availability of the opportunistic resources. In addition, we apply the mean residual service time concept to achieve an analytical solution for the queueing delay. Moreover, inspired by the slotted-Aloha system, we model the packets service mechanism, and determine the manner in which it depends on the resources availability. The delay becomes unbounded if the spectrum availability dynamics are not carefully considered in network design.**

I. Introduction

The increasing demand and usage of wireless technologies and applications are causing a spectrum shortage [1]. This shortage in spectrum supply is, however, shown to be due not to the scarcity of the spectrum resource, but rather to the current static allocation methods [2]. The dynamic spectrum access provided through cognitive radios is considered as a promising solution for overcoming this spectrum shortage problem [3]–[5]. In addition to primary users (PUs), which have the priority to access a number of communication channels, secondary users (SUs), which implement cognitive radios, may access the channels opportunistically [6]. That in turn improves spectrum utilization and gives opportunities to SUs to access licensed bands in an economical way. However, the lack of access priority could cause a drastic performance degradation to the cognitive users.

A. Motivation

The random availability of PUs channel leads to intermittent SU transmissions, which in turn affects delay performance [7]. It is important to analyze SU's average packet delay and determine the manner in which it depends on the randomly available resources.

The delay analysis has its consequences on cognitive network design. For a given PUs activity level, SUs might not be able to meet a certain performance criteria. The required transparency of the SUs activity could result in excessive packet drops or queue instability in buffered networks [8]. Hence, modifications need to be introduced to network settings to make sure it achieves the desired quality of service (QoS). The parameters that could be affected may include SUs' data rates, their numbers, their number of interfaces, their packet lengths, the number of channels they can access to, etc.

In this paper we are analyzing the performance of clustered cognitive network, where number of nodes along with a cluster head, which equips the cognitive radio, form a cluster. The cluster heads from different clusters contend to access the opportunistic resources. This model can very well apply to a cognitive radio sensor network, where the sensor nodes send their data to a sink, the cluster head in our model, that accesses to channels opportunistically. Sensor applications usually generate data in small rates and hence there is no need for acquiring a licensed band, and having opportunistic spectrum access can be enough to achieve a desired QoS.

B. Summary of Contributions

The complexity of analyzing cognitive networks delay performance and the broad aspects of such analysis seem to be the reason behind the area being not well investigated. The SUs need to adapt their operating conditions to the PUs which poss different transmission characteristics. Different delay components come to the picture as a result of that. A SU experiences a delay while identifying and exploiting spectrum access opportunities. To the best of our knowledge, there are no comprehensive delay models for cognitive networks in the literature. Hence, more thorough investigations need to be done in this area.

The availability of the opportunistic resources varies over time depending on PUs behavior. It turns out that different important cognitive network characteristics, e.g., the switching process parameters, are modeled analytically by establishing the model of the process that describes this availability. To the best of our knowledge, this research has never been addressed. In addition, most of the related work do not provide closed-form solutions for the average delay. Also, there is no much work about modeling SU's packets service mechanism, in spite of its importance in delay performance analysis.

In this paper, we analyze the performance of clustered cognitive radio network modeled as a discrete-time queueing system where the data queues up at the cluster head. The channel occupancy is modeled as a Markov chain. The contributions of this paper are summarized as follows.

- We characterize the properties of the random process that describes the opportunistically available resources. The properties of this process lead to the analytical characterization of the switching performed and the outage experienced by cognitive users.
- We apply the mean residual service time concept to derive the SUs' packet queueing delay for single as well as batch arrivals.
- Inspired by the slotted-Aloha system, we statistically characterize the packets service time distribution and hence the average service delay for single as well as multi-cluster networks. This delay results from the lack of access priority. The derived closed-form expression captures the dependence of this delay on the PUs behavior.
- By providing some numerical results, we show the importance of our analytical analysis in achieving a desired QoS and maintaining network stability.

II. Related Work

Due to its great potentials in addressing the spectrum shortage problem, the cognitive radio network paradigm has attracted significant research focus over the past decade, addressing various different aspects, such as protocol design [9]–[13], spectrum sensing [14]–[17], resource allocation and management [18]–[22], performance modeling and analysis [7], [23]–[28], and spectrum trading and auction [29]–[31], just to name a few. Delay performance analysis has also received some attention, but not as much [32]–[38]. In [32], the authors presented queuing analysis to study delay in cognitive networks. The authors in [32] obtain the solution of the queues average length for SUs that content to access PUs channel. The authors in [33] analyzed the stationary queue distribution for a constant SUs arrival process. They provided a closed-form expression for the distribution for the case of two channels, and upper and lower bounds for an arbitrary number of channels. In [34], the authors analyzed the delay for a clustered cognitive network. They considered the service time to be random and not following a standard distribution. They ended up analyzing the delay through approximating the average length of SUs queue size. The authors in [35] considered a cooperation between secondary and primary users to enhance the delay performance. In [36], the authors analyzed a cognitive network transmission delay by considering the distribution of time through which some opportunistic resources are available. In our paper, however, we characterize the properties of the process that describe the evolution of the resource availability over time. We use this process not only to understand the nature of delay, but also to obtain the analytical characterization of the switching mechanism and outage. In [37] the authors proposed centralized and distributed spectrum access schemes for SUs with different priority classes. The authors analyzed the blocking probability and average switching delay. The authors in [38] developed an admission control technique that guarantee a QoS requirements in terms of SUs' packets queueing delay. The authors assumed the availability of channels holds over a slot duration. In [39], the authors suggested a pricing strategy for reusing cellular networks spectrum. The cellular primary usage in [39] is modeled as a Poisson process. In our paper, we model the primary users behavior similarly. In [40], the authors used fluid flow models and effective bandwidth approximation to analyze the queueing system in a cognitive network. [41] presented the residual service time concept and applied it for continuous systems queueing analysis. The authors in [42] characterized the service time mechanism of a slotted-Aloha system with either finite or infinite user population where each user has finite or infinite buffer capacity.

In our paper, we capture the nature of the volatile spectrum in cognitive radio networks and its impact on the quality of service measured in terms of delay.

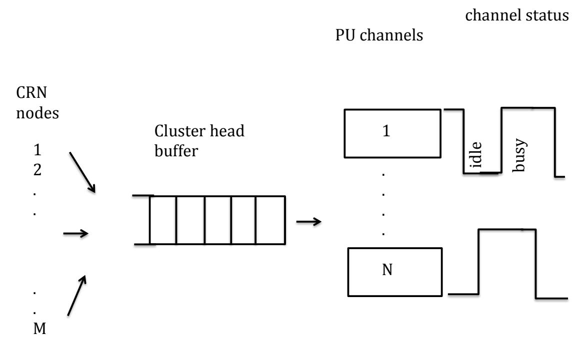

Fig. 1. A single cluster cognitive network

III. Network Model

The cognitive radio network (CRN) has access to N channels licensed to some PUs. The occupancy of each channel is modeled as a two-state Markov chain. We are considering a clustered network, where number of nodes along with a cluster head, which equips cognitive radio, form a cluster. There are L clusters, each contend with probability P_c to access the spectrum. Each cluster is modeled as a discrete-time queueing system where the data queues up at the cluster head buffer. Fig. 1 illustrates a single cluster network.

