Squeezing the Most Out of Cognitive Radio: A Joint MAC/PHY Perspective

Sai Shankar N
Qualcomm Inc.,
5775 Morehouse Drive, San Diego, 92121, CA, USA
Email: nsai@qualcomm.com

***Abstract*— Cognitive radios provide means to improve the spectrum utilization by allowing opportunistic usage of allocated spectrum today if they are not utilized by the primaries[1]. We first outline that there are two types of cognitive radios [9], namely, static and dynamic cognitive radio.Then we explain the role of physical (PHY) and medium access control (MAC) layer of the cognitive radio and outlines the interaction to perform optimal operation. We model this total interaction by means of a queueing system and analyze the performance. It was found that dynamic cognitive nodes exhibit higher statistical multiplexing gain and lower delay compared to static cognitive radio network.**

Keywords: *Cognitive Radio, Queueing, MAC, PHY, Cross layer*

I. Introduction

Wireless technologies have had an explosive growth with the opening of unlicensed spectrum. The growth of wireless technologies in the unlicensed spectrum has resulted in significant congestion in these bands. These technologies are either power or bandwidth limited and operate sub-optimally. Also congestion in unlicensed bands translates to requiring additional spectrum allocation or better spectrum utilization. Historically, spectrum licensing and access have been static, targeted at a particular application, leading to a low spectral efficiency as shown in a number of studies, e.g., [6]. Opening certain portions of spectrum for unlicensed technologies such as WLAN, Bluetooth and UWB has revolutionized the wireless industry leading to over utilization while the static allocation such as TV bands are under utilized. Since a large amount of (unused or lightly occupied) white space exists in time, space, and frequency, it is often the absence of dynamic channel **access** instead of the true spectral scarcity that limits the growth of wireless communications systems. With this realization, the FCC has taken important initiatives toward flexible and dynamic spectrum policies, including regulation recommendations [4], secondary market spectrum leasing rulings [7], [5], and technical model proposals [3], [2]. At the same time, advanced semiconductor and RF technologies have produced devices that are more intelligent and less expensive with stronger sensing and signal processing capability. Driven by necessity and enabled by new technological advances, now is the perfect time to develop cognitive radio networks. Cognitive radio is a wireless technology that addresses the problem of spectrum utilization by opportunistically accessing portions of spectrum that are not used. Also they have the capability to change their transmission characteristics dynamically based on the policy guidelines specified by regulatory agencies. For details of cognitive radio, the readers are referred to [11].

This paper analyzes the performance of cognitive radio in the presence of cross layer interactions among different layers, namely, PHY, MAC and NET. While the presence of cognitive radio does'nt guarantee that the emerging multimedia applications would meet their required throughput/QoS. We take a high level view of the cognitive radio system wherein the three layers PHY and MAC interact to meet the QoS needs of the application. To model the interaction, we will first analyze the important functionalities that each of the above layers perform in the context of cognitive radio.

- The **PHY layer** has the responsibility to sense the channel for the presence of any primary[2] user and determine if the channel is available or not. Based on the detection scheme, number of sensors employed and diversity techniques, the PHY layer will conclude with sufficient probability whether the sensed channel is occupied by the primary or not. It also has the responsibility to detect the amount of interference present in the current channel if it is occupied by a secondary[3]. Finally it needs to shape the transmitter waveform such that it will cause no or tolerable interference at the primary and meet its bit error rate requirement. The channel occupancy by the primary is represented as T_{on} and non occupancy is represented as T_{off}.
- The **MAC layer** gets the information about the channel occupancy (T_{on} and T_{off}) and will use that information to determine when to switch to a new channel and how to disseminate the information among other secondary nodes in the network. Also the MAC translates the T_{on} and T_{off} to a guaranteed rate that a particular channel can provide and uses that information during admission control and scheduling different applications.

The goal of all the layers is to extract information coming from the PHY layer and exchange it to maximize the QoS expectations of the application. The rest of the paper is organized as follows: Section II outlines the maximum capacity achievable by cognitive radio and the possible architectures. Section III outlines the statistical multiplexing gain achieved by the cognitive radio for different architectures.

[1]Primaries are the owners of the particular spectral band and are allocated by the Federal Communications Commission by auction. Commercial examples include TV bands, cellular bands etc.

[2]These are the owners of the spectrum. Examples of such users are TV transmitters and receivers, cell phones and base stations, police radio etc.

