Chapter Number

Multi-User Multimedia Transmission over Cognitive Radio Networks Using Priority Queuing

Hsien-Po Shiang, Mihaela van der Schaar
University of California Los Angeles
USA

1. Introduction

The emergence of cognitive radio networks have spurred both innovative research and ongoing standards (Mitola et al., 1999; Haykin, 2005; Cordeiro et al., 2006). Cognitive radio networks have the capability of achieving large spectrum efficiencies by enabling interactive wireless users to sense and learn the surrounding environment and correspondingly adapt their transmission strategies. In this context, there exist three main challenges for multimedia users to efficiently transmit their delay-sensitive traffic over the cognitive radio networks. The first problem is how these multimedia users should sense the spectrum and timely model the behavior of the primary licensees. The second problem is how these users should manage the available spectrum resources and share the resource to the license-exempt users to satisfy their multimedia traffic requirements while not interfering with the primary licensees. The third problem is how to maintain seamless communication during the transition (hand-off) of selected frequency channels. In this chapter, we focus on the second challenge regarding the resource management problem. For the remaining two challenges, one can find relevant discussions in other existing literatures as in Akyildiz et al., 2006 and Brown, 2005, etc.

Due to the informationally-decentralized nature of cognitive radio networks (Shiang & van der Schaar, 2007b), the complexity of the optimal centralized solutions (Zekavat & Li, 2006; Fu & van der Schaar, 2007; van der Schaar & Fu, 2009) for spectrum allocation is prohibitive for large systems. In addition, the centralized solution might require a large amount of time to process and to collect the required information, which induces delay that can be unacceptable for the delay-sensitive applications, e.g. multimedia streaming. Hence, it is important to implement distributed solutions as in (Shiang & van der Schaar, 2009) for dynamic resource management by relying on the wireless users’ capabilities to sense, adapt, and coordinate themselves. Importantly, for the distributed solutions, the coordinated interactions (information exchanges) across the autonomous wireless users are essential, since the decisions of an autonomous wireless user will impact and be impacted by the other users. Without explicit coordinated interactions, the heterogeneous users will consume additional resources and respond slower to the time-varying environment. Such information
exchange can rely on a dedicated control channel for all users (Brik et al., 2005), or using a

group-based coordination scheme without a common control channel (Zhao et al., 2005).

In recent years, the research focus regarding multimedia transmission in wireless

networks has been to adapt existing multimedia compression (Stockhammer et al., 2003),

error protection algorithms (Mohr et al., 2000), and rate-distortion optimized transmis-

sion (Chou & Miao 2006) to the rapidly varying resources of wireless networks (see van der

Schaar & Chou, 2007 for more references). Significant contributions have been made to

enhance the separate performance of the various OSI layers, or jointly for the MAC, PHY,

and application layers (van der Schaar et al., 2003; Setton et al. 2005). However, these

solutions cannot provide an integrated and realistic cross-layer optimization framework in

the cognitive radio networks to support delay-sensitive multimedia streaming applications.

Importantly, the cross-layer optimization has been performed in an autonomous, selfish and

isolated manner, at each multimedia source/user, and does not consider its impact on the

overall wireless infrastructure and the interactions with other information streams. As such,

existing solutions do not provide adequate support for multi-user multimedia streaming

ever spectrum agile network.

In this chapter, we introduce an integrated cross-layer optimization framework for multi-

user multimedia transmission over cognitive radio networks. The traffic of the users

(including the licensed users and the license-exempt users) and the channel conditions (e.g.

Signal-to-Noise Ratio, Bit-Error-Rate) are modeled using stationary stochastic models

(Shanker et al., 2005). Based on these models, a novel priority virtual queue analysis

(Shiang & van der Schaar, 2008) for cognitive radio networks is introduced. This analysis enables a

coordination interface to the license-exempt wireless users without requiring to change

existing communication protocols, e.g. IEEE 802.22 (Cordeiro et al., 2006). The virtual

queues are priority queues for each of the frequency channels. They are emulated in a

distributed manner by each autonomous wireless user in order to estimate the delay of

selecting a specific frequency channel for transmission. Unlike the majority of prior works

assuming the available frequency channels as spectrum holes (Haykin, 2005; Akyildiz et al.,

2006) that can be accessed using a 2-state on-off channel model (Shanker et al., 2005), we

adopt priority queuing to model the users’ interactions by assigning the highest priority to

the licensed users and adapt the channel access of the license-exempt users (with lower

priorities). Importantly, instead of making the primary licensees passively exclude the

license-exempt users from using the occupied frequency channels, the introduced approach

allows the primary licensees to share the frequency channels and also endows these primary

licensees with the preemptive priority to delay the transmissions of the license-exempt users

in the same frequency channel. Hence, the introduced approach can further improve the

spectrum utilization and reduce the impact of the license-exempt users. The proposed

concept can also be applied to the leased network as in Akyildiz et al., 2006 and Stine, 2005.

Based on the priority queuing analysis, each wireless user builds an abstraction of the
dynamic wireless environment (e.g. wireless condition) and the competing users’ behaviors

using the same frequency channel (including the primary licensees, to which the highest

priority is assigned). Note that the abstraction is important in order to enable intelligent

wireless users to learn and adapt their cross-layer transmission strategies (Haykin, 2005;

Mitola et al., 1999). Additionally, the necessary multi-agent interactions (information

exchanges) are also determined for the priority queuing analysis. This chapter focuses on

the delay-sensitive applications such as surveillance, multimedia conferencing, and media
streaming etc., since these applications are most impacted by inefficient spectrum usage. Moreover, this chapter only focuses on the multimedia transmission over a single-hop network infrastructure. Discussion regarding to multimedia transmission over a multi-hop cognitive radio network can be found in Shiang & van der Schaar, 2009.

