Chapter 11
Real-Time Multimedia Transmission over Cognitive Radio Networks

Haiyan Luo, Song Ci, Dalei Wu, Zhiyong Feng, and Hui Tang

Abstract Cognitive radio (CR) has been proposed as a promising solution to improve connectivity, self-adaptability, and efficiency of spectrum usage. When used in video applications, user-perceived video quality experienced by secondary users is a very important performance metric to evaluate the effectiveness of CR technologies. However, most current research only considers spectrum utilization and effectiveness at MAC and PHY layers, ignoring the system performance of upper layers. Therefore, in this chapter we aim to improve the user experience of secondary users for wireless video services over cognitive radio networks. We propose a quality-driven cross-layer optimized system to maximize the expected user-perceived video quality at the receiver end, under the constraint of packet delay bound. By formulating network functions such as encoder behavior, cognitive MAC scheduling, transmission, as well as modulation and coding into a distortion-delay optimization framework, important system parameters residing in different network layers are jointly optimized in a systematic way to achieve the best user-perceived video quality for secondary users in cognitive radio networks. Furthermore, the proposed problem is formulated into a MIN-MAX problem and solved by using dynamic programming. The performance enhancement of the proposed system is evaluated through extensive experiments based on H.264/AVC.

11.1 Introduction

With the fast development of wireless communication technologies, the limited unlicensed spectrum bands can no longer meet the increasing requirement. Wireless multimedia applications require significant bandwidth with relatively tight delay constraints. With limited radio spectrum, bandwidth is considered to be one of the major bottlenecks for high-quality multimedia wireless applications. Frequency spectrum has become the scarcest resource in the next-generation wireless multimedia networks. However, statistics indicate that a large amount of the licensed spectrum bands are under-utilized, due to the spectrum strategy of static allocation and centralized management. Research also shows that the utilization of both licensed

S. Ci (✉)
College of Engineering, University of Nebraska-Lincoln, Omaha, NE, USA
e-mail: sci@engr.unl.edu

F.R. Yu (ed.), *Cognitive Radio Mobile Ad Hoc Networks*,
DOI 10.1007/978-1-4419-6172-3_11,

and unlicensed frequency bands changes with time and location. Thus, efficient methods are necessary for spectrum sharing among different systems, applications, and services in a dynamic wireless environment [1]. Therefore, "cognitive radio" has emerged as a new design paradigm for next-generation wireless networks, aiming to increase the utilization of radio spectrum for both licensed and unlicensed bands.

Generally speaking, cognitive radio is an intelligent wireless communication system that is aware of its surrounding environment and uses the methodology of understanding-by-building to learn from the environment. Then, its internal states are adapted to respond to statistical variations of the incoming RF stimuli by changing certain operating parameters in real time. Here, there are two primary objectives: (1) to provide highly reliable communications whenever and wherever needed and (2) to achieve efficient utilization of the radio spectrum [2]. Figure 11.1 illustrates a typical scenario of wireless video transmission over cognitive radio networks, where primary users and secondary users share a block of frequency spectrum. In cognitive radio networks, only primary users are authorized to use the radio spectrum. Thus, secondary users have to search the idle channels to use at the beginning of every slot by performing channel sensing. Based on the sensing outcomes, secondary users will decide whether or not to access the sensed channels. It has been reported that some frequency bands in the radio spectrum are largely unused, while some are

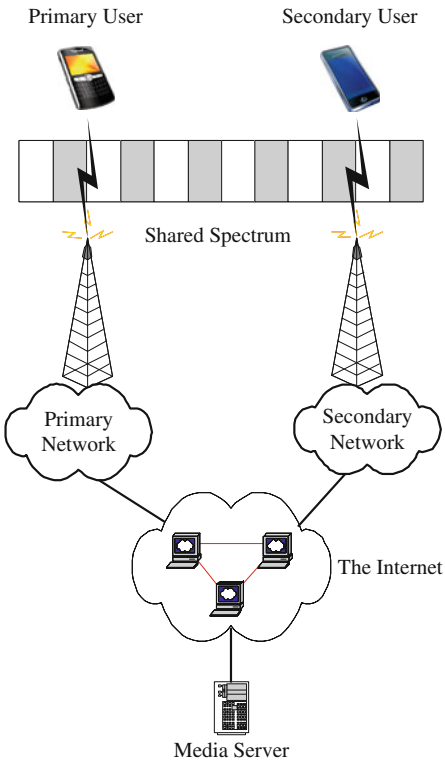

Fig. 11.1 Wireless video transmission over cognitive radio networks, where primary users and secondary users access the network through shared radio frequency band

heavily used. In particular, while a frequency band is assigned to a primary wireless system/service at a particular time and location, the same frequency band is unused by this wireless system/service in other times and locations. This results in spectrum holes (a.k.a spectrum opportunities) [3]. Therefore, by allowing secondary users to utilize these spectrum holes, spectrum utilization can be improved substantially.