The cognitive radio system works as follows:

- The system is time slotted.
- The traffic arrives to the cluster head follows a Bernoulli process. The arrivals are independent of each other. Packets arrive in a batch within each slot. The average number of arrived packets per slot is λ, which can be viewed of as the arrival rate per slot.
- Over any available channel, the cluster head sends the data on a first-in first-serve basis. It switches from a spectrum to another whenever the last assigned becomes unavailable.
- The service times are assumed to be independent and identically distributed with an unspecified general distribution. The service process is assumed to be independent from the arrival process.

The slotted system assumption is a reasonable widely-used assumption, e.g. [32], [34], and [38], and many others consider a similar assumption. In addition to dedicating a part of the slot for packets transmission, another part is usually assumed to be dedicated to spectrum sensing and opportunities detection.

The Bernoulli arrival process assumption is more realistic than the Poisson process in our setting. Unless some sort of reservations is assumed, which might not be possible to achieve, assuming a Poisson arrivals for multi-packet messages is not reasonable.

In addition, our PU models applies to the users in cellular networks. It is popular to model the calls arrival as a Poisson process (i.e., exponentially distributed interarrival times), and the call durations probability distributions as exponential [39]. Successive interarrival times and call durations are independent of each other in this model.

The SUs' packet arrival process is independent from the service process. As we will explain in Section V, the service process depends mainly on the size of packets, contention methodology and the dynamics of the spectrum availability.

As the cluster head plays the role of identifying spectrum opportunities, we are not concerned with the delay that might result from sensing errors, as it remains low. This error is usually ignored when there is a central point involved in detecting spectrum opportunities (that is the case in [38] and [40] for example). Our work can serve as the basis for achieving other analytical delay models that include this delay.

Our model applies to cognitive radio sensor networks, where sensor nodes send their data to a sink, the cluster head in our model. The sensors send their data over unlicensed channel in a triggered based. In other words, once an event is sensed, sensors report to the cluster head. Sensor applications usually require large number of sensors to be implemented, each generates data in small rate. Hence, a scheduled access scheme might not be suitable and it suffices for nodes to randomly access the channel shared among them. Within each slot λ packets arrive the cluster head on average. Through reusing cellular bands, L number of clusters within the network transmit data, received from the nodes associated with them, to backbone network.

IV. Resource Availability Process

The availability of the opportunistic resources vary over time depending on the primary users activity and their spectrum usage pattern. In this section, we model and derive a number of statistics that describe the random availability of the resources. We model a number of random processes that are used to define the resource availability process and characterize its properties.

Single-Channel Availability Model: The evolution of the availability of a channel i over time is a random process CH_i. This process is a family of random variables $\{CH_i(t) : t \geq 0\}$, where each random variable takes a value zero if the channel is idle and one otherwise. By assumption, this process is modeled as a continuous-time Markov Chain with two states, labeled 0 and 1. The states 0 and 1 represent the idle and busy events, respectively. The transition time of the CH_i states, denoted by T_{CH_i}, is exponentially distributed with parameters u for the zero state and v otherwise. The state-transition diagram a channel occupancy appears in Fig 2(a). All the channels are assumed to be independent of each other and identical.. The analysis can be developed similarly if the channels are unidentical. Through out our analysis we are assuming the transition rates u and v are known. In practice, if they are not known, they can be estimated by observing the PUs behavior.

Multi-Channels Availability Model: Studying the process corresponding to the joint channels availability is important for our subsequent analysis. The joint availability process CH_{joint} is a family of vectors of random variables $\{(CH_1(t), \ldots, CH_N(t)) : t \geq 0\}$. Since at any time instant, the realization of CH_i, $\forall i \in \{1, .., N\}$, is either zero or one, the state space of CH_{joint} has a size of 2^N states. The transition time between the states is denoted by $T_{CH_{joint}}$.

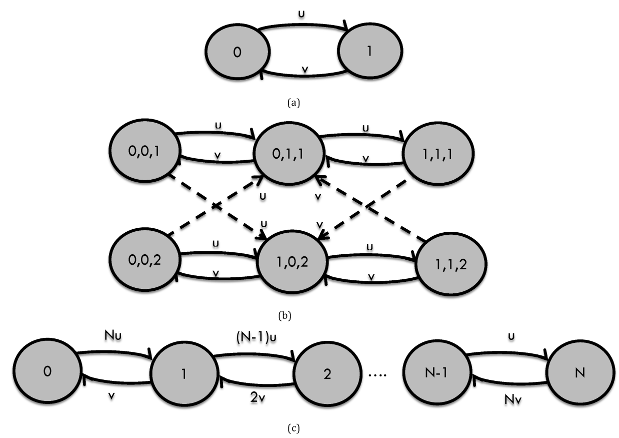

Fig. 2. (a) Single channel Markov chain. (b) Extended model of the two-channel availability. (c)Simultaneously occupied channels model.

Lemma 1. *The joint availability process is Markov.*

Proof. The state transition time for a given state $\{(CH_1 = ch_1, .., CH_N = ch_N)\}$, where $ch_i \in \{0, 1\}$ $\forall i \in \{1, .., N\}$, is given by $T_{CH_{joint}} = min\{T_{CH_1}, \ldots, T_{CH_N}\}$. Since T_{CH_i}, for $\forall i \in \{1, .., N\}$, is exponentially distributed with parameter $u^{1-ch_i} v^{ch_i}$, $T_{CH_{joint}}$ has also exponential distribution with parameter equals $\sum_{i=1}^{N} u^{1-ch_i} v^{ch_i}$.

□

A. Resource Availability Process Model

The resource availability process is basically a process that describes if there are any resources can be accessed by the cognitive network at any given time. Modeling this process is required for achieving the analytical analysis for some performance measures.

Definition 1. *The resource availability process for N-channel system ($N \geq 2$) is a process $CH = \{CH(t) : t \geq 0\}$ such that:* $CH(t) = CH_1(t)CH_2(t) \ldots CH_N(t)$.

At any instant of time, the value of $CH \in \{0, 1\}$. When CH equals zero, it means that there are some resources can be used by cognitive users. In other words, there is at least one PU channel idle. However, when CH equals one, that indicates all the channels are simultaneously busy and the cognitive network is going through an outage.