The organization of the chapter is as follows. In Section 2, we provide the network settings of the cognitive radio network. Section 3 presents the cross-layer problem formulation for multi-user multimedia streaming over such network through a multi-agent interaction. In Section 4, we show that the multi-user multimedia streaming problem over such network can be analyzed using priority queue modeling and hence, facilitate the optimal cross-layer transmission strategies of the multimedia streaming problem through appropriate information exchange. Finally, Section 5 concludes the chapter.

2. Network Settings for Multi-User Multimedia Transmission over Cognitive Radio Networks

2.1 Multimedia traffic characteristics

Assume that there are \( N \) multimedia applications \( V_1, \ldots, V_N \) with distinct sources destinations. In this chapter, we define the users as the transmission pairs between predetermined source wireless stations and destination wireless stations. Each encoded multimedia stream is separated into a certain number of classes (quality layers) as in van der Schaar et al., 2006. Assume that the packets within each multimedia class have the same delay deadline. The number of priority classes for a multimedia sequence \( V_i \) equals \( K_i \). Assume that the total number of priority classes across all users in the network is \( K \). The priority classes in the network are denoted as \( C_1, \ldots, C_K \). For the purpose of analysis, we reserve the highest priority class \( C_1 \) for the primary users in each frequency channel, i.e. \( \lambda_1 \gg \lambda_k, 2 \leq k \leq K \). The secondary users can be categorized into the rest of \( K - 1 \) priority classes \( (C_2, \ldots, C_K) \) to access the frequency channels. Hence, the total number of classes across all users in the network equals \( K = \sum_{i=1}^{N} K_i + 1 \). The users in higher priority classes can preempt the transmission of the lower priority classes to ensure an interference-free environment for the primary users (Kleinrock, 1975). The priority of a user affects its ability of accessing the channel. Each multimedia class \( C_k \) is characterized by:

- \( \lambda_k \), the expected quality impact of receiving the packets in the class \( C_k \). We prioritize the multimedia classes based on this parameter. In the subsequent part of the chapter, we label the \( K \) classes (across all users) in descending order of their priorities, i.e. \( \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_K \).
- \( L_k \), the average packet lengths of the class \( C_k \). The expected quality improvement for receiving a multimedia packet in the class \( C_k \) is defined as \( \lambda_k \cdot L_k \) (see e.g. Wang & van der Schaar, 2006 for more details).
- \( N_k \), the number of packets in the class \( C_k \) in one GOP duration of the corresponding multimedia sequence.

\(^1\) The prioritization of the secondary users can be determined based on their applications, prices paid for spectrum access, or other mechanism design based rules. In this chapter, we will assume that the prioritization was already performed.
\( P_k^{\text{succ}} \), the probabilities of successfully receiving the packets in the class \( C_k \) at the destination. Thus, the expected number of the successfully received packets of the class \( C_k \) is \( N_k \cdot P_k^{\text{succ}} \).

- \( d_k \), the delay deadlines of the packets in the class \( C_k \). Due to the hierarchical temporal structure deployed in 3D wavelet multimedia coders (see Wang & van der Schaar, 2006; van der Schaar & Turaga, 2007), for a multimedia sequence, the lower priority packets also have a less stringent delay requirement. This is the reason why we prioritize the multimedia bitstream in terms of the quality impact. However, if the used multimedia coder did not exhibit this property, we need to deploy alternative prioritization techniques \( \lambda_k^{\text{video}}(\lambda_k, d_k) \) that jointly consider the quality impact and delay constraints (see more sophisticated methods in e.g. Jurca & Frossard, 2007; Chou & Miao, 2006).

At the client side, the expected quality improvement for multimedia \( V_i \) in one GOP can be expressed as:

\[
\begin{align*}
\hat{u}_i = \sum_{C_k \in V_i} \lambda_k L_k N_k P_k^{\text{succ}}.
\end{align*}
\]  

(1)

Here, we assume that the client implements a simple error concealment scheme, where the lower priority packets are discarded whenever the higher priority packets are lost (van der Schaar & Turaga, 2007). This is because the quality improvement (gain) obtained from decoding the lower priority packets is very limited (in such embedded scalable multimedia coders) whenever the higher priority packets are not received. For example, drift errors can be observed when decoding the lower priority packets without the higher priority packets (Wang & van der Schaar, 2006). Hence, we can write:

\[
\begin{align*}
P_k^{\text{succ}} &= \begin{cases} 
0, & \text{if } P_k^{\text{succ}} = 1 \text{ and } C_{k'} \prec C_k \\
(1 - P_k) E[I(D_k \leq d_k)], & \text{otherwise,}
\end{cases}
\end{align*}
\]  

(2)

where we use the notation in (Chou & Miao, 2006) - \( C_{k'} \prec C_k \) to indicate that the class \( C_k \) depends on \( C_{k'} \). Specifically, if \( C_k \) and \( C_{k'} \) are classes of the same multimedia stream, \( C_{k'} \prec C_k \) means \( k' < k \) due to the descending priority \( (\lambda_{k'} > \lambda_k) \). \( P_k \) represents the end-to-end packet loss probability for the packets of class \( C_k \). \( D_k \) represents the experienced end-to-end delay for the packets of class \( C_k \). \( I(\cdot) \) is an indicator function. Note that the end-to-end probability \( P_k^{\text{succ}} \) depends on the network resource, competing users’ priorities as well as the deployed cross-layer transmission strategies. In addition, for the multimedia \( V_i \), let us assume that the multimedia packets are scheduled in a specific order \( \pi_i \) according to the prioritization associated with the multimedia content characteristics.