[3]Cognitive radios are also mentioned as secondaries in this paper

II. Capacity of Cognitive Radios

Shannon [1] has shown that the channel capacity C of any medium can be written as:

$$C = B\log_2\Big(1+\frac{S}{N}\Big) \quad (1)$$

Here B is the bandwidth in Hz, and $\frac{S}{N}$ is the signal to noise power ratio.

A. Types of cognitive radio network

There are two types of cognitive radios [?], namely, **static bandwidth cognitive radios** and **dynamic bandwidth cognitive radios**. In static bandwidth cognitive radio, all devices use fixed spectral bandwidth to transmit their data and exploit the available holes (whose sizes fit the operational spectral bandwidth) in the spectrum by opportunistically hopping on those holes in the absence of a primary. Examples of the such systems are current wireless technologies which have cognitive radio functionalities embedded in them. IEEE 802.11 WLANs, IEEE 802.15.3 bluetooth, IEEE 802.16e WiMAX are the best examples of such systems. In the dynamic bandwidth cognitive radio system, the system may be able to expand or contract its spectral bandwidth at any time instant and changes its transmission waveform accordingly. As an example, this can be achieved by switching the carriers "ON" and "OFF" according to spectrum availability in Orthogonal Frequency Division Multiplexing (OFDM) modulation or in Multi-Carrier Code Division Multiple Access (MC-CDMA).

B. Utilization of Cognitive Radio

To motivate the need for such systems, let us consider that the primary occupancy is represented by T_{on} and primary absence is represented by T_{off}. The cognitive radio device would be able to access the system only if the channel state is T_{off}. Assuming that the RF frontend of the cognitive radio is capable of operating in a finite spectral bandwidth W which allows it to operate in any one of $N = \frac{W}{B}$ channels. Here B represents the channel bandwidth of the primary. Also assume that the occupancy in each primary channel is independent of each other and T_{on} and T_{off} are independent identically distributed (i.i.d) with exponentially distribution, one can arrive at the utilization (U) [10] as:

$$U = 1 - \Big(\frac{T_{on}}{T_{on}+T_{off}}\Big)^N \quad (2)$$

For a general case with each channel having its own T_{on}^i and T_{off}^i, we have U as [10]:

$$U = \sum_{k=0}^{N} \frac{\min(M,k) r_k}{M} \quad (3)$$

Here M represents the number of competing cognitive radio networks[4] (CRNs) competing for N channels and $\tau_i = \frac{T_{on}^i}{T_{on}^i + T_{off}^i}$. r_k in the above equation represents the probability of k available channels and is given by:

$$r_k = \sum_{c=1}^{\frac{N!}{k!(N-k)!}} \Bigg[\prod_{i\in S_c^k}(1-\tau_i) \prod_{j\in\{1,2,\cdots,N\}-S_c^k} \tau_j\Bigg] \quad (4)$$

[4]A cognitive radio network consists of more than on cognitive radio device who exchange information among themselves

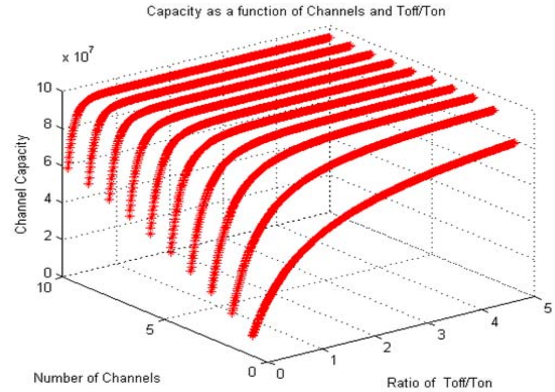

Fig. 1. Capacity as a function of primary occupancy and number of channels

We plot the capacity of CRN as a function of the ratio of $\frac{T_{off}}{T_{on}}$ and number of channels N. From the figure 1 it is clear that the capacity increases as a function of N and T_{off}.

C. Capacity of static and dynamic cognitive radios

Having calculated the utilization, we will now calculate the capacity of static and dynamic cognitive radio. Since the capacity static cognitive radio is just dependent on the number of availability of the channel, it can be expressed as:

$$C_{SCR} = \frac{T - T_{scan}}{T} \times U \times B \times \log_2\Big(1+\frac{P}{N_0 B}\Big) \quad (5)$$

Here C_{SCR} is the capacity of static cognitive radio, T_{scan} represents the scanning time over a interval T and N_0 represents the noise power spectral density. The above equation has a scaling factor of U which represents the utilization of N channels and amount of time spent on scanning for primary channel availability thus reducing the throughput.

