Performance of a Cognitive Radio Network with Tolerable Service Degradation

Shensheng Tang
Dept. of Engineering Technology
Missouri Western State University
St. Joseph, Missouri

Brian L. Mark
Dept. of Electrical and Computer Eng.
George Mason University
Fairfax, Virginia

Abstract—We analyze the performance of a cognitive radio wireless network, where secondary users opportunistically share the radio spectrum with primary users through spectrum sensing and may cause service degradation to the primary users due to unreliable spectrum sensing. When a secondary user occupying a channel detects the presence of a primary call arriving to the channel, it switches from its current channel to another channel, if one is available, to continue its service; however, if the secondary user fails to detect the presence of the primary call, it remains on the channel and both calls will receive degraded service. Based on a queueing network model, we derive closed-form solutions for the equilibrium system state in terms of generating functions, and derive several performance metrics of interest. In particular, we introduce the system supportability metric to evaluate the extent to which the system can support secondary users and the interference factor to evaluate the degree of service degradation imposed on primary users by the secondary users. Numerical results are presented to show the impact of system parameters on the derived performance metrics.

I. INTRODUCTION

Spectrum measurement studies have indicated that large portions of the currently allocated spectrum are highly underutilized [1], [2]. Cognitive radios have been proposed as a promising approach to promote the efficient use of the spectrum by exploiting the existence of spectrum holes (i.e., unused spectrum) [3], [4]. The IEEE 802.22 standard defines a cognitive radio as a radio transmitter/receiver designed to intelligently detect whether a particular segment of the radio spectrum is currently in use and to jump into (or out of) the temporarily-unused spectrum very rapidly without interfering with the transmissions of other users. Cognitive radios offer a great number of benefits in commercial, government, and military applications.

In cognitive radio (CR) networks, users equipped with CRs are referred to as secondary users. Secondary users are unlicensed users of the spectrum who opportunistically share the spectrum resources with the licensed users (i.e., primary users) of the existing system. By allowing secondary users to reclaim idle channels, much higher spectrum efficiency can be achieved [5]. In CR networks, the spectrum availability for the secondary users depends on the spectrum occupancy of the primary users. Prior to initiating a call, each secondary user performs spectrum sensing to determine the channel occupancy status. In addition, an active secondary user occupying a given channel must continue to sense the spectrum occupancy periodically in case a primary user attempts to use the channel. In this case, the active secondary user must leave the current channel and switch to another channel, if one is available.

Recent research on CR networks has used statistical signal processing techniques (cf. [6], [7]), optimization techniques (cf. [8], [9]), and queueing theoretic techniques (cf. [5], [10]), as tools for design and analysis. In [6], a measurement-based model was proposed to characterize the busy and idle periods statistically. Both energy-based detection and feature-based detection strategies were explored to identify spectrum opportunities. In [7], a multiband joint detection approach was proposed for wideband spectrum sensing. The joint detection problem was formulated as a class of optimization problems to improve spectral efficiency and reduce interference.

In [8], Lyapunov optimization was used to design an online flow control and resource allocation algorithm that maximizes throughput subject to constraints on collisions with the primary users. In [9], an approach to the discovery of spectrum opportunities was developed involving a sensing-period optimization mechanism, a channel-sequencing algorithm, and an adaptive channel-usage pattern estimation method. In [10], a queueing-theoretic framework was developed to study important performance measures experienced by secondary users in a CR network employing an opportunistic channel allocation scheme. In [5], an analytic model of opportunistic spectrum sharing with unreliable spectrum sensing was developed using a two-dimensional Markov process.

In this paper, we analyze the performance of a CR network in which service degradation experienced by primary users due to spectrum sensing errors by secondary users can be tolerated. Thus, the reliability of the primary system is dependent on the characteristics of the secondary system. The CR network is modeled using a queueing network, which affords greater flexibility in modeling channel assignment strategies than the single global queue model used in [5]. In [5], a global queue is used only to store secondary calls that have been preempted by primary calls. In the queueing network model developed here, secondary calls are never blocked or dropped; rather, each secondary call enters a queue associated with a particular channel, where it waits (possibly for zero time) until the
channel becomes free. Thus, the queueing network model is appropriate for delay-tolerant data traffic.