### 2.3 Cognitive radio network settings

Assume that there are a total of \( M \) frequency channels \( F = \{F_1, \ldots, F_M\} \) in the cognitive radio network and there are \( M \) primary users \( \text{PU} = \{PU_1, \ldots, PU_M\} \), each in a separate frequency channel. These primary users can only occupy their assigned frequency channels. Since the primary users are licensed users, they are provided with an interference-free environment (Haykin, 2005; Akyildiz et al., 2006). Assume that there are \( N \) secondary users...
\( \text{SU} = \{SU_1, \ldots, SU_N\} \) in the system. These secondary users are able to operate their applications across various frequency channels for transmission. Hence, they need to time share the chosen frequency channel. Moreover, these secondary users are usually the license-exempt users, and hence, the resulting impact on the primary users must be eliminated.

Let us further assume that there is a Network Resource Manager (NRM) that coordinates multiple access control scheme for sharing the spectrum resource (by assigning transmission opportunities), while ensuring that the corresponding interference on the primary users is eliminated. The role of the NRM is similar to the coordinator in the current IEEE 802.11e Hybrid Coordination Function (HCF) solutions for multimedia applications (van der Schaar et al., 2006). Note that the NRM will not make decisions for the secondary users, but it will only manage the transmission opportunities of the frequency channels based on the priority classes to avoid interference. In this chapter, we investigate the dynamic resource management problem for the multimedia streaming of the secondary users that are associated with this specific NRM. Primary users in the highest priority class \( C_1 \) can always access their corresponding channels at any time. Secondary users, on the other hand, require transmission opportunities from the NRM for transmission based on their priorities.

Multiple users can time share the same frequency channel. Note that even if the same time sharing fractions are assigned to the users choosing the same frequency channel, the experienced channel conditions can be different for the users. A wireless user needs to stream its multimedia over an appropriate frequency channel to minimize the transmission delay \( D_k \) and thereby, increase the multimedia quality in equation (1). For a certain frequency channel \( F_j \), the secondary users can experience various channel conditions for the same frequency channel. Besides, the frequency selections of the secondary users also mutually affect each other. Let us denote \( T_{ij} \) and \( p_{ij} \) as the resulting physical transmission rate and packet error rate for the secondary user \( SU_i \) transmitting through a certain frequency channel \( F_j \). Let \( R_{ij} = [T_{ij}, p_{ij}] \in \mathcal{R} \) be the channel conditions of the channel \( F_j \) for the secondary user \( SU_i \) and denote the channel condition matrix as \( R = [R_{ij}] \in \mathcal{R}^{M \times N} \).

The effective transmission rate \( T_{ij}^e = T_{ij}(1 - p_{ij}) \) depends on 1) the channel conditions, i.e. Signal-to-Noise-Ratio (SNR) \( x_{ij} \), 2) the modulation and coding scheme \( \theta \) that is adopted by the user, and 3) the multiple access scheme of sharing the channel. For simplicity, in this chapter, let us assume that the NRM adopts the simplest polling mechanism (Bertsekas & Gallager, 1987) similar to IEEE 802.11e that assigns transmission opportunity to the secondary users from the users in higher priority class to the lower priority class. The users in the same priority class will be polled in a round-robin fashion and have the same chance to transmit their packets. More sophisticated MAC protocols are proposed to deal with the spectrum heterogeneity (such as HD-MAC in Zhao et al., 2005). However, the contention-based MAC protocol are not preferable for delay-sensitive applications. Hence, a simple polling-based MAC similar to that used in IEEE 802.11e is considered in this chapter. The NRM only ensures the priority order among the users, but will not decide the secondary users’ actions. The expected physical transmission rate \( T_{ij}^e \)
and packet error rate $p_{ij}$ can be approximated as sigmoid functions of measured and the adopted modulation and coding scheme as in (Krishnaswamy, 2002):

$$p_{ij}(\theta_i, x_{ij}) = \frac{1}{1 + \exp(\zeta(\theta_i)(x_{ij} - \delta(\theta_i)))},$$  (3)

$$T_{ij}^e(\theta_i, x_{ij}) = \frac{T_{ij}(\theta_i)}{1 + \exp(-\zeta(\theta_i)(x_{ij} - \delta(\theta_i)))},$$  (4)

where $\zeta(\theta_i)$ and $\delta(\theta_i)$ are empirical constants for multimedia user $V_i$ using frequency channel $F_j$ corresponding to the modulation and coding schemes $\theta_i$ for a given packet length $L_k$ of class $C_k$. Thus, even though coordinated by the same NRM, the expected $T_{ij}$ and $p_{ij}$ of the same frequency channel can be different for various secondary users.


3.1 Cross-layer transmission problem formulation

The considered actions of a secondary user $V_i$ as the following cross-layer transmission strategies $a_i = [\pi_i, \gamma_i, \alpha_i, \theta_i] \in A_{tot} = A_{\pi} \times A_{\gamma} \times A_{\alpha} \times A_{\theta}$ including:

- The physical layer modulation and coding scheme $\theta_i \in A_{\theta}$.
- The MAC layer retransmission limit $\gamma_i \in A_{\gamma}$.
- The application layer multimedia packet scheduling $\pi_i \in A_{\pi}$.
- The selection of the frequency channel for multimedia transmissions $\alpha_i \in A_{\alpha}$.

Denote the frequency selection of a secondary user $SU_i$ using $\alpha_i = [a_{i1}, a_{i2}, \ldots, a_{iM}] \in A_{\alpha} = \{0,1\}^M$, where $a_{ij} = 1$ represents the fact that the secondary user $SU_i$ chooses the frequency channel $F_j$. Otherwise, $a_{ij} = 0$. Let $\alpha_{-i}$ denote the actions of the other secondary users except $SU_i$. Let $A = [a_1, \ldots, a_N]$ denote the total action set across all secondary users.