For the Dynamic Cognitive Radio, the capacity is expressed as:

$$C_{DCR} = \frac{T - T_{scan}}{T} \times B_{avg} \times \log_2\Big(1+\frac{P}{N_0 B_{avg}}\Big) \quad (6)$$

Here B_{avg} is the average bandwidth and is given by $B_{avg} = r_k B$. The equations 5 and 6 represent the upper bound in the information theoretic sense but we need to exactly compute the channel rate that will be used by the MAC to determine its effective bandwidth and admit streams. In order to do that the MAC obtains the following information from the PHY regarding the following parameters, namely, T_{on}^{min}, T_{on}^{max}, T_{off}^{min}, T_{off}^{max} and average values of T_{on}^{avg} and T_{on}^{avg}. From the above estimated values it is easy to infer that ON and OFF periods can be easily related to leaky bucket characteristics. Here the the time T_{on} represents the time that is not available to the cognitive radio and T_{off} represents the time available for cognitive radio. These times can be related to the peak (P), mean (ρ) and burstiness (σ) of the application as $T_{off} = \frac{\sigma}{P-\rho}$ and $T_{on} = \frac{\sigma}{\rho}$. Additionally T_{on}^{max} represents the maximum scheduling delay and the buffer requirement at the higher layer. The maximum burst size that can be generated by the source is $\sigma_{max} = T_{off}^{min}\sigma$. Using the relationships shown in this paragraph, admission control is done for a arriving

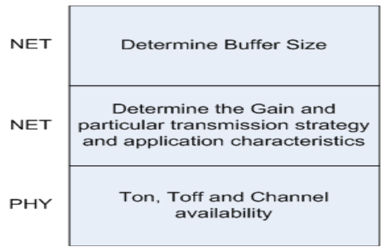

Fig. 2. Functionalities done at different layer in cognitive radio

traffic coming from the higher layer. Also the buffer size at the network layer is determined by the σ_{max}. The following figure 2 outlines the functionalities done at different layers. Having calculated the capacity of the two types of cognitive radios, we will not use that in determining their statistical multiplexing gains.

III. Multiplexing Gain of Using Multiple Channels

A. Delay of dynamic cognitive radio system

Let us consider the dynamic cognitive radios (DCRs) wherein they aggregate the available channel opportunities and use those aggregate channels as if it was one continuous large spectral channel. As an example consider an cognitive radio system that has RF front end capable of operating between 3 and 5 GHz. Also assume that the PHY layer is made up of OFDM with each sub-carrier occupying a bandwidth of 500 KHz. So there are 4000 carriers in this channel starting from 3 GHz. Assume that some of the channels are occupied by primaries. Then the cognitive radio system does not transmit in those carriers that would interfere with the primary transmission and transmits on carriers sensed to be non used by primaries. For the MAC it is one large channel made up of $x(\leq 4000)$ number of carriers. Since the channel availability varies over time, care must be taken to switch ON and OFF the corresponding carriers and hence the effective bit rate would also vary over time as determined by the Eqn.(6).

To analyze the performance of the above system, we construct a single server queue whose server capacity is varying over time. The arrivals to the queue could be modeled as the aggregated arrivals from all cognitive radio networks that are sharing all available channels. Here we have assumed that all cognitive radio devices and networks sense the same channel conditions in their locations. This aggregated arrivals from all cognitive radio networks sees the aggregated "channel" with a varying transmission capacity as shown in Figure 3. The transmission capacity of this queue depends on how many primary networks are using the channels and the distribution of the transmission capacity is determined by Eqn. (4). It is also assumed that the individual arrival rate of the packets[5] from a single cognitive radio network is Poisson and is independent of arrivals from other cognitive radio networks. Hence the aggregated arrivals to the available spectrum from all cognitive radio networks is also Poisson with rate $M\lambda$. The server capacity to this queue is dependent on the occupancy distribution of the primary and is assumed to be generally distributed. Therefore, the system can be modeled as an simple M/G/1 queueing system. However, it is possible that all channels are occupied by primary networks with probability r_0 in Eqn. (4) and for an average duration of $\frac{1}{\sum_i^N \frac{1}{T_{on}^{(i)}}}$, resulting in the transmission capacity of the considered M/G/1 queue reduced to 0. This case is included in our queuing system as an high priority arrival which has preemptive access to the server. The blocking for the spectral agile networks is modeled as another arrival process with rate $\frac{1}{r_0}$, and its "average occupancy" is given by $\frac{1}{\sum_i^N \frac{1}{T_{on}^{(i)}}}$. The resulting M/G/1 queue with preemptive priority is illustrated in Figure 3-(b). The average packet waiting time of a cognitive radio network is then computed by using the results in [12] as