Indeed, improving system performance of secondary users is a vital factor that hinges on the success of cognitive radio technologies for wireless video applications. Actually, if upper layer system performance is not well considered, the video quality degradation perceived by secondary users can largely impede the successful deployment of cognitive radio technologies. Also, user experience is by far the most important performance metric for end users in wireless video transmission. Therefore, it is critical to improve the user-perceived video quality at the receiver side for secondary users in cognitive radio networks. Although a lot of research activities have been conducted in cognitive radio networks, most of them only consider sensing effectiveness and spectrum utilization at the system design level. Other design metrics of the upper layers, such as the user-perceived video quality, have been mostly ignored.

In this chapter, we first briefly introduce two major new design technologies that can be used to improve the overall network performance of video applications over cognitive radio networks. Then, to achieve the best user-perceived video quality at the receiver side for secondary users of real-time wireless video transmission over cognitive radio networks, we present a quality-driven cross-layer system for joint optimization of system parameters residing in the entire network protocol stack. In the presented system, time variations of primary network usage and wireless channels are modeled, based on which, the encoder behavior, cognitive MAC scheduling, transmission, and modulation, and coding are jointly optimized for secondary users in a systematic way under a distortion-delay framework for the best video quality perceived by secondary users. The presented problem is formulated as a MIN-MAX problem and solved by using dynamic programming [4, 5]. Furthermore, the experiments prove that the proposed joint optimization system can greatly improve user experience at the receiver side for real-time wireless video services over cognitive radio networks.

11.2 Design Background

To increase resource utilization and improve the overall performance of cognitive radio networks, different design ideas have been proposed in recent years. Generally, two schools of approaches were presented: game theoretic approach and cross-layer-based optimization. In game theoretic approach, the behaviors of different competing users in the same frequency are modeled and a maximized utility function is formulated to dynamically adapt the channel selection strategy. In cross-layer optimization strategy, the interactions of different layers are jointly considered to adapt to the varying wireless channel quality to make better usage of the wireless channel resources.

11.2.1 Game Theory

Game theory is a discipline aimed at modeling situations in which decision makers have to make specific actions that have mutual, possibly conflicting, consequences. It has been used primarily in economics to model competition between companies. For example, should a given company enter a new market, considering that its competitors could make similar (or different) moves [6, 7]?

Game theory has also been applied to other areas, including politics and biology. Not surprisingly, it has also been applied to networking, in most cases to solve routing and resource allocation problems in a competitive environment. In recent years, it has been applied by an increasing number of researchers to resolve traditional issues in wireless networks. The common scenario is that the decision makers in the game are rational users or networks operators who control their communication devices. These devices have to cope with a limited transmission resource (i.e., the radio spectrum) that imposes a conflict of interests. In an attempt to resolve this conflict, they can make certain moves such as transmitting now or later, changing their transmission channel, or adapting their transmission rate. Felegyhazi et al. in [8] summarized how game theory can be used to model the radio communication channel. By leveraging on four simple running examples, the authors in this chapter have introduced the most fundamental concepts of non-cooperative game theory.

In [9], the authors proposed a dynamic channel-selection solution for autonomous wireless users transmitting delay-sensitive multimedia applications over cognitive radio networks, in which various rate requirements and delay deadlines of heterogeneous multimedia users were also considered. An information exchange mechanism was proposed to manage available spectrum resources in a decentralized manner. Based on this, a priority virtual queue interface was proposed that determines the required information exchanges and evaluates the expected delays experienced by various priority traffic. Then, a dynamic strategy learning (DSL) algorithm is deployed at each user that exploits the expected delay and dynamically adapts the channel selection strategies to maximize the user's utility function. In their simulation, almost 2 dB of video quality improvement can be achieved by reducing packet loss rate.

In another paper [10], the scalable and delay-sensitive characteristics of multimedia data and the resulting impact on users' viewing experiences of multimedia content are explicitly involved in the proposed utility function. The spectrum allocation problem was then formulated as an auction game and a distributively auction-based spectrum allocation scheme.

Generally, using game theoretic approach in multimedia applications over cognitive radio networks is a new research direction that was emerging a few years ago. The coming years is expected to see more research findings in this regard.