As stated in equation (1), each secondary user has its own multimedia quality function as the utility function to maximize. Conventionally, the utility function of a specific user is often modeled solely based on its own action, i.e. $u_i(a_i)$ without modeling the other secondary users (Wang & Pottie, 2002; van der Schaar & Shanker, 2005). However, in fact, the utility function in cognitive radio networks depends on the effective delay and throughput that the secondary user can get from the frequency channel it selects, and this is related to the channel sharing, which is coupled with other secondary users. Hence, the utility function $u_i$ is also influenced by the action of other secondary users that select this frequency channel. In other words, the utility function can be regarded as $u_i(a_i, \alpha_{-i})$.

Specifically, each autonomous multimedia user selects its optimal action $a_i^{opt}$ by solving the optimization problem with the knowledge of $\alpha_{-i}$:

$$a_i^{opt} = \arg \max_{a_i \in A_{\alpha}} u_i(a_i, \alpha_{-i})$$  (5)
Importantly, in an informationally-decentralized cognitive wireless network that consists of decentralized secondary users, the secondary user $SU_i$ may not know the exact actions of other secondary users $\alpha_{-i}$. Moreover, even if all the actions are known, it is unrealistic to assume that the exact action information can be collected timely to compute and maximize the actual utility function $u_i(a_i, \alpha_{-i})$. Hence, a more practical solution is to dynamically model the other secondary users’ behavior by updating their probabilistic frequency selection strategy profile based on some available information exchange, and then maximizes the expected utility function of $SU_i$.

Hence, we define a frequency selection strategy profile of a secondary user $SU_i$ as a vector of probabilities $s_i = [s_{i1}, s_{i2}, \ldots, s_{iM}] \in S^i = S^M$, where $s_{ij} \in S$ ($S \in [0,1]$) represents the probability of the secondary user $SU_i$ choosing the frequency channel $F_j$ ($a_{ij} = 1$). Hence, the summation over all the frequency channels, $\sum_{j=1}^{M} s_{ij} = 1$. Note that $s_{ij}$ can also be viewed as the fraction of data from $SU_i$ transmitted on frequency channel $F_j$, and hence, multiple frequency channels are selected for a secondary users with $s_{ij} > 0$.

Let $S = [s^T_1, \ldots, s^T_N] \in S^{M \times N}$ denote the total strategy profile across all the secondary users. The expected utility function, given a fixed strategy profile $(s_i, s_{-i})$ is

$$U_i(a_i(s_i), s_{-i}) = E_{(s_i, s_{-i})}[u_i(a_i, \alpha_{-i})],$$

where $s_{-i}$ denotes the collected frequency selection profile. $a_i(s_i) = [\pi_i, \gamma_i, s_i, \theta_i] \in \mathcal{A}_{od} = \mathcal{A}_i \times \mathcal{A}_i \times S^i \times \mathcal{A}_i$ represents the transmission strategy using $s_i$ instead of $\alpha_i$. Then, the optimization problem in equation (5) becomes:

$$\tilde{a}_i^{opt} = \arg \max_{\tilde{a}_i \in \mathcal{A}_{od}} U_i(\tilde{a}_i(s_i), s_{-i}) = \arg \max_{\tilde{a}_i \in \mathcal{A}_{od}} \sum_{C_i \in V_i} \lambda_i L_i N_i P^{succ}(\tilde{a}_i(s_i), s_{-i})$$

Figure 1 provides the system architecture of the secondary users. In Section 4, we will discuss how to model the strategy (behavior) $s_{-i}$ and the impact of the other secondary users using priority queuing analysis.
### 3.2 Cross-layer optimization examples

We first look at the case with 6 secondary users with multimedia streaming applications ("Coastguard", frame rate of 30Hz, CIF format, delay deadline 500ms) sharing 10 frequency channels \( N = 6, M = 10 \). We compare the discussed cross-layer optimization using priority queuing analysis with other two resource allocation algorithms – the “Static Assignment” (Tekinay & Jabbari, 1991) and the “Dynamic Least Interference” (Wang & Pottie, 2002). In the “Static Assignment” algorithm, a secondary user will statically select a frequency channel with the best effective transmission rate without interacting with other secondary users. In the “Dynamic Least Interference” algorithm, a secondary user will dynamically select a single frequency channel that has the least interference from the other users (both secondary users and primary users). We consider 100 random frequency channel conditions as well as the traffic loadings and then compute the average the video PSNR and the standard deviation of the PSNR over the one hundred cases in Table 1 for the 6 video applications. There are primary users randomly appears in each frequency channel (occupying different frequency channels with a fixed loading). The results show that the cross-layer optimization outperforms the other two algorithms for delay-sensitive multimedia applications in terms of video quality (PSNR). The standard deviations of the cross-layer optimization are also smaller than the other two solutions. Unlike the “Dynamic Least Interference” that only considers how a single user adapts to the experienced environment, the multi-agent cross-layer optimization allows the secondary users to track the frequency selection strategies of the other users and adequately optimize the cross-layer transmission strategies. Hence, the users will be able to self-organize into various cognitive radio channels while adapting to the new environmental conditions. In Table 2, we see that the multi-agent cross-layer optimization approach still outperforms the other two approaches with different loadings of the primary users in each frequency channel.