The main contribution of this paper is the formulation of a CR data network with tolerable service degradation in terms of a queueing network model and the derivation key performance metrics of interest for network design and analysis. The analytical results derived from the queueing network model can be used to dimension secondary traffic flows subject to constraints imposed by the primary traffic. In particular, the CR network can be dimensioned such that the tolerable service degradation incurred on primary users is kept below a certain threshold.

The remainder of the paper is organized as follows. Section II describes a queueing network model of a CR data network with tolerable service degradation. Section III develops the analysis for a given channel by considering the channel occupancy characteristics of primary and secondary calls. Section IV presents numerical results, illustrating the system performance with respect to various metrics derived in the preceding section. Finally, the paper is concluded in Section V.

II. Model Description

We consider a CR network operating over a given service area. Secondary users equipped with cognitive radios are able to sense channel usage and switch between different channels using appropriate communication mechanisms, while causing negligible interference to the primary users. The licensed spectrum band is divided into \( N \) frequency channels serving primary users. Secondary users detect the presence or absence of primary calls and maintain records of the channel occupancy status. Spectrum detection is performed by a secondary user or by a group of cooperating secondary users. Alternatively, the spectrum detection may involve information exchange with an associated base station in an infrastructured network. Thus, the proposed model is applicable to both infrastructured and infrastructureless network architectures.

Primary users operate as if there are no secondary users in the service area. Primary calls arriving to different channels form independent Poisson processes. An incoming primary call occupies a channel if it is free; otherwise, the call is blocked. We remark that a channel being used by a secondary user is still a free channel as seen by the primary network. A secondary user has to sense the channel availability before accessing a channel. After a secondary user accesses an idle channel, say channel \( i \), the secondary user must continue to sense the channel periodically, in case a primary user attempts to use the channel.

A secondary call occupying a channel may encounter three situations:

1) The call completes without interruption and leaves the system. This occurs with probability \( r_{i0} \).
2) The secondary user senses the arrival of a primary call to the channel and switches the call to another channel \( j \neq i \) with probability \( r_{ij} \).
3) The secondary user fails to sense the arrival of a primary call to the channel (i.e., a sensing error occurs) and continues to the channel. This happens with probability \( r_{ii'} \), \( r_{ii'} = 1 - \sum_{i=0}^{N} r_{ij}, j \neq i \). In this case, both calls receive degraded service.

In the queueing network model, there is no feedback link at a given channel, i.e., a secondary call preempted by a primary call does not join the queue associated with the same channel. By assumption, the radio spectrum is highly underutilized by primary users, such that preempted secondary calls find an alternative idle channel with high probability. If channel \( j \) is being used at that time, the secondary call joins the queue associated with channel \( j \). The queued secondary calls access the channel in FCFS (first-come first-served) order as the channel becomes available.

Fig. 1 illustrates our proposed queueing network model of the CR network, where each channel is effectively a single server queueing system. The external arriving secondary traffic to channel \( i \) is assumed to form a Poisson process with rate \( \gamma_i^S \), \( 1 \leq i \leq N \). The service time of secondary traffic at channel \( i \) is assumed to be exponentially distributed with parameter \( \mu_i^S \), \( 1 \leq i \leq N \). The dynamics of the primary traffic are not shown explicitly in Fig. 1, however, the constraints imposed by the primary traffic on the secondary traffic are embedded in the model. For example, routing probabilities depend on the activity of the primary users. From Fig. 1, the total secondary call arrival rate to channel \( i \), \( \lambda_i^S \), can be obtained from the external arrival rate \( \gamma_i^S \) and the internal arrival rates from other channels.