11.2.2 Cross-layer Optimization

Cross-layer design over cognitive radio networks was briefly discussed in [11, 12]. In the proposed frameworks, each node in the network can sense and learn from the

wireless environment and then respond to environment changes by adapting system parameters. However, how to integrate the different layers was not discussed. Also, there were no experimental results to verify the proposed frameworks.

In [13, 14], a rake optimized power-aware scheduling (ROPAS) architecture was proposed for mobile ad hoc networks (MANs) to deal with the utilization of cognitive radio (CR) for dynamic channel allocation among the requesting applications while limiting the average power transmitted in each sub-band. In that work, the cross-layer interaction between medium access control (MAC) and physical (PHY) layers of cognitive radio networks was considered. The joint power control and link scheduling can reduce adjacent channel and multi-access interference.

Su et al. in [15] have proposed an opportunistic multi-channel MAC protocol, in which the spectrum sensing at the physical layer and the packet scheduling at the MAC layer are integrated for wireless ad hoc networks. The proposed MAC protocol enables secondary users to identify and utilize the leftover frequency spectrum to reduce the interference level. Two different policies on channel sensing, which includes random sensing policy and negotiation-based sensing policy, have been proposed to detect the availability of unused licensed channels.

In [16], a framework called cognitive resource manager (CMR) was proposed to optimize the network protocol stack as a whole. The exchange of network information between CMRs can avoid harmful interactions arising from local optimization methods. In addition, the proposed framework can adapt MAC and link parameters to choose the best possible settings for the applications running on top. However, no implementation details and experimental analysis were discussed.

Overall, most current research efforts on cross-layer design over cognitive radio networks only focus on the joint consideration of MAC layer and physical layer. The design objectives are limited due to the fact that only sensing effectiveness and spectrum utilization are used as the design criteria, while the performance at the upper layers has been largely ignored. In this chapter, we will focus on addressing this issue.

11.3 System Model for Video Transmission

In this section, we present a quality-driven cross-layer optimized system that includes different modules, such as video encoder module, cognitive MAC module, modulation and coding module, cross-layer optimization module, as well as wireless video transmission module. These system modules actually represent different network functions residing in different network layers. For example, the video encoder resides in the application layer. The cognitive MAC module resides in the MAC layer, while the modulation and coding module is in the physical layer. As shown in Fig. 11.2, the cross-layer optimization module is able to communicate with other system modules to adjust the network functions, by selecting the optimal system parameters within a distortion-delay optimization framework. In this way, the major network functions are jointly optimized to achieve the best user-perceived video quality over cognitive radio networks under the current network conditions.

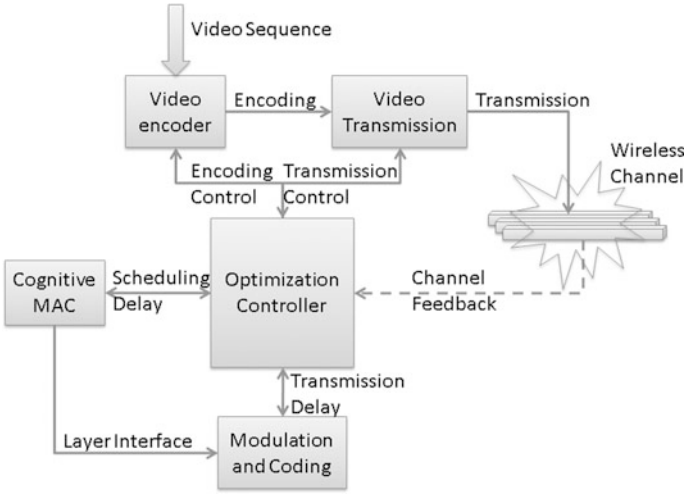

Fig. 11.2 The quality-driven cross-layer optimized system model for real-time wireless video transmission over cognitive radio networks

In the presented system, the expected video distortion calculated at the video encoder of the application layer is adopted as the objective function, which can be represented by encoding parameters (such as quantization step size q) and packet loss rate ρ. Furthermore, the expected packet delay, which is jointly represented by system parameters that affect encoder behavior, MAC scheduling, transmission, and modulation and channel coding, is used as the design constraint. Therefore, with the feedback information from the network such as RTT, queue length, and packet loss rate, the distortion-delay optimization module can choose the optimal set of parameters through the proposed cross-layer optimized system to achieve the best user-perceived video quality.