<table>
<thead>
<tr>
<th>SU ( i )</th>
<th>Average Y-PSNR ( \theta_{ij}^{opt} ) (dB)</th>
<th>Average Y-PSNR Standard Deviation (dB)</th>
<th>Average Y-PSNR Standard Deviation (dB)</th>
<th>Average Y-PSNR Standard Deviation (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SU ( 1 )</td>
<td>29.48</td>
<td>4.94</td>
<td>29.89</td>
<td>4.32</td>
</tr>
<tr>
<td>SU ( 2 )</td>
<td>29.90</td>
<td>4.89</td>
<td>30.35</td>
<td>4.29</td>
</tr>
<tr>
<td>SU ( 3 )</td>
<td>29.69</td>
<td>5.02</td>
<td>30.37</td>
<td>4.41</td>
</tr>
<tr>
<td>SU ( 4 )</td>
<td>30.59</td>
<td>4.98</td>
<td>30.87</td>
<td>4.37</td>
</tr>
<tr>
<td>SU ( 5 )</td>
<td>29.48</td>
<td>4.98</td>
<td>29.87</td>
<td>4.41</td>
</tr>
<tr>
<td>SU ( 6 )</td>
<td>30.01</td>
<td>5.04</td>
<td>30.65</td>
<td>4.46</td>
</tr>
</tbody>
</table>

Table 1. Comparisons of the channel selection algorithms for video streaming with \( N = 6, M = 10 \).
Next, let us take a look at the impact of different numbers of secondary users with video streaming applications. Figure 2 shows the average packet loss rate and the average PSNR over the $N$ video streams (instead of over 100 different channel conditions in the previous simulation). The empirical average $T_i^j(\theta_i^{opt})$ of the frequency channels is shown to be 3 Mbps, instead of 1 Mbps in the previous example. Larger $N$ reduces the available resources that can be shared by the video streams, and hence, increases the application layer packet loss rate (due to the expiration of the delay deadline) and hence, decreases the received video quality. The result shows that the cross-layer optimization using the priority queuing analysis outperforms the other two algorithms for multi-user video streaming applications.

![Fig. 2 Average Y-PSNR versus number of secondary users using different resource management schemes](image-url)
4. Priority Queuing Analysis for Multimedia Transmission in Cognitive Radio Networks

In this section, we discuss the priority queuing analysis for the multimedia streaming problem over cognitive radio networks. The goal is to provide an abstraction of the dynamic wireless environment and the competing wireless users’ behaviors that impact the secondary user’s utility. It is important to note that the packets of the competing wireless users are physically waiting at different locations. Figure 3 gives an example of the physical queues for the case of $M$ frequency channels and $N$ secondary users. Each secondary user maintains $M$ physical queues for the various frequency channels, which allows users to avoid the well-known head-of-line blocking effect (Wang et al., 2004).

![Diagram of physical queues at secondary users](image)

Fig. 3 Frequency selection of the secondary users $a_{ij}$ and physical queues and virtual queues for each frequency channel.

4.1 Traffic models of the users

- **Traffic model for primary users** – Assume that the stationary statistics of the traffic patterns of primary users can be modeled by all secondary users. The packet arrival process of a primary user is modeled as a Poisson process with average packet arrival rate $r_{j}^{PU}$ for the primary user $PU_j$ using the frequency channel $F_j$. We denote the $m_{th}$ moments of the service time distribution of the primary user $PU_j$ in frequency
channel $F_j$ as $E[(X_j^{SU})^m]$. Note that this traffic model description is more general than a Markov on-off model (Shanker et al., 2005), which is a sub-set of the introduced queuing model with an exponential idle period and an exponential busy period. As in Shanker et al., 2005, an M/G/1 model is adopted in this chapter for the traffic descriptions.

- **Traffic model for secondary users** – Assume that the average rate requirement for the secondary user $SU_i$ is $B_i$ (bit/s). Let $r_{ij}$ denote the average packet arrival rate of the secondary user $SU_i$ using the frequency channel $F_j$. Since the strategy $s_{ij}$ represents the probability of the secondary user $SU_i$ taking action $a_{ij}$ (transmitting using the frequency channel $F_j$), we have $r_{ij} = s_{ij}B_i / L_i(\pi_i)$, where $L_i(\pi_i)$ denotes the average packet length of the secondary user $SU_i$ depending on which priority class of multimedia packets is chosen in $\pi_i$. If a certain secondary user $SU_i$ can never use the frequency channel $F_j$, we fix its strategy to $s_{ij} = 0$, and hence, $r_{ij} = 0$. For simplicity, we also model the packet arrival process of the secondary users using a Poisson process.

Since packet errors are unavoidable in a wireless channel, let us assume that packets will be retransmitted, if they are not correctly received. This can be regarded as a protection scheme similar to the Automatic Repeat Request protocol in IEEE 802.11 networks. Hence, the service time of the users can be modeled as a geometric distribution. Let $E[X_{ij}(\pi_i, \gamma_i, \theta_i)]$ and $E[X_{ij}^2(\pi_i, \gamma_i, \theta_i)]$ denote the first two moments of the service time of the secondary user $SU_i$ using the frequency channel $F_j$. For simplicity, we denote $E[X_{ij}(\pi_i, \gamma_i, \theta_i)]$ and $E[X_{ij}^2(\pi_i, \gamma_i, \theta_i)]$ using $E[X_{ij}]$ and $E[X_{ij}^2]$ hereafter in this chapter. Based on the physical layer link adaptation (Krishnaswamy, 2002), $R_{ij} = [T_{ij}, p_{ij}]$ in equation (3) and (4), we have:

$$ E[X_{ij}] = \frac{(L_i(\pi_i) + L_o)(1 - p_{ij}(\theta_i))^{\gamma_i+1}} {T_{ij}(\theta_i)(1 - p_{ij}(\theta_i))} \approx \frac{(L_o(\pi_i) + L_o)} {T_{ij}(\theta_i)(1 - p_{ij}(\theta_i))}, $$

(8)

$$ E[X_{ij}^2] \approx \frac{(L_o(\pi_i) + L_o)^2(1 + p_{ij}(\theta_i))} {T_{ij}(\theta_i)^2(1 - p_{ij}(\theta_i))^2}, $$

(9)

where $L_i$ is the average packet length of the secondary user $SU_i$ and $L_o$ represents the effective control overhead including the time for protocol acknowledgement, information exchange, and NRM polling delay, etc.