$$W_{DCR} = \frac{\frac{1}{\mu_p}(1-\rho_p-\rho_{CR}) + R_{CR}}{(1-\rho_p)(1-\rho_p-\rho_{CR})}, \qquad (7)$$

where $\mu_p = \sum_i^N \frac{1}{T_{on}^{(i)}}$, $\rho_p = \frac{1}{r_0\mu_p}$, ρ_{CR} represents the server utilization of the cognitive radio network, and R_{CR} represents the average residual service time seen by the packets of cognitive radio networks.

If we assume the average packet size is L and the transmission capacity of a single channel is C, ρ_{CR} is computed as

$$\rho_{CR} = \frac{M\lambda}{\mu_{CR}}, \qquad (8)$$

where $\frac{1}{\mu_{CR}} = \sum_{i=1}^N \frac{L}{i\cdot C} r_i$ is the average service time of a packet from spectral-agile networks. Finally, the residual time R_{CR} is computed as

$$R_{CR} = \frac{1}{2}\left[M\lambda \sum_{i=1}^N (\frac{L}{i\cdot C})^2 r_i + \frac{2}{r_0\mu_p^2}\right], \qquad (9)$$

as derived in [12]. It is clear that when the primary load is very low, then the multiplexing gain of both the systems are close with each other resulting in similar delay performance. When the primary load increases, then the difference becomes significant.

B. Delay of static cognitive radio system

Let us now compare the above queuing system to a static cognitive radio system. Again, we will have a M/G/1 queueing

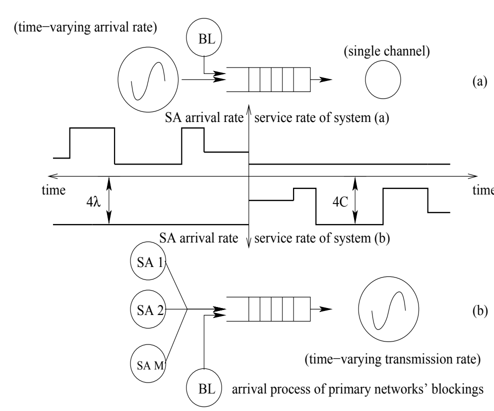

Fig. 3. Queueing models for statistical multiplexing gain: $N = 4$

[5]Packets and frames are used interchangeably in this paper

model with preemptive priority but the parameters of arrival and server capacities are different. Since the cognitive radio network uses utmost one channel, the "service rate" is constant (from the perspective of packets of a secondary network), and is equal to the transmission capacity of a single channel. However, packet arrivals in a channel changes with the number of active primary networks. That is, the packet arrivals in a channel are dependent on the state of the primary network's occupation of the spectrum. If the number of available channels in the spectrum is limited then more and more spectral agile radio networks try to occupy the same channel resulting in more packet arrivals than the available capacity. We can model this arrival process as a Markov-Modulated Poisson Process (MMPP) using Eq. (4), but for the sake of simplicity we approximate the arrival process as a Poisson process, which gives us an M/D/1 queue with preemptive priority. In order to model it as a Poisson process, we need to calculate the average arrival rate.

C. Case I: $M < N$

If there are at least M channels available, then the arrival rate at the $M/G/1$ queue is just λ. If the number of available channels is $M-1$, then one of the spectral-agile networks joins the channel which has already been "occupied" by another spectral-agile network. That is, multiple spectral-agile networks share one channel (in this case 2 spectral agile networks share one channel). The average arrival rate is then $\frac{M}{M-1}\lambda r_k$. Proceeding similarly, we have the average arrival rate computed as

$$\lambda_{new} = \sum_{i=M}^{N} r_i \lambda + \sum_{i=1}^{M-1} r_{M-i}\Big(\frac{M}{M-i}\Big)\lambda. \qquad (10)$$

D. Case II: $M > N$

In this case, we have more cognitive radio networks than the total number of available channels. If none of the channels are occupied by the primary network, then the best-case arrangement occurs when each channel has $\lceil \frac{M}{N} \rceil$ spectral-agile networks. Proceeding similarly to the previous subsection, we have the arrival rate λ computed as

$$\lambda_{new} = \sum_{i=0}^{N-1} r_{N-i}\Big(\frac{M}{N-i}\Big). \qquad (11)$$

Finally, we can use Eqs. (7) and (9) with the new average arrival rate and constant packet service time $\frac{L}{C}$.