When the system is in steady-state, the output rate of channel \( i \) is equal to its arrival rate. Hence, we have the
following traffic equation [11]:

$$\lambda^S_i = \gamma^S_i + \sum_{j=1}^{N} \lambda^S_j r_{ji}, \ j \neq i, i = 1, 2, \cdots, N,$$

(1)

where $r_{ji}$ is the routing probability from channel $j$ to channel $i$. The traffic equation (1) can be expressed in vector form:

$$\Lambda_S = \Gamma_S + P' \Lambda_S,$$

(2)

where $\Lambda_S = (\lambda^S_1, \lambda^S_2, \cdots, \lambda^S_N)'$, $\Gamma_S = (\gamma^S_1, \gamma^S_2, \cdots, \gamma^S_N)'$, and the routing matrix $P$ is

$$P = \begin{bmatrix}
0 & r_{12} & \cdots & r_{1N} \\
r_{21} & 0 & \cdots & r_{2N} \\
\vdots & \vdots & \ddots & \vdots \\
r_{N1} & r_{N2} & \cdots & 0
\end{bmatrix}.$$  

(3)

Since the matrix $I - P'$ is invertible, (2) has a unique solution given by

$$\Lambda_S = (I - P')^{-1} \Gamma_S.$$  

(4)

III. Performance Analysis

In this section, we analyze the performance of the queue associated with a target channel $i$ under unreliable spectrum sensing, i.e., an active secondary user on a given channel may fail to detect the presence of an arriving primary call on the channel. Such detection errors are assumed to occur with probability $p_e$. For simplicity, we assume that detection errors are committed only by active secondary users occupying a channel; a secondary user accessing a new channel for a call should have enough time to identify the channel status correctly.

Missed detection events can negatively impact the performance of the primary system, since when such an error occurs, the secondary user transmits on the channel simultaneously with a primary user, causing interference to the primary user. We assume that once a sensing error occurs, both calls will use the same channel until one completes its service and leaves the entire channel for the other. On the other hand, missed detection events can also provide a chance for the active secondary user to continue its call. When this happens, both the primary and secondary call will experience degraded service. For example, each call may occupy part of the channel bandwidth via sub-rating (cf. [12]); hence, each call will have a reduced rate, say, $\alpha P \mu^P_i$ for the primary call and $\alpha_S \mu^S_i$ for the secondary call, where $0 \leq \alpha, \alpha_S \leq 1$, $\alpha + \alpha_S = 1$. We assume that the reduced rate of a primary call satisfies a predefined threshold for the quality of service of primary users. Thus, the system will provide higher supportability to secondary users at the expense of some tolerable service degradation for primary users.

Let the arrival rate and service rate of primary calls to channel $i$ be $\lambda^P_i$ and $\mu^P_i$, respectively. Let $\{X(t), Y(t); t \geq 0\}$ represent the state of the queue associated with channel $i$ at time $t$, where $X(t)$ is set to 1 if the system is in the primary mode serving a primary call, 2 if the system is in the secondary mode serving a secondary call, and 3 if the system is in the degraded mode serving both calls with reduced rates, and $Y(t)$ denotes the number of secondary calls in the system including the one in service (if any). The process $(X(t), Y(t))$ is a two-dimensional Markov process with state space $\{(k, j)|k = 1, 2, 3; j = 0, 1, 2, \cdots\}$ and state transition diagram shown in Fig. 2.

Let $\pi_i(k, j)$ denote the steady-state probability that the channel $i$ system is in state $(k, j)$. The global balance equations of the system are given as follows:

$$\pi_i(1, 0)(\lambda^P_i + \mu^P_i) = \pi_i(2, 0)\lambda^P_i + \pi_i(2, 1)(1 - p_e)\lambda^P_i + \pi_i(3, 1)p_e \alpha_S \mu^S_i,$$

(5)

$$\pi_i(2, 0)(\lambda^P_i + \lambda^S_i) = \pi_i(2, 1)(1 - p_e)\mu^S_i + \pi_i(1, 0)\mu^P_i,$$

(6)

$$\pi_i(1, j)\lambda^P_i = \pi_i(2, j + 1)(1 - p_e)\lambda^P_i + \pi_i(1, j + 1)\lambda^S_i + \pi_i(3, j + 1)p_e \alpha_P \mu^P_i + \pi_i(2, j - 1)\lambda^P_i + \pi_i(2, j + 1)(1 - p_e)\mu^S_i,$$