The process CH is not Markovian. The state zero in the state space corresponds to all states in the CH_{joint} state space except the state representing the event of all busy channels, i.e., the state with $CH_i = 1, \forall i \in \{1, .., N\}$. Hence, the transition time of the state zero, denoted by T_0, is a random sum of the transition times of the corresponding CH_{joint} states. As proven in Lemma 1, the probability distribution of the transition time of each CH_{joint} state is exponential, hence T_0 can not be exponential.

B. Resource Availability Process Statistics

The previously described models are used to derive some of the process statistics, which are important in making decisions concerning the applications that are admissible by the cognitive network. The cognitive users rate of switching, the cognitive network probability of outage, and the rate of outage, are important performance measures. It is of interest to obtain the analytical relationship between those statistics and the dynamics of the spectrum availability.

Switching Model: To find the rate of switching, denoted by r_{sw}, we extend the CH_{joint} chain such that each state is represented by $(CH_1, CH_2, \ldots, CH_N, n)$ where $n \in \{1, .., N\}$ indicates the last channel the cluster head was assigned to. For a two-channel system, we show in Fig. 2(b) the state-transition diagram for this extended Markov chain. The dashed arrows in the figure represent transitions that involve channel switching. By analyzing the chain for N-channel, we get

$$r_{sw} = \frac{u}{(u/v+1)^2} + \frac{(N-1)v}{(v/u+1)^N} \quad (1)$$

Where u is a channel idle-to-busy transition rate and v is a channel busy-to-idle transition rate.

Outage Model: The cognitive radio network outage rate, the outage probability, and the average outage time formulas can be derived by modeling the number of channels that are simultaneously busy. The evolution of the number of occupied channels over time is a random process $\{i(t) : t \geq 0\}$ where $i \in \{0, 1, \ldots, N\}$. This process is modeled as a Markov chain with $N+1$ states labeled 0 to N. The state label indicates the number of simultaneously occupied channels. The i^{th} state transition time distribution is exponential with parameter $iv + (N-i)u$ (the proof follows from the proof of Lemma 1). Fig. 2(c) shows the state-transition diagram for this chain.

The outage rate, denoted by r_{outage}, is defined as the rate at which all channels become simultaneously busy. It can be obtained by determining the rate at which the state N is visited. By analyzing the chain we obtain

$$r_{outage} = Nv/(1+v/u)^N \quad (2)$$

The outage probability, denoted by P_{outage}, is defined as the percentage of time during which no resources are available. That is the probability that the system is in the state N. It can be expressed as follows

$$P_{outage} = 1/(1+v/u)^N \quad (3)$$

The average outage time, denoted by $\overline{T}_{outage}$, is the average time spent in the state N. It is given by

$$\overline{T}_{outage} = 1/(Nv) \quad (4)$$

V. Delay Modeling and Characterization

In this section we are interested in analytically characterizing the average time required to deliver a packet within the cognitive radio network. We are considering two delay components, waiting and service delay.

The waiting delay is the time a packet spends at the queue until it starts being served. If a packet arrives to the system while there is a packet under service, the remaining of this service time is included in its waiting time. In addition, if a packet arrives while the queue is not empty, then the waiting time also includes the service time of all the packets ahead of it in the queue. In this section, we achieve the analytical solution for the expected waiting delay and show the way it is related to the service delay and hence to the opportunistically available resources.

The service delay is defined as the time between the instant the packet reaches the head of the queue to the instant it successfully departs the queue. If the cognitive network has access priority, it takes only one slot to serve a packet. The service time of any packet starts and ends at the slot boundaries. However, since the cluster head has only an opportunistic access to the channel, it takes integral (random) multiple of the slot duration to successfully transmit a packet. In this section, we determine the service time distribution and obtain analytically the manner in which the expected service time depends on the dynamics of the spectrum availability.

A. *Residual Service Time*

We derive the average waiting delay for our system using the service residual time concept. The concept of the mean residual service time has been considered for evaluating the performance of some continuous-time queueing systems [41]. However, to the best of our knowledge, it has not been considered for evaluating the performance of discrete-time systems. The analysis made for continuous-time systems can not be readily applied to discrete-time systems. In this section, we determine the mean residual service time for the discrete systems and use it to analyze the delay performance.

1) Residual Service Time Concept: An arrival to the system may experience some delay resulting from the residual service time of one of the packets arrived ahead of it. Let R_i denotes the residual service time seen by the i^{th} arrival. If the j^{th} packet is being served when the i^{th} packet arrives, then R_i corresponds to the remaining time until packet j completes its service. When packet i arrives while the system is empty, then R_i equals zero.

Fig. 3 illustrates by example the concept of residual time. In this figure we draw the number of arrivals and departures over time and show the residual service time corresponding to each arrival. X_i denotes the service time of the i^{th} arrival. t_i represents the time at which the i^{th} arrival arrives, and $t_i^{'}$ represents the time at which the i^{th} arrival leaves the system. The residual time can take a non zero value only at the instants at which an arrival occurs.

2) Residual Service Time in Discrete Systems: The evolution of the residual time over time is random, we showed a sample path for a simple example in Fig. 3. In continuous systems, the residual time can take a non zero value at any instants since an arrival can occur at any time. However, in discrete systems, the residual time can take a non zero value only at the slot boundaries. Also, since a service time is integral multiples of slot duration and it starts and ends at the slot boundaries, the remaining of a service time as seen by an arrival can only equal integral multiples of the slot duration.

Let the service time of the i^{th} arrival starts at the beginning of the k^{th} slot. Assume this service lasts for X_i slots. Let's refer to the residual time at the end of a slot k by r_k, where r_k is measured in slots. The residual times corresponding to the arrivals arrive during the service of the i^{th} arrival are denoted by $r_k, r_{k+1}, \ldots, r_{k+X_i-1}$. Their values are $X_i-1, X_i-2, \ldots, 1, 0$ slots, respectively. At the end of the first slot of the service time, the residual time is $X_i - 1$ slots, and its value decreases by one slot at the end of the next slot, and keeps doing so until the