11.4 Video Quality Performance Metric

For video applications, the estimated video distortion is popularly used as the most important design metric. In this section, we present a formulation of end-to-end video distortion at the video encoder that can be used as the objective function in the proposed design.

Video encoder situates at the application layer. Without losing generality, we consider H.264 video codec in this chapter. In H.264 codec, each video frame is represented by block-shaped units of the associated luminance and chrominance samples (16×16 pixel region) called macroblocks (MBs). Furthermore, macroblocks can be both intra-coded and inter-coded from samples of previous video frames [17].

Intra-coding is performed in the spatial domain by referring to neighboring samples of previously coded blocks which are on the left and/or above the block to be predicted. Meanwhile, inter-coding is performed with temporal prediction from samples of previous video frames.

To estimate the end-to-end video distortion accurately, we need to consider all possible factors, which include source coding, error propagation, channel coding. Many research efforts can be found in literature on distortion estimation for hybrid motion-compensated video coding and transmission over lossy channels [18–21]. For real-time source coding, the estimated distortion caused by quantization, packet loss, and error concealment at the encoder can be calculated by using the "Recursive Optimal Per-pixel Estimate" (ROPE) method [18], providing an accurate optimization metric to the proposed system based on video quality. For the source coding parameter, we consider quantization step size (QP) q in this chapter.

According to the H.264 standard, one packet is set to be one row of macroblocks, which is also called one slice [17]. Therefore, slice and packet are two interchangeable concepts in this chapter. Given the dependencies introduced by error concealment scheme, the expected distortion of packet x of video frame n can be calculated at the encoder by using ROPE method as

$$\begin{aligned} E[D_{n,x}] = (1-\rho_x)E\left[D^r_{n,x}\right] \\ +\rho_x(1-\rho_{x-1})E\left[D^{lr}_{n,x}\right] + \rho_x\rho_{x-1}E\left[D^{ll}_{n,x}\right] \end{aligned} \tag{11.1}$$

where ρ_x is the loss probability of packet x with consideration of packet delay bound $T_n^{\max}$. $E\left[D^r_{n,x}\right]$ is the expected distortion of packet x when it is successfully received. Furthermore, depending on whether packet $(x-1)$ is received or lost, $E\left[D^{lr}_{n,x}\right]$ and $E\left[D^{ll}_{n,x}\right]$ are the corresponding expected distortion after concealment when packet x is lost. Therefore, the expected distortion of the whole video frame n can be represented as

$$E[D_n] = \sum_{x=1}^{X_n} E[D_{n,x}] \tag{11.2}$$

where X_n is the total number of packets in the video frame n. Thus, the expected end-to-end video distortion is accurately calculated by ROPE under instantaneous network conditions, which becomes the objective function in the proposed optimized system. For a given video packet x, the expected packet distortion only depends on packet error rate ρ_x and QP q. Considering the fact that the individual contribution of each path is continuously updated, this parameter is updated after each packet is encoded. In addition, the prediction and calculation of packet loss rate ρ_x will be discussed in Section 11.7. Readers can also refer to [4, 18] for detailed information regarding the calculation of the expected video distortion.

11.5 Channel Model

In this chapter, we assume that wireless channels are frequency flat, remaining time-invariant during a packet, but may vary from packet to packet. The channel quality is captured by the received signal-to-noise ratio (SNR) ξ. We adopt the Rayleigh channel model to describe ξ statistically. Therefore, the received SNR per packet is a random variable with a probability density function (pdf):

$$p(\xi) = \frac{1}{\bar{\xi}} \exp\left(-\frac{\xi}{\bar{\xi}}\right), \qquad \xi \geq 0 \tag{11.3}$$

where $\bar{\xi} := E\{\xi\}$ is the average received SNR. We also assume that the receiver has perfect channel side information and hence knows the instantaneous values of channel state information (CSI), while the transmitter has no such knowledge [22].

Moreover, we employ a cognitive channel model in which secondary users will try to transmit data when primary users are in presence. Secondary users will first perform channel sensing to detect the activity of primary users and then decide whether to transmit the data immediately or wait for the next available time slot depending on the detection result.

The different possible channel states for the primary usage are defined in Table 11.1.