To describe the traffic model, we define the traffic specification$^2$ for the secondary user $SU_i$ as $\mathbf{T}_i = [C_k, B_i, L_i, X_i, X_i^2]$, if $SU_i \in C_k$, where $X_i = [E[X_{ij}] \mid j = 1, \ldots, M]$ and $X_i^2 = [E[X_{ij}^2] \mid j = 1, \ldots, M]$ are the vectors of the first two moments of service time when user $SU_i$ transmitting in each frequency channel. This information needs to be exchanged among the secondary users, which will be discussed in more details in Section 4.4.

$^2$The traffic specification is similar to the TSPEC in current IEEE 802.11e for multimedia transmission.
4.2 Queuing analysis for the secondary users with the same priority

In this subsection, we first consider the case that all packets have the same priority by ignoring the impact of the primary users. In the next subsection, we will generalize these results by considering the impact of the primary users using priority queuing.

In order to solve the dynamic resource management problem, we need to calculate the distribution of the end-to-end delay $D_i(a_i, \alpha_{-i})$ for the secondary user $SU_i$ to transmit its packets. The expected end-to-end delay $E[D_i]$ of the secondary user $SU_i$ can be expressed as:

$$E[D_i(a_i, \alpha_{-i})] = \sum_{j=1}^{M} s_{ij} E[D_j(R_{ij}(a_i, \alpha_{-i}))],\quad (10)$$

where $E[D_j(R_{ij}(a_i, \alpha_{-i}))]$ is the average end-to-end delay if the secondary user $SU_i$ chooses the frequency channel $F_j$. Recall that $R_{ij} = [T_{ij}, p_{ij}] \in \mathcal{R}$ is the channel condition of the channel $F_j$ for the secondary user $SU_i$. Note that $s_{ij}$ is the strategy of the action $a_{ij}$.

We use the queuing model in Figure 3 and apply queuing theory to calculate the end-to-end delay.

In Figure 3, for a specific secondary user $SU_i$, each arriving packet will select a physical queue to join (action $a_{ij}$) according to the strategy $s_{ij}$. Assume that once a packet selects a physical queue, it cannot jockey to another queue (change position to the other queues). Thus, a queued packet waits in line to be served on the selected frequency channel.

Note that there are $N$ physical queues from $N$ secondary users for a frequency channel $F_j$. Only one of them can transmit its packets at any time. Hence, we form a “virtual queue” for the same frequency channel (see Figure 4). In a virtual queue, the packets of the different secondary users wait to be transmitted. Importantly, the total sojourn time (queue waiting time plus the transmission service time) of this virtual queue now becomes the actual service time at each of the physical queues. The concept is similar to the “service on vacation” (Bertsekas & Gallager, 1987) in queuing theory, and the waiting time of the virtual queue can be regarded as the “vacation time”.

![Diagram](image.png)

Fig. 4 Priority virtual queue for a specific frequency channel.
Since the number of the secondary users in a regular cognitive radio network is usually large, we can approximate the virtual queue using the FIFO (First-In-First-Out) M/G/1 queuing model (i.e. when $N \to \infty$, the input traffic of the virtual queue can be modeled as a Poisson process). The average arrival rate of the virtual queue of the frequency channel $F_j$ is $\sum_{i=1}^{N} r_{ij}$. Let us denote the first two moments of the service time for the virtual queue of the frequency channel $F_j$ as $E[\tilde{X}_j]$ and $E[\tilde{X}_j^2]$. For a packet in the virtual queue of frequency channel $F_j$, we determine the probability of the packet coming from the secondary user SU$_i$ as:

$$f_{ij} = \frac{s_{ij}}{\sum_{k=1}^{N} s_{kj}}.$$  \hspace{1cm} (11)

Hence,

$$E[\tilde{X}_j] = \sum_{i=1}^{N} f_{ij} E[X_{ij}], \quad E[\tilde{X}_j^2] = \sum_{i=1}^{N} f_{ij} E[X_{ij}^2].$$  \hspace{1cm} (12)

Let $\tilde{W}_j$ and $\tilde{D}_j$ represent the queue waiting time and sojourn time of the virtual queue of the frequency channel $F_j$, respectively. $E[\tilde{W}_j]$ can be obtained from the Pollaczeck-Khinchin formula (Bertsekas & Gallager, 1987). Then, $E[\tilde{D}_j]$ can be obtained as:

$$E[\tilde{D}_j] = E[\tilde{W}_j] + E[\tilde{X}_j] = \frac{\sum_{i=1}^{N} r_{ij} E[\tilde{X}_j]}{2 \left(1 - \sum_{i=1}^{N} r_{ij} E[\tilde{X}_j]\right)} + E[\tilde{X}_j].$$  \hspace{1cm} (13)

Then, we apply G/G/1 approximation based on the work of (Abate et al., 1995; Jiang et al., 2001) for the virtual queuing delay distribution:

$$\text{Prob}(\tilde{D}_j > t) = \sum_{i=1}^{N} r_{ij} E[\tilde{X}_j] \exp\left(-\frac{t \sum_{i=1}^{N} r_{ij} E[\tilde{X}_j]}{E[\tilde{D}_j]}\right).$$  \hspace{1cm} (14)

This virtual queuing delay distribution is the service time distribution of the physical queues at the secondary users. Since the service time of the physical queue is an exponential distribution (see equation (14)), the average end-to-end delay of the secondary user SU$_i$ sending packets through frequency channel $F_j$ is approximately:

$$E[D_{ij}] = \frac{E[\tilde{D}_j]}{1 - r_{ij} E[\tilde{D}_j]}, \quad \text{for } r_{ij} E[\tilde{D}_j] < 1.$$  \hspace{1cm} (15)

Strategies $(s_i, s_{-i})$ such that $r_{ij} E[\tilde{D}_j] \geq 1$ will result in an unbounded delay $E[D_{ij}]$, which is undesirable for multimedia streaming. The advantage of this approximation is that once the average delay of the virtual queue $E[\tilde{D}_j]$ for a certain frequency channel $F_j$ is known...
by the secondary user $SU_i$, the secondary user can immediately calculate the expected end-to-end delay $E[D_{ij}]$ of a packet transmitting using the frequency channel $F_j$.