(7)

$$\pi_i(2, j)(\lambda^P_i + \lambda^S_i + (1 - p_e)\mu^S_i) = \pi_i(3, j)p_e \alpha_P \mu^P_i + \pi_i(2, j - 1)\lambda^P_i + \pi_i(2, j + 1)\lambda^S_i + \pi_i(3, j - 1)\lambda^S_i + \pi_i(2, j + 1)\mu^P_i,$$

(8)

where $j \geq 1$ and $\pi_i(3, 0) \equiv 0$. It is worth noting that state $(2, 0)$ corresponds to the empty system, $(1, 0)$ corresponds to a primary call in service and no secondary call in the system, and $(3, 1)$ corresponds to both a primary call and a secondary call being served on the same channel with no other secondary call in the system.

In states $(2, j), j \geq 1$, an active secondary call contributes to the transition to $(3, j)$ with rate $p_e \lambda^P_i$ when it fails to detect the presence of an arriving primary call, and contributes to the transition to $(1, j - 1)$ with rate $(1 - p_e)\lambda^P_i$ when it detects correctly. From Fig. 2, it is easy to find the routing probabilities as

$$r_{i0} = \frac{(1 - p_e + p_e \alpha_S)\mu^S_i}{(1 - p_e)\lambda^P_i + (1 - p_e + p_e \alpha_S)\mu^S_i},$$  

(10)

$$r_{ij} = \frac{p_e (1 - p_e) \lambda^P_i \beta_{ij}}{(1 - p_e)\lambda^P_i + (1 - p_e + p_e \alpha S)\mu^S_i}, \ j \neq i.$$  

(11)

where $\beta_{ij}$ denotes the conditional probability that the secondary call switches from channel $i$ to channel $j$, given that it has to switch out from channel $i$. The value of $\beta_{ij}$ depends on the channel assignment strategy and real-time measurements. For simplicity, we choose the equal-probability channel assignment strategy, i.e., $\beta_{ij} = 1/(N - 1)$, which can be substituted into (11) to evaluate $\Lambda_S$ in (4). We remark that alternative channel assignment strategies (cf. [13]) can also be represented within the queueing network framework presented here.

We shall use generating function techniques (cf. [14], [15]) to solve the above system equations. Define the partial gener-
Solving (13), (14), and (15), we have
\[ G_{k,i}(z) = \sum_{j=0}^{\infty} \pi_i(k,j) z^j, \quad |z| \leq 1, \quad k = 1, 2, 3. \] (12)
Applying \( G_{k,i}(z) \) to the system equations, we obtain the following equations:
\[
\begin{align*}
\mu_i^P z G_{1,i}(z) &+ p_e \alpha S_i^S G_{3,i}(z) = \pi_i(2,0) \mu_i^P (1 - p_e - z), \\
\mu_i^P z^2 G_{1,i}(z) &+ p_e \alpha S_i^S G_{3,i}(z) + \mu_i^P z G_{2,i}(z) = \pi_i(2,0) \mu_i^P z (1 - z).
\end{align*}
\] (13)
\[
\begin{align*}
\mu_i^P z G_{2,i}(z) &+ (\lambda_i^S z - \lambda_i^S - p_e \alpha P_i^P - p_e \alpha S_i^S) G_{3,i}(z) \\
&= \pi_i(2,0) p_e \lambda_i^P.
\end{align*}
\] (14)
Solving (13), (14), and (15), we have
\[
G_{k,i}(z) = \frac{\pi_i(2,0) g_{k,i}(z)}{g_{0,i}(z)}, \quad k = 1, 2, 3,
\] (16)
where
\[
\begin{align*}
g_{0,i}(z) &= (\lambda_i^S z - \lambda_i^S - \mu_i^P) (\lambda_i^S z - (1 - p_e) \mu_i^P) (\lambda_i^S z - C_0) \\
&\quad - \lambda_i^S z (\lambda_i^S z - C_0 - \mu_i^P) + C_1 - C_0, \\
g_{1,i}(z) &= -\lambda_i^P (\lambda_i^S z)^2 + \lambda_i^P \lambda_i^S (C_0 + C_1) z \\
&\quad - \lambda_i^P (C_0 C_1 - C_2 - C_5), \\
g_{2,i}(z) &= -(1 - p_e) \lambda_i^S \lambda_i^S (\lambda_i^S z - \lambda_i^P - \lambda_i^S - \mu_i^P)^2 \\
&\quad + \lambda_i^S [(1 - p_e) \mu_i^P C_0 + C_3 - C_4] z - C_0 C_3 + \mu_i^P C_2, \\
g_{3,i}(z) &= p_e \lambda_i^P \lambda_i^S (\lambda_i^S z - \lambda_i^P - \lambda_i^S - \mu_i^P) z.
\end{align*}
\] (17)
Using the normalization condition \( \sum_{k=1}^{3} G_{k,i}(1) = 1 \), we obtain
\[
\pi_i(2,0) = \frac{g_{0,i}(1)}{g_{1,i}(1) + g_{2,i}(1) + g_{3,i}(1)}.
\] (21)
The generating functions \( G_{k,i}(z) \), \( k = 1, 2, 3 \), are then determined by substituting (21) into (16).