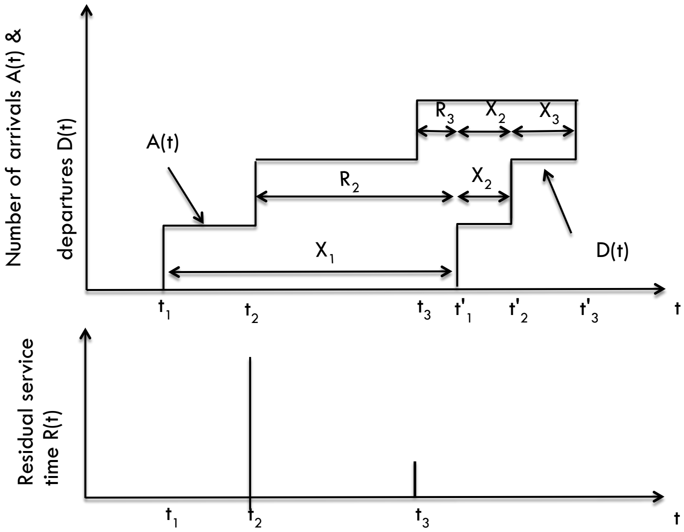

Fig. 3. The concept of residual service time

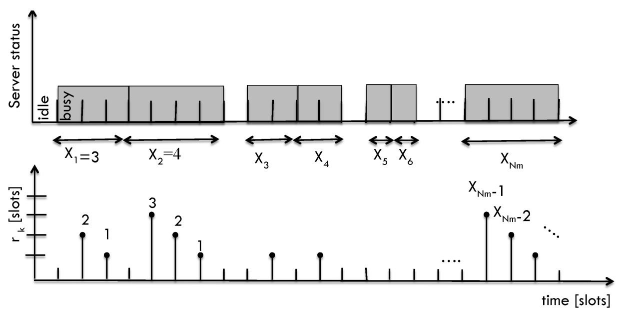

Fig. 4. A sample path of a server status and the corresponding residual service time.

service time completes. It equals zero at the end of the last slot of the service time. See Fig. 4 for illustration.

For an outside observer, any service corresponds to an arrival arrives right prior to the start of the service. The service that starts at the k^{th} slot corresponds to an arrival arrives at the begining of that slot. Since one packet at most can arrive at any given time slot, no other arrivals can arrive at this particular arrival instant. The residual times at the beginning of any slot at which a service starts is zero. See Fig. 4 for illustration.

3) Mean Residual Service Time: According to [41] , the mean residual time as seen by an arrival is equal to the mean residual time seen by an outside observer at a random time. This is valid for any arrivals satisfying the Poisson Arrivals See Time Averages (PASTA) property, which is the case for the queueing systems with Poisson arrival process. The question arises here is what about the Bernoulli arrivals system. Since the Bernoulli Arrivals See Time Averages (BASTA) property for those systems is analogous to the PASTA property in continuous-time systems, we can also define the mean residual time as seen by an arrival (denoted by $\overline{R}$) to be the mean residual time seen by an outside observer at a random time. We use a graphical argument to obtain $\overline{R}$. The analysis we make applies for single and batch arrivals. The residual time of an arrival depends on the arrival instant and not on the number of arrivals arrive at that instant.

The cluster head status over time is random. During any time slot, it could be either busy serving a packet or idle. When a packet i starts to be served, the cluster head stays busy for X_i slots. In Fig. 4, we plot a sample path of the cluster head status. We also plot the corresponding residual service time sample path, which we use to obtain the time average of the residual service time. Consider the time interval $[0,\tau]$, where τ is the time instant corresponding to the end of the m^{th} slot. We are assuming that up to the m^{th} slot, N_m packets have already been served. The time average of the residual time (measured in slots) in this interval is given by $E_m = \frac{1}{m}\sum_{k=1}^{m} r_k$.

Since we know the values of r_k's during the service time of each packet (as we explained earlier), the sum of the r_k over the m slots can be determined by summing the r_k's corresponding to the service times. The average time of the residual time can then be rewritten as

$$E_m = \frac{1}{2}\frac{N_m}{m}\left(\frac{\sum_{i=1}^{N_m} {X_i}^2}{N_m} - \frac{\sum_{i=1}^{N_m} X_i}{N_m}\right)$$

Taking the limit as $m \to \infty$, assuming it exists, we obtain

$$\lim_{m\to\infty} E_m = \frac{1}{2}\lim_{m\to\infty}\frac{N_m}{m}\lim_{m\to\infty}\left(\frac{\sum_{i=1}^{N_m} {X_i}^2}{N_m} - \frac{\sum_{i=1}^{N_m} X_i}{N_m}\right)$$

The left-hand side limit is the time average of the residual time. The limits on the right-hand side are the departure rate (which equals the arrival rate), the service time second and first moments respectively. Assuming that the time averages can be replaced by the ensemble averages, the average residual time can then be expressed as

$$\overline{R} = \frac{1}{2}\lambda(\overline{X^2} - \overline{X}) \quad (5)$$

where $\overline{X}$ and $\overline{X^2}$ denote the service time first and second moment respectively.

B. Waiting Delay

We derive the average waiting delay for our system in terms of the average service residual time.

1) Single Arrival Systems: The per-packet average waiting time $\overline{W}$ can be expressed in terms of the average residual time as $\overline{W} = \overline{R}/(1-\rho)$, where $\rho = \lambda\overline{X}$ is the utilization factor [41]. ρ should be less than unity for a stable system [43] and λ is the arrival rate per slot. Replacing $\overline{R}$ with its expression presented in Equation (5) yields

$$\overline{W} = \frac{1}{2}\frac{\lambda(\overline{X^2} - \overline{X})}{(1-\rho)} \quad (6)$$

2) Batch Arrival Systems: The average waiting time of an arbitrary chosen packet in batch arrival systems is consisting of two independent components. One is the average waiting time of the batch that the packet belongs to, $\overline{W_b}$. The other is the average waiting time within the batch $\overline{W_w}$. The average waiting time $\overline{W_b}$ is the same as the average waiting time of the first packet in the batch. $\overline{W_b}$ equals average residual time of the first packet arrive in the bach plus the average service time of all the packets ahead of the batch in the queue. $\overline{W_b}$ can be expressed as

$$\overline{W_b} = \frac{\overline{R} + \rho\overline{W_w}}{(1-\rho)} \quad (7)$$

Denote by A the batch size. The first moment and second moment of A are denote by λ, and $\overline{\lambda^2}$ respectively. For a fixed batch size a, the average waiting of a packet within a batch is given by $\overline{X}\frac{(a^2-a)}{2a}$. The probability that an arbitrary chosen packet is in a batch of size a is expressed as aP_a/λ, where P_a is the probability that a batch has a size a. Therefore, $\overline{W_w}$ for an arbitrary packet is expressed as

$$\overline{W_w} = \frac{1}{2}\overline{X}(\overline{\lambda^2} - \lambda) \tag{8}$$

From Equation (7) and (8), the per-packet mean waiting time of a batch arrival system can be written as

$$\overline{W} = \frac{\lambda^2(\overline{X^2} - \overline{X}) + \overline{X}(\overline{\lambda^2} - \lambda)}{2\lambda(1-\rho)} \tag{9}$$

C. Service Delay

The service time distribution is a prerequisite for analyzing the delay performance. The analytic solution of the expected waiting delay given in Equations (6) and (9) involves both the first and second moments of the service time. Delay analysis can still be made if the service time distribution is not realized. However, the exact analysis appears to be very difficult. Depending on the model of the system under consideration, the service time can turn out to be not following any standard distribution. Let's assume that a channel needs to be available for an S amount of time continuously so that a packet can be transmitted. Let's also assume that the cluster head starts to serve packets whenever there is a channel available. It is possible that a cluster head starts to serve a packet and then before it completes its transmission, the channel gets occupied by a PU. This could happen many times in a random manner. This causes the service time to be random and not following any standard distribution.