### 4.3 Queuing analysis with the impact of higher priority users

Based on the derivations in the previous subsection, we now consider the impact of primary users. First, let us consider the case that there are only two priority classes (i.e. $K = 2$, $PU \in C_1$, $SU \in C_2$). Note that in the introduced queuing model in Figure 4, the packets from the primary users will not be seen at the physical queues of the secondary users, but only have impact on the virtual queues of the frequency channels. Since the primary users are the first priority in each of the frequency channels, we modeled the virtual queues for a particular frequency channel using a priority M/G/1 queue instead of a FIFO M/G/1 queue. Recall that the average input rate of the primary user $PU_j$ is $\lambda_{j}^{PU}$, and the first two moments of the service time is $E[X_{j}^{PU}]$ and $E[X_{j}^{PU^2}]$. By applying the Mean Value Analysis (MVA) in queuing theory (Kleinrock, 1975), we modify equation (13) into a priority M/G/1 queuing case:

$$E[\hat{D}_j] = E[\hat{W}_j] + E[\hat{X}_j]$$

$$= \frac{r_{j}^{PU}E[(X_{j}^{PU})^2] + \sum_{i=1}^{N} r_{ij}E[\hat{X}_j]}{2 \left(1 - r_{j}^{PU}E[X_{j}^{PU}] \right) \left(1 - \sum_{i=1}^{N} r_{ij}E[\hat{X}_j] - r_{j}^{PU}E[X_{j}^{PU}] \right)} + E[\hat{X}_j],$$

$$= \frac{\rho_j^2 + \mu_j^2}{2 (1 - \rho_j)(1 - \rho_j - \mu_j)} + E[\hat{X}_j]$$

$\rho_j$ represents the normalized loading of the primary user $PU_j$ for the frequency channel $F_j$, and

$$\rho_j = r_{j}^{PU}E[X_{j}^{PU}], \quad \mu_j = r_{j}^{PU}E[(X_{j}^{PU})^2].$$

$\mu_j$ represents the summation of the normalized traffic loading of all the secondary users using the frequency channel $F_j$, and

$$\mu_j = \sum_{i=1}^{N} r_{ij}E[\hat{X}_j], \quad \mu_j^2 = \sum_{i=1}^{N} r_{ij}E[\hat{X}_j^2].$$

Hence, we substitute the $E[\hat{D}_j]$ of equation (16) into equation (15), and determine the average end-to-end delay $E[D_{ij}]$ of the secondary user $SU_i$ sending packets using frequency channel $F_j$ while considering the impact of the primary user $PU_j$.

The derivation can be generalized to $K$ priority classes among users ($K > 2$, $PU \in C_1$, $SU \in \{C_2, ..., C_K\}$). Similar to the two priority classes’ case, the priority queuing model only affects the virtual queues for different frequency channels. Since the secondary users now have different priorities, the secondary users in different priority classes will experience different virtual queuing delay. Let $E[\hat{D}_{ijk}]$ represent the virtual queuing delay.
experienced by the secondary user in class $C_k$ in the virtual queue for the frequency channel $F_j$. Let $\mu_{jk}$ represent the normalized traffic loading of all the class $C_k$ secondary users using the frequency channel $F_j$. Based on the definition in the two priority users’ case, we have:

$$\mu_{jk} = \sum_{\forall SU_i \in C_k} r_{ij}E[\tilde{X}_j], \text{ and } \mu_{jk}^2 = \sum_{\forall SU_i \in C_k} r_{ij}E[\tilde{X}_j^2].$$

(19)

By applying the Mean Value Analysis (MVA) (Kleinrock, 1975), we have:

$$E[\tilde{D}_{jk}] = E[\tilde{W}_{jk}] + E[\tilde{X}_j] = \frac{\rho_j^2 + \sum_{l=2}^{k} \mu_{jk}^2}{2(1 - \rho_j - \sum_{l=2}^{k} \mu_{jl})(1 - \rho_j - \sum_{l=2}^{k} \mu_{jl})} + E[\tilde{X}_j].$$

(20)

Hence, for a secondary user $SU_i \in C_k$ using the frequency channel $F_j$, its end-to-end delay $E[D_{ij}]$ and probability of packet loss $P_{ij}(\bar{a}_i(s_i), s_{-i})$ become:

$$E[D_{ij}] = \frac{E[\tilde{D}_{jk}]}{1 - r_{ij}E[\tilde{D}_{jk}]} \text{ for } r_{ij}E[\tilde{D}_{jk}] < 1, \text{ } SU_i \in C_k,$$

(21)

$$P_{ij}(\bar{a}_i(s_i), s_{-i}) = r_{ij} \exp\left(-\frac{r_{ij}(1 - r_{ij}E[\tilde{D}_{jk}])d_i}{E[\tilde{D}_{jk}]\rho_j}\right), \text{ for } SU_i \in C_k.$$ (22)

Therefore, we can approximate the objective function in equation (7) for the multimedia streaming of the secondary user $SU_i$ as (note that $\sum_{j=1}^{M} s_{ij} = 1$):

$$\max_{\bar{a}_i \in \mathcal{A}_i} \sum_{C_k \in C_i} \lambda_k L_k N_k \cdot P_k^{\text{succ}}(\bar{a}_i(s_i), s_{-i})$$

$$\Rightarrow \max_{\bar{a}_i \in \mathcal{A}_i} \sum_{C_k \in C_i} \lambda_k L_k N_k \cdot \sum_{j=1}^{M} s_{ij}(1 - P_{ij}(\bar{a}_i(s_i), s_{-i}))$$

(23)

$$\Rightarrow \max_{\bar{a}_i \in \mathcal{A}_i} \sum_{C_k \in C_i} \lambda_k L_k N_k \sum_{j=1}^{M} s_{ij}(1 - r_{ij} \exp\left(-\frac{r_{ij}(1 - r_{ij}E[\tilde{D}_{jk}(\bar{a}_i(s_i), s_{-i})])d_i}{E[\tilde{D}_{jk}(\bar{a}_i(s_i), s_{-i})]}\right))$$

Note that only $E[\tilde{D}_{jk}]$ depends on the strategies $s_{-i}$ of other secondary users.