A. Mean number of secondary calls

The mean number of secondary calls in the channel \( i \) system, denoted by \( L_i \), can be expressed as
\[
L_i = \sum_{k=1}^{3} \sum_{j=1}^{\infty} j \pi_i(k,j) = \sum_{k=1}^{3} G_{k,i}(1). \] (22)
B. System supportability

The supportability of the channel \( i \) system for secondary users, denoted by \( S_i \), is defined as the sum of the steady-state probabilities that the channel supports secondary calls with both normal service (e.g., full-rate) and degraded service (e.g., sub-rate). Thus,
\[
S_i = \sum_{k=2}^{\infty} \pi_i(k,j) = \sum_{k=2}^{3} G_{k,i}(1) - \pi_i(2,0). \] (23)
C. Interference factor

The system supportability of secondary users is achieved at the expense of some service degradation for primary users. We introduce a metric, interference factor of the secondary to the primary system, denoted by \( I_{sp} \), to characterize the impact of service degradation to primary users. The interference factor is defined as the sum of the probabilities of the states in which primary calls receive degraded service:
\[
I_{sp} = \sum_{j=1}^{\infty} \pi_i(3, j) = G_{3,i}(1). \] (24)
By varying the input parameters, e.g., arrival rates, service rates, detection error, etc., one can maximize the carried load of secondary calls subject to limiting the interference factor to a predefined value \( I_0 \).
IV. NUMERICAL RESULTS

We present numerical results in terms of the above performance metrics. We assume a system consisting of \( N = 12 \) channels. The remaining system parameters are set as follows: the primary call arrival rate vector is \( \lambda_P = \lambda_{P0} v_p \), and its associated service rate vector is \( U_P = \mu_{P0} v_p \); the secondary call external arrival rate vector is \( \Gamma_S = \gamma_{S0} v_s \), and the secondary call service rate vector (to the \( N \) channels) is \( U_S = \mu_{S0} v_s \). For the purpose of this numerical study, the constant vectors \( v_p \) and \( v_s \) are assigned as follows:

\[
\begin{align*}
\mathbf{v}_p & \triangleq [1, 1.2, 1.3, 1.5, 1.7, 1.4, 1.3, 1.6, 1.8, 1.9, 1.5, 1.4]; \\
\mathbf{v}_s & \triangleq [1.6, 1.8, 1.9, 1.8, 2.1, 1.9, 1.8, 2.2, 2.3, 2.4, 1.9, 1.8].
\end{align*}
\]

The parameters \( \lambda_{P0}, \gamma_{S0}, \mu_{P0}, \) and \( \mu_{S0} \) are either varied or set to default fixed values as follows: \( \lambda_{P0} = 1 \) or \( 2 \), \( \gamma_{S0} = 1 \); \( \mu_{P0} = 6 \); \( \mu_{S0} = 12 \) or \( 14 \). We set \( \alpha_P = \alpha_S = 0.5 \). The error probability \( p_e \) is indicated in the figure captions. Without loss of generality, we only show the numerical results for channel 1.