Inspired by the slotted-Aloha system presented in [41] and [42], we make the following arguments. Let S denotes the slot duration. Corresponding to our model, at any given slot the cluster head transmits a packet ready for service if there is idle channel. Given that the availability of spectrum is time-variant with some probability channel remains available over the entire slot duration and transmission succeeds. If transmission failed, cluster head retransmits (with probability P_c in case there are number of clusters contend for channels) the packet in the successive slot until transmission succeeds. Denote by μ the probability that the time spent in serving a packet is one slot only. μ can be viewed as the service rate per slot. The service time (measured in slots) needed by the cluster head to successfully transmit a packet is geometric random variable with parameter μ. We derive the expression of μ for the single and multi-cluster systems.

1) Single Cluster Systems: A packet transmission is successful within a time slot if there is at least one channel available during at least the slot duration. The probability of transmission success can be written as $\mu = Pr\{no\ outage\}Pr\{channel\ idle\ time > S\}$.

Using Equation (3) which gives the probability of no cognitive network outage and considering the exponential distribution of channel idle time, we obtain

$$\mu = (1 - \frac{1}{(1+v/u)^N})e^{-uS} \tag{10}$$

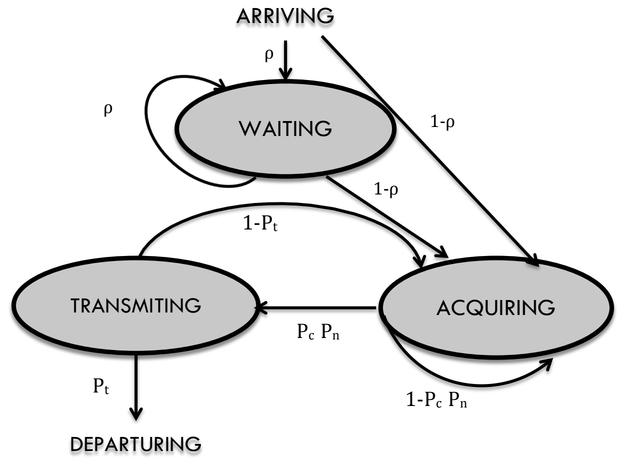

Fig. 5. Arrivals state transition diagram

The packet average delay (denoted by $\overline{T}$) consists of the average waiting delay and the service time average delay. From Equation (6), $\overline{T}$ for the single arrival system can be written as

$$\overline{T} = \overline{X} + \frac{1}{2}\frac{\lambda(\overline{X^2} - \overline{X})}{(1-\lambda\overline{X})} \tag{11}$$

2) Multi-cluster Systems: Clusters contend with probability P_c to access primary user channels in multi-cluster system. We illustrate in Fig. 5 arrivals state transition diagram for single-arrival multi-cluster system. At any given slot, a packet at cluster head queue waits (i.e., it is in the state labeled WAITING in Fig. 5) for service with probability ρ, which is the probability that the cluster head is loaded. When the packet reaches the head of queue (i.e., it makes transition to state ACQUIRING), with probability P_n the cluster head acquires an idle channel. P_n is the probability that there is no outage. The cluster head contends over the acquired channel with probability P_c. The cluster transmits the packet with probability P_nP_c (makes transition to state TRANSMITTING). With probability P_t the transmission succeed and the packet leaves the system. P_t is the probability that no collision occurred over the acquired channel and no primary user reclaims the channel usage right. The average service time $\overline{X}$ corresponds to the average time spent in ACQUIRING and TRANSMITTING states. $\overline{X}$ is given by $1/\mu$. Denote the number of clusters by L. μ can be expressed as

$$\mu = e^{-uS}\Big(\frac{u/v}{1+u/v}\Big)^N (1-P_c)^L \sum_{l=1}^{L} \binom{L-1}{l} \Big(\frac{P_c}{1-P_c}\Big)^l$$
$$\sum_{i=1}^{N} \binom{N}{i} (u/v)^i \Big(\frac{i-1}{i}\Big)^{l-1} \tag{12}$$

VI. Performance Evaluation and Analysis

In this section, we numerically analyze the impact of the PU behaviors on the statistics of the resource availability process. We also measure the delay performance and study its dependence on the dynamics of the spectrum availability. For convenience,

TABLE I
DESCRIPTIONS OF FREQUENTLY USED SYMBOLS

Parameter	Description
u	A channel idle-to-busy transition rate
v	A channel busy-to-idle transition rate
N	Number of channels
r_{outage}	Outage rate
$\overline{T}_{idle}$	Average channel idle interval
$\overline{T}_{busy}$	Average channel busy interval
λ	Arrival rate per slot
S	Time slot duration

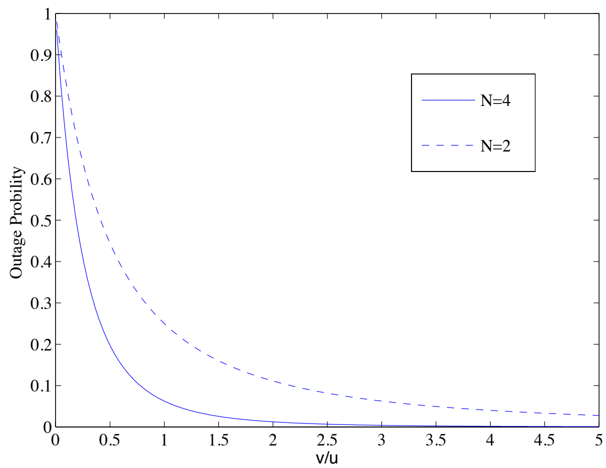

Fig. 6. Outage probability vs $\overline{T}_{idle}/\overline{T}_{busy}$

a reference of used network parameters and their descriptions is given in Table I.

A. Resource Availability Process Statistics Analysis

In Fig. 6, we plot the probability of outage (Equation (3)) versus the ratio between the average channel idle time to the busy time. Here we refer to the average channel idle and busy interval by $\overline{T}_{idle}$ (which equals $1/u$, where u is the PU idle to busy transition rate) and $\overline{T}_{busy}$ (which equals $1/v$, where v is the PU busy to idle transition rate) respectively.