Figure 7 plots the average packet waiting time of a spectral-agile network when it uses a single channel and multiple channels. We fix the value of T_{off} at 1 second while varying the value of T_{on}, so as to vary each channel's average load imposed by the primary network. Obviously, the average packet waiting time in the case of $M < N$ is less than that in the case of $M > N$ as there are less spectral-agile networks seeking spectral opportunities in the case of $M < N$. However, the packet waiting time of using multiple channels is always less than that of using a single channel in both cases. The improvement is even more significant in the case of $M > N$ as expected. These numerical results demonstrate the potential advantages of using multiple channels in a spectral-agile network, especially when $M > N$.

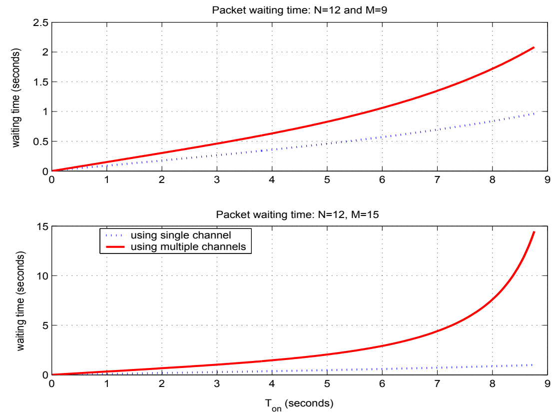

Fig. 4. Average waiting time of packets from a spectral-agile network: $T_{off} = 1$ and $\frac{L}{C} = 0.1$

IV. Conclusions

This paper explains the functionalities of different layers and how to treat them in a unified way so that performance of a cognitive radio network can be characterized by a simple queueing system. We obtain the capacities of two types of cognitive radios and determine the statistical multiplexing gain for those networks. Further we model the available time of the channel (T_{off}) and the non available times (T_{on}) to the traffic characteristics that can be sustained in the cognitive radio. We model this total interaction by means of a queueing system and analyze the performance. It was found that dynamic cognitive nodes exhibit higher statistical multiplexing gain and lower delay compared to static cognitive radio network.

References

[1] C.E. Shannon: Communication in the Presence of Noise Proc. IRE, Vol. 37, pps 10 - 21, January 1949.

[2] FCC. Establishment of an interference temperature metric to quantify and manage interference and to expand available unlicensed operation in certain fixed, mobile and satellite frequency bands, ET Docket No. 03-237, FCC 03-289.

[3] FCC. Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies, notice of proposed rule making and order, FCC 03-322.

[4] FCC. FCC spectrum policy task force report, ET Docket No. 02-135.

[5] FCC. In the matter of promoting efficient use of spectrum through elimination of barriers to the development of secondary markets, WT Docket No. 00-230, FCC Rcd 20604.

[6] Mark McHenry. Spectrum white space measurements. presented to New America Foundation BroadBand Forum, June 2003. http://www.newamerica.net/Download_Docs/pdfs/DocFile_185_1.pdf.

[7] FCC. Promoting efficient use of spectrum through elimination of barriers to the development of secondary markets, WT Docket No. 00-230, report and order and further notice of proposed rulemaking, FCC 03-113 (secondary markets report and order and FNPRM).

[8] FCC. Unlicensed operation in the TV broadcast bands, ET Docket No. 04-186; additional spectrum for unlicensed devices below 900 mhz and in the 3 ghz band, ET Docket No. 02-380, FCC 04-113.

[9] Sai Shankar N, Chun-Ting Chou, Kiran Challapali, Stefan Mangold, "Spectrum Agile Radio: Capacity and QoS Implications of Dynamic Spectrum Assignment", *IEEE GLOBECOM 2005*, St. Louis, USA.

[10] Chun-Ting Chou, Sai Shankar N, Hyoil Kim and Kang G Shin, "What and How Much to Gain From Spectrum Agility", Accepted in *IEEE JSAC, Special Issue on Cognitive Radios*, August 2006.

[11] http://www.darpa.mil/ato/programs/XG/rfc_vision.pdf — *Vision RFC*.

[12] D. Bertsekas and R. Gallager, "Data Networks", *Prentice-Hall*, New York, 1992.