Fig. 3 shows the variation in the probability \( r_{10} \) of secondary calls leaving the channel 1 system and the total secondary call arrival rate \( \lambda^S_1 \), when the parameters \( \lambda_{P0} \) and \( \mu_{S0} \) are both varied. As \( \lambda_{P0} \) is increased, the channels of the system will be occupied by primary calls more frequently. Correspondingly, ongoing secondary calls will be interrupted more frequently and have fewer opportunities to complete their service. Consequently, the secondary calls will have to switch to other channels more frequently, resulting in an increase in the internal secondary call arrival rates. As can be seen in Fig. 3, as \( \lambda_{P0} \) increases, \( r_{10} \) decreases and \( \lambda^S_1 \) increases. As \( \mu_{S0} \) increases, \( r_{10} \) increases and \( \lambda^S_1 \) decreases. An increase in \( U_S \) leads to more opportunities for secondary calls to complete their service before the next arrival of a primary call; equivalently, the secondary calls switch to other channels less frequently.

In Fig. 4, we plot the mean number of secondary calls at channel 1, \( \lambda_1 \), vs. the external secondary call arrival rate \( \gamma^S_1 \) for different values of \( \lambda_{P0} \) and \( \mu_{S0} \). We observe that \( \lambda_1 \) increases as \( \gamma^S_1 \) or \( \lambda_{P0} \) increases and decreases as \( \mu_{S0} \) increases. As \( \lambda_{P0} \) is increased, channel 1 is used more frequently by primary calls, leading to fewer opportunities for secondary calls to use the channel. Conversely, an increase in \( \mu_{S0} \) leads to more opportunities for other secondary calls in the queue to use the channel.

In Fig. 5, we show the relationship of channel 1 system supportability \( S_1 \) with respect to \( \gamma^S_1 \) and error probability \( p_e \). We observe that \( S_1 \) increases as \( \gamma^S_1 \) or \( p_e \) increases. As \( p_e \) is increased, more secondary users will remain on channel 1 with degraded service rather than drop out. When \( p_e = 0 \), the secondary call in service must vacate its current channel upon an arrival of primary call, thus leading to the lowest system supportability.

Although an increase of detection error probability can improve the system supportability, it also causes interference to the primary users, especially at relatively heavy primary traffic loads, which is clearly illustrated in Fig. 6. For example, when \( \gamma^S_1 = 11.2 \), the interference factor \( I_{sp} \) is 0.0482 at \( \lambda_{P0} = 1 \), \( p_e = 0.01 \) and 0.0527 at \( \lambda_{P0} = 1 \), \( p_e = 0.01 \); while \( I_{sp} \) is 0.0989 \( \lambda_{P0} = 2 \), \( p_e = 0.01 \). Thus, the improvement of system supportability should be subject to an interference constraint for the primary users. On the other hand, given that the interference constraint is kept below a predefined threshold \( I_p \), an appropriate system supportability gain can be achieved by artificially controlling the detection error \( p_e \).

V. CONCLUSION

We presented an analytic performance model of a cognitive radio (CR) wireless network, where secondary users opportunistically share the radio spectrum with primary users through spectrum sensing. When a secondary user detects the presence of a primary call, it will switch out from its current channel to some other channel to continue its service; however,
when a secondary user fails to detect the primary call, it will remain on the channel with the primary call and both calls will receive degraded service. By modeling the CR network as a queueing network, we derived explicit solutions for the equilibrium system state in terms of generating functions. We derived several performance metrics of interest, including the system supportability to secondary users and the interference factor to primary users. We presented numerical results, which showed the system performance in terms of these performance metrics.

The analytical results derived in this paper can be used to design cognitive radio data networks such that the service degradation caused to primary users is limited below a certain threshold. In this paper, we assumed a simple equal probability channel assignment strategy for secondary users. Other channel assignment strategies for CR networks can also be evaluated within the proposed queueing network framework.