As the ratio $\overline{T}_{idle}/\overline{T}_{busy}$ increases, the outage probability decreases. When this ratio is much less than one, the outage probability can be close to one. Also, as the number of channels N increases, the outage probability decreases. Increasing the number of channels gives the cognitive network more chances to find an idle one. However, as $\overline{T}_{idle}/\overline{T}_{busy}$ increases, the outage occurs with less probability and the effect of having more channels on this probability becomes less. This is an important observation to consider when it comes to making decision about the number of channels a cluster head needs to be able to access to.

Fig. 7 plots the effect of the PUs activity on the outage rate, as expressed by Equation (2). The observation we make here is when $\overline{T}_{idle}$ is low ($1/u < 1$) and $\overline{T}_{idle}$ is less than $\overline{T}_{busy}$ (i.e., $v < u$), as the value of $\overline{T}_{busy}$ decreases, so does the outage rate r_{outage}. However, for large $\overline{T}_{idle}$, as the $\overline{T}_{busy}$ increases, the r_{outage} decreases. One thinks that the lower the value of the $\overline{T}_{busy}$ is, the better the performance. For example, if the

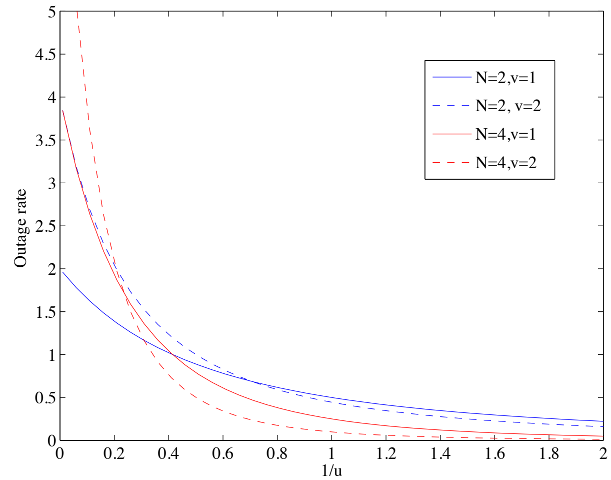

Fig. 7. Outage rate vs $\overline{T}_{idle}$

PUs are a cellular network users, one thinks the smaller the average call duration, the better the SUs performance. However, Fig. 7 indicates for a large call interarrival time, the longer the call duration, the less the outage rate. The trend for the outage probability change is different though, as shown Fig. 6. In other words, as the calls durations increases, the outage probability decreases. That should make sense since the outage rate illustrates how often the network goes through outage, but does not tell how long the outage is or what its probability is. The different changing trends shows the importance of considering all the different resource availability statistics if one desires to determine if an application is admissible by the cognitive network.

The switching rate is also an important parameter to consider when it comes to designing a multi-channel cognitive network. In practice, there is a communication overhead and energy consumption associated with switching. The graph of switching rate behavior, as described by Equation (1), versus $\overline{T}_{idle}$ is shown in Fig. 8. An important conclusion we point out from this figure is that the impact of the number of channel on the performance varies depending on the value of $\overline{T}_{idle}$. For a given v value (i.e., $\overline{T}_{busy}$), when $u > 1$ ($\overline{T}_{idle}$ is small) and $\overline{T}_{idle}$ is less than $\overline{T}_{busy}$ (i.e., $v < u$), the larger the number channels N, the higher the switching rate. This is because when the time in which a channel remains idle gets smaller on average, the cluster head needs to switch across the channels more often. As $\overline{T}_{busy}$ increases, the switching rate decreases. It might seem appealing to decrease the switching rate, however, as v and N decrease, the outage probability increases. In fact, the increase in the outage probability is the reason behind the decrease in the switching rate as no switching occurs during the outage. Similar to the observation we made about the outage rate, the trend of switching rate change is different than that for the outage probability. This confirms our conclusion about the importance of considering all these statistics for making designing decisions concerning the cognitive network.

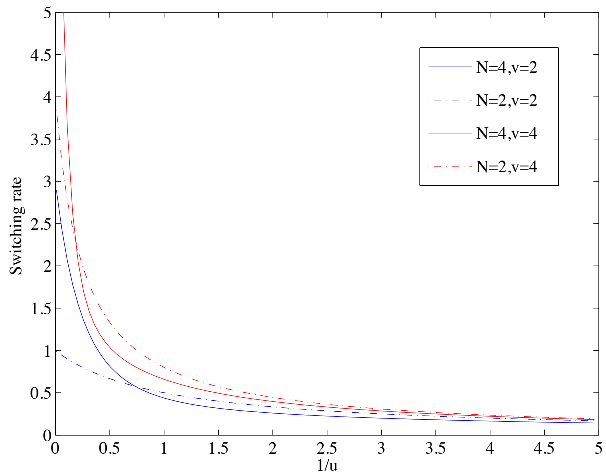

Fig. 8. Switching rate vs $\overline{T}_{idle}$

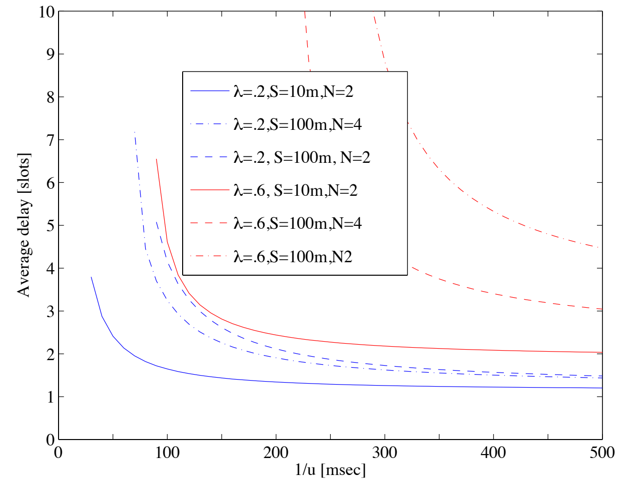

Fig. 10. Average delay vs average of channel availability time ($\overline{T}_{idle}$)

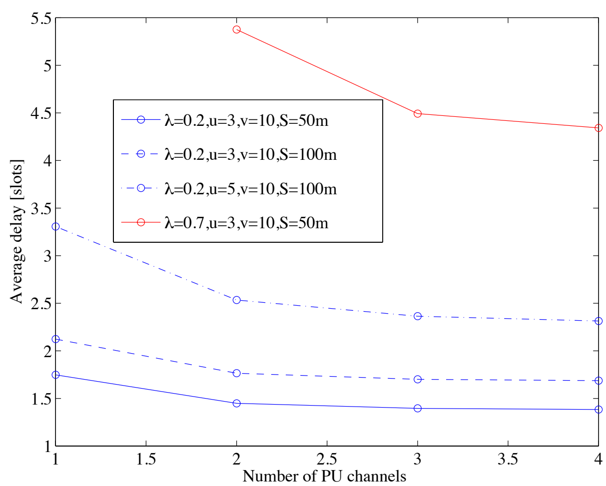

Fig. 9. Average delay vs. number of PU channels

B. Delay Analysis

For a given SUs' and PUs' traffic parameter, we plot in Fig. 9 the delay performance for the single-arrival single-cluster system (Equation (11)) versus number of channels. The average delay decreases as number of channels increases, which is intuitive. However, how fast delay decreases and what value the delay converges to depend on the slot duration, the SUs' packet arrival rate, and the PUs usage of the channels reflected via the rate at which a channel make transition from idle to busy (u) and from busy to idle (v).