We provide here an example that considers a simple network with two secondary users and three frequency channels (i.e. $N = 2$, $M = 3$). In the simple example, the behavior of the proposed cognitive radio model can be clearly understood. Assume that each secondary user can choose all three frequency channels. The two secondary users are in the same priority class. The simulation parameters of the secondary users are presented in Table 3 including the channel conditions $R_{ij} = [T_{ij}, p_{ij}]$, and initial strategies $s_i(0)$, etc. The normalized traffic statistics of the primary users are given in Table 4.
<table>
<thead>
<tr>
<th>Secondary users</th>
<th>Physical transmission rate $T_{ij}(\theta^{opt})$ (Mbps)</th>
<th>Physical packet error rate $p_{ij}(\theta^{opt})$</th>
<th>Rate requirement $B_i$ (Mbps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SU_1$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$SU_2$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Considered parameters of the secondary users in the example.

<table>
<thead>
<tr>
<th>Primary users</th>
<th>Normalized loading $\rho_j$</th>
<th>Second moment normalized loading $\rho_j^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$PU_1$</td>
<td>0.2</td>
<td>$1 \times 10^{-4}$</td>
</tr>
<tr>
<td>$PU_2$</td>
<td>0.1</td>
<td>$1 \times 10^{-4}$</td>
</tr>
<tr>
<td>$PU_3$</td>
<td>0.3</td>
<td>$1 \times 10^{-4}$</td>
</tr>
</tbody>
</table>

Table 4. Considered parameters of the primary users in the example.

Fig. 5 Analytical expected delay of the secondary users with various strategies in different frequency channels, shadow part represents a bounded delay below the delay deadline (stable region).
Given the statistics, Figure 5 provides the different strategy pairs \((s_{1,j}, s_{2,j})\) in the three frequency channels that keep the analytical experienced delays \(E[D_{ij}]\) (using equation (21)) bounded by the delay deadlines for the two secondary users. Importantly, a strategy pair \((s_{1,j}, s_{2,j})\) that results in an unbounded \(E[D_{ij}]\) will make the multimedia quality drop abruptly for the delay-sensitive applications, which is undesirable for these secondary users. Figure 5 clearly shows when the channel conditions become worse (from \(F_{1}\) to \(F_{3}\)), the selectable frequency strategy pairs becomes less. Hence, equation (21) provides the analytical operation points for the frequency selection strategy pairs.

4.4 Realistic framework for multimedia transmission over cognitive radio networks using queuing analysis

The priority virtual queue analysis requires the following information to compute \(\mu_{i,j}\) and \(\mu_{j,i}^2\) in (20):

- Priority: the secondary users' priorities.
- Normalized loading: the secondary users' normalized loading parameters \(r_{ij}E[\bar{X}_{j}]\), which not only include the information of \(s_{i,j}\), but also reflects the input traffic loading and the expected transmission time using a specific frequency channel.
- Variance statistics: the secondary users' variance statistics with the normalized parameter \(r_{ij}E[\bar{X}_{j}^2]\).

Hence, two kinds of information exchange are defined for the priority virtual queue analysis:

- Other secondary users' traffic specification \(TS_{i}\).
- The frequency selection information of the other secondary users to model the strategies \(s_{-i}\).

Figure 6 shows the block diagram of the introduced priority virtual queue interface (Shiang & van der Schaar, 2008) together with the cross-layer optimization approach in Section 3.1. Since the traffic specification \(TS_{i}\) only varies when the frequency channels change dramatically, the traffic specification can be exchanged only when a secondary user joins the network. On the other hand, the frequency selection information can be exchanged more frequently (e.g. once per service interval in van der Schaar et al., 2006). Note that since the users in the higher priority classes will not be affected by the users in the lower priority classes, they do not need the information from the users in a lower priority class. Hence, the information exchanges (overheads) and computational complexity will be scalable and will increase as the traffic priority decreases, thereby benefiting the high priority and low-delay traffic.
5. Conclusions

In this chapter, we discussed the priority virtual queuing architecture for heterogeneous and autonomous secondary users in cognitive radio networks, based on which they can time-share the various frequency channels in a distributed fashion. With the information exchange defined by the proposed interface, the secondary users can build an abstraction of the dynamic wireless environment as well as the competing wireless users’ behaviors and learn how to efficiently adapt their transmission strategies for multimedia streaming. Importantly, unlike conventional channel allocation schemes that select the least interfered channel merely based on the channel estimation, the introduced multi-agent priority queue modeling allows the secondary users to track the other users and adequately adapt their own transmission strategies to the changing multi-user environment. It can be shown that the introduced cross-layer optimization that applies priority queuing analysis significantly outperforms the fixed channel allocation and the current dynamic channel allocation that selects the least interfered channel, in terms of multimedia quality. Finally, we discuss the required information exchange that is required for the queuing analysis and present a realistic framework for the secondary users to transmit multimedia traffic over cognitive radio networks.

References

13, Sep 2006.