As we explained earlier, as u increases, the outage probability increases. Also, the probability that a channel remains idle for an entire slot duration S decreases. Hence, the average service rate decreases. Since the cluster head uses only one channel at any given time to send data, having more channels while u is large does not improve the performance considerably.

Note that in this and all the subsequent figures, if the results are not shown for a range of values within the horizontal axis, this implies that the delay is unbounded. For a given network setting, if the cluster head can not keep up with the arrival rate, in other words service rate is less than the arrival rate, the network becomes unstable and the delay increases without bounds.

Fig. 10 illustrates how Equation (11) behaves when $\overline{T}_{idle}$ changes. As $\overline{T}_{idle}$ increases, the performance improves. However, the significance of the improvement depends on some other network settings. We observe from this figure that increasing the number of channels for highly loaded network has more impact on the delay performance than the lightly loaded. Similarly, the impact of increasing the slot duration becomes more severe as the value of λ increases. Increasing the slot duration decreases the probability that a channel can remains idle over the slot duration. Hence, the average service time increases, and so consequently does the average waiting time, especially for large values of λ. The range of $\overline{T}_{idle}$ through which the system is unstable varies depending on the primary and cognitive network settings. In order to maintain the network stability, it is necessary to design the cognitive network such that the traffic arrival rate does not exceed the service rate, which is as expressed in Equation (10) is a function of primary users' traffic parameter.

VII. Conclusions and Future Work

In this paper, we have analyzed the dynamics of the spectrum availability and studied the delay performance of a clustered cognitive network. We have introduced the concept of the resource availability process and characterized its properties as a way to measure the network performance. We have also obtained the analytical characterization of the relationship between the packets delay and the dynamics of the spectrum availability. The opportunistically available resources need to be carefully considered in making design decisions regarding the cognitive network to maintain its stability. Studying the gain of the multiple-interface transmission on the performance for both discrete-time and continuous-time systems is underway.

Acknowledgment

This work was supported in part by the US National Science Foundation under NSF award CNS-1162296.

REFERENCES

[1] F. S. P. T. Force, "Report of the spectrum efficiency working group," 2002.

[2] D. Datla, A. M. Wyglinski, and G. J. Minden, "A spectrum surveying framework for dynamic spectrum access networks," *Vehicular Technology, IEEE Transactions on*, vol. 58, no. 8, pp. 4158–4168, 2009.

[3] B. Hamdaoui, "Optimal discovery of bandwidth opportunities in spectrum agile networks," in *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*. IEEE, 2008, pp. 1–5.

[4] P. Venkatraman, B. Hamdaoui, and M. Guizani, "Opportunistic bandwidth sharing through reinforcement learning," *Vehicular Technology, IEEE Transactions on*, vol. 59, no. 6, pp. 3148–3153, 2010.

[5] B. Hamdaoui, "Adaptive spectrum assessment for opportunistic access in cognitive radio networks," *Wireless Communications, IEEE Transactions on*, vol. 8, no. 2, pp. 922–930, 2009.

[6] B. Hamdaoui and K. G. Shin, "OS-MAC: An efficient MAC protocol for spectrum-agile wireless networks," *Mobile Computing, IEEE Transactions on*, vol. 7, no. 8, pp. 915–930, 2008.

[7] M. NoroozOliaee, B. Hamdaoui, X. Cheng, T. Znati, and M. Guizani, "Analyzing cognitive network access efficiency under limited spectrum handoff agility," *Vehicular Technology, IEEE Transactions on*, vol. 63, no. 3, pp. 1402–1407, 2014.

[8] N. Adem and B. Hamdaoui, "Delay performance modeling and analysis in clustered cognitive radio networks," in *Global Communications Conference (GLOBECOM), 2014 IEEE*, Dec 2014, pp. 193–198.

[9] S. Ehsan, B. Hamdaoui, and M. Guizani, "Radio and medium access contention aware routing for lifetime maximization in multichannel sensor networks," *Wireless Communications, IEEE Transactions on*, vol. 11, no. 9, pp. 3058–3067, 2012.

[10] R. Hamdi, M. Ben Ghorbel, B. Hamdaoui, and M. Guizani, "Design and implementation of distributed dynamic spectrum allocation protocol," in *Communications Workshops (ICC), 2014 IEEE International Conference on*. IEEE, 2014, pp. 274–278.

[11] C. Cormio and K. R. Chowdhury, "A survey on mac protocols for cognitive radio networks," *Ad Hoc Networks*, vol. 7, no. 7, pp. 1315–1329, 2009.

[12] M. Maiya and B. Hamdaoui, "imac: improved medium access control for multi-channel multi-hop wireless networks," *Wireless Communications and Mobile Computing*, vol. 13, no. 11, pp. 1060–1071, 2013.

[13] A. Sivanantha, B. Hamdaoui, M. Guizani, X. Cheng, and T. Znati, "Emmac: An energy-aware multi-channel mac protocol for multi-hop wireless networks," in *Wireless Communications and Mobile Computing Conference (IWCMC), 2012 8th International*. IEEE, 2012, pp. 1159–1164.

[14] K. Arshad and K. Moessner, "Robust collaborative spectrum sensing based on beta reputation system," in *Future Network & Mobile Summit (FutureNetw), 2011*. IEEE, 2011, pp. 1–8.

[15] J. Shen, T. Jiang, S. Liu, and Z. Zhang, "Maximum channel throughput via cooperative spectrum sensing in cognitive radio networks," *Wireless Communications, IEEE Transactions on*, vol. 8, no. 10, pp. 5166–5175, 2009.

[16] M. Grissa, A. Yavuz, and B. Hamdaoui, "LPOS: locaion privacy for optimal sensing in cognitive radio networks," in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*. IEEE, 2015, pp. 1–6.

[17] O. Altrad, S. Muhaidat, A. Al-Dweik, A. Shami, and P. Yoo, "Opportunistic spectrum access in cognitive radio networks under imperfect spectrum sensing," *Vehicular Technology, IEEE Transactions on*, vol. 63, no. 2, pp. 920–925, 2014.

[18] M. NoroozOliaee, B. Hamdaoui, and K. Tumer, "Achieving optimal elastic traffic rewards in dynamic multichannel access," in *High Performance Computing and Simulation (HPCS), 2011 International Conference on*. IEEE, 2011, pp. 155–161.

[19] M. NoroozOliaee and B. Hamdaoui, "Distributed resource and service management for large-scale dynamic spectrum access systems through coordinated learning," in *Wireless Communications and Mobile Computing Conference (IWCMC), 2011 7th International*. IEEE, 2011, pp. 522–527.

[20] O. Alsaleh, P. Venkatraman, B. Hamdaoui, and A. Fern, "Enabling opportunistic and dynamic spectrum access through learning techniques," *Wireless Communications and Mobile Computing*, vol. 11, no. 12, pp. 1497–1506, 2011.

[21] M. Ben Ghorbel, B. Hamdaoui, R. Hamdi, M. Guizani, and M. NoroozOliaee, "Distributed dynamic spectrum access with adaptive power allocation: Energy efficiency and cross-layer awareness," in *Computer Communications Workshops (INFOCOM WKSHPS), 2014 IEEE Conference on*. IEEE, 2014, pp. 694–699.

[22] T. Alshammari, B. Hamdaoui, M. Guizani, and A. Rayes, "Malicious-proof and fair credit-based resource allocation techniques for dsa systems," *Wireless Communications, IEEE Transactions on*, vol. 14, no. 2, pp. 606–615, 2015.

[23] S. Ehsan, B. Hamdaoui, and M. Guizani, "Feasibility conditions for rate-constrained routing in power-limited multichannel wsns," in *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*. IEEE, 2011, pp. 1–6.

[24] O. Filio-Rodriguez, S. Primak, V. Kontorovich, and A. Shami, "On a cognitive radio networks random access game with a poisson number of secondary users," *IEEE Communications Letters*, 2015.

[25] B. Hamdaoui, K. G. Shin, and M. Maiya, "Constraint design and throughput evaluation in multi-band wireless networks using multiple-input multiple-output links," *IETE Technical Review*, vol. 26, no. 2, pp. 101–107, 2009.

[26] L.-C. Wang, C.-W. Wang, and C.-J. Chang, "Modeling and analysis for spectrum handoffs in cognitive radio networks," *Mobile Computing, IEEE Transactions on*, vol. 11, no. 9, pp. 1499–1513, 2012.

[27] M. Elmachkour, A. Kobbane, E. Sabir, J. Ben-othman *et al.*, "Data traffic-based analysis of delay and energy consumption in cognitive radio networks with and without resource reservation," *International Journal of Communication Systems*, vol. 28, no. 7, pp. 1316–1328, 2015.

[28] M. NoroozOliaee, B. Hamdaoui, T. Znati, and M. Guizani, "Forced spectrum access termination probability analysis under restricted channel handoff," in *Wireless Algorithms, Systems, and Applications*. Springer Berlin Heidelberg, 2012, pp. 358–365.

[29] M. Elmachkour, I. Daha, E. Sabir, A. Kobbane, and J. Ben-Othman, "Green opportunistic access for cognitive radio networks: A minority game approach," in *Communications (ICC), 2014 IEEE International Conference on*. IEEE, 2014, pp. 5372–5377.

[30] M. A. Taleghan and B. Hamdaoui, "Efficiency-revenue optimality tradeoffs in dynamic spectrum allocation," in *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*. IEEE, 2010, pp. 1–5.

[31] O. F. Rodriguez, S. Primak, V. Kontorovich, and A. Shami, "A game theory interpretation for multiple access in cognitive radio networks with random number of secondary users," *arXiv preprint arXiv:1305.5222*, 2013.

[32] S. Wang, J. Zhang, and L. Tong, "Delay analysis for cognitive radio networks with random access: A fluid queue view," in *INFOCOM, 2010 Proceedings IEEE*. IEEE, 2010, pp. 1–9.

[33] A. Laourine, S. Chen, and L. Tong, "Queuing analysis in multichannel cognitive spectrum access: A large deviation approach," in *INFOCOM, 2010 Proceedings IEEE*. IEEE, 2010, pp. 1–9.

[34] S. Feng and D. Zhao, "Supporting real-time cbr traffic in a cognitive radio sensor network," in *Wireless Communications and Networking Conference (WCNC), 2010 IEEE*. IEEE, 2010, pp. 1–6.

[35] K. Zheng, J. Luo, J. Zhang, W. Wu, X. Tian, and X. Wang, "Cooperation improves delay in cognitive networks with hybrid random walk."

[36] Z. Liang, S. Feng, D. Zhao, and X. S. Shen, "Delay performance analysis for supporting real-time traffic in a cognitive radio sensor network," *Wireless Communications, IEEE Transactions on*, vol. 10, no. 1, pp. 325–335, 2011.

[37] V. K. Tumuluru, P. Wang, D. Niyato, and W. Song, "Performance analysis of cognitive radio spectrum access with prioritized traffic," *Vehicular Technology, IEEE Transactions on*, vol. 61, no. 4, pp. 1895–1906, 2012.

[38] M. M. Rashid, M. J. Hossain, E. Hossain, and V. K. Bhargava, "Opportunistic spectrum access in cognitive radio networks: A queueing analytic model and admission controller design," in *Global Telecommunications Conference, 2007. GLOBECOM'07. IEEE*. IEEE, 2007, pp. 4647–4652.

[39] A. Al Daoud, M. Alanyali, and D. Starobinski, "Secondary pricing of spectrum in cellular cdma networks," in *New Frontiers in Dynamic Spectrum Access Networks, 2007. DySPAN 2007. 2nd IEEE International Symposium on*. IEEE, 2007, pp. 535–542.

[40] Z. Fan, "Performance analysis of dynamic spectrum access networks," in *Pervasive Computing and Communications, 2008. PerCom 2008. Sixth Annual IEEE International Conference on*. IEEE, 2008, pp. 336–341.

[41] D. P. Bertsekas, R. G. bertsekas1992data, and P. Humblet, *Data networks*. Prentice-Hall International New Jersey, 1992, vol. 2.

[42] T. Wan and A. U. Sheikh, "Performance and stability analysis of buffered slotted aloha protocols using tagged user approach," *Vehicular Technology, IEEE Transactions on*, vol. 49, no. 2, pp. 582–593, 2000.

[43] H. Takagi, "Queueing analysis. discrete-time systems, vol. 3," 1993.