Based on this definition, we depict the state transition model for cognitive radio transmission as shown in Fig. 11.3. In this figure, the state transition model is completely described by its stationary distribution of each channel state i and the state transition probability from state s_i to state s_j at the beginning of each time slot, which is denoted as p_{ij} $(1 \leq i, j \leq 4)$. Given the knowledge of channel fading, state transition probabilities, and primary user usage, the channel transition matrix $\mathbf{R}$ can be written as

$$\mathbf{R} = \{p_{ij}\} \quad (1 \leq i \leq 4, 1 \leq j \leq 4) \tag{11.4}$$

where

$$p_{ij} \geq 0, \forall i, j$$
$$\sum_{j=1}^{4} p_{ij} = 1, \forall i$$

Next, we will derive the transition matrix $\mathbf{R}$. In this chapter, we adopt noisy observations y_i under the Neyman–Pearson formulation to detect correlated random signals [23–25]. Therefore, channel sensing can be formulated as a hypothesis testing problem between the noise w_i and the signal s_i in noise.

$$\begin{aligned} H_0 &: y_i = w_i, \quad i = 1, ..., S_{\text{tot}} \\ H_1 &: y_i = s_i + w_i \quad i = 1, ..., S_{\text{tot}} \end{aligned} \tag{11.5}$$

Table 11.1 Possible channel states for the primary channel

State ID	State name	State description
s_1	correct detection	channel is idle, detected as idle
s_2	false alarm	channel is idle, detected as busy
s_3	missed detection	channel is busy, detected as idle
s_4	correct detection	channel is busy, detected as busy

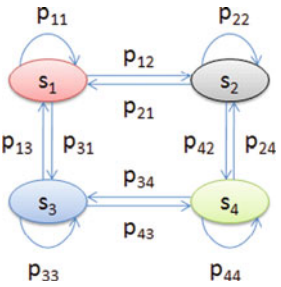

Fig. 11.3 The proposed state transition model for cognitive radio channel

where S_{tot} is the total symbols within a duration that is allocated to sense the channel. Received noise $\{w_i\}$ is a zero-mean, complex Gaussian random variable with variance σ_w^2 for all i, denoted as $\{w_i\} \backsim S_{\text{tot}}\left(0, \sigma_w^2\right)$. $\{s_i\}$ is the sum of the active primary users' faded signals arriving at the secondary receiver, which has a circularly symmetric complex Gaussian distribution with zero-mean and variance σ_s. Both $\{s_i\}$ and $\{w_i\}$ are assumed to be independent and identically distributed (i.i.d) [25].

Therefore, the optimal Neyman–Pearson detector of the above detection problem is given by

$$Y = \frac{1}{S_{\text{tot}}} \sum_{i=1}^{S_{\text{tot}}} \left| y_i^2 \right| \underset{H1}{\overset{H0}{\gtrless}} t_{\text{d}} \tag{11.6}$$

where t_{d} is the detection threshold. Thus, test statistic Y is chi-square distributed with $2S_{\text{tot}}$ degrees of freedom. Therefore, the probabilities of detection p_{d} and false alarm p_{f} can be represented as follows, respectively.

$$p_{\text{d}} = P_r(Y > t_{\text{d}}|H_1) = 1 - P\left(\frac{S_{\text{tot}} t_{\text{d}}}{\sigma_w^2 + \sigma_s^2}, S_{\text{tot}}\right) \tag{11.7}$$

$$p_{\text{f}} = P_r(Y > t_{\text{d}}|H_0) = 1 - P\left(\frac{S_{\text{tot}} t_{\text{d}}}{\sigma_w^2}, S_{\text{tot}}\right) \tag{11.8}$$

where $P(x, z)$ denotes the regularized lower gamma function. Denote $\Gamma(z)$ and $\gamma(x, z)$ the Gamma function and the lower incomplete gamma function, respectively, then $P(x, z)$ can be represented as

$$P(x, z) = \frac{\gamma(x, z)}{\Gamma(z)} \tag{11.9}$$

Therefore, the probability of channel being idle and detected as idle can be calculated from

$$p_{11} = (1 - p_{\mathrm{f}})(1 - \zeta_{\mathrm{b}}) \tag{11.10}$$

where ζ_{b} is the prior probability of channel being busy. Equation (11.10) indicates that the transition probability only depends on the current state, regardless of the original state. This also holds true for the other state transition probabilities. Denote

$$\begin{aligned} &p_{i1} = p_1; \quad p_{i2} = p_2 \\ &p_{i3} = p_3; \quad p_{i4} = p_4 \\ &\text{where} \qquad 1 \le i \le 4 \end{aligned} \tag{11.11}$$

Then, the transition probabilities can be represented as

$$\begin{aligned} p_1 &= (1 - p_{\mathrm{f}})(1 - \zeta_{\mathrm{b}}) \\ p_2 &= p_{\mathrm{f}}(1 - \zeta_{\mathrm{b}}) \\ p_3 &= (1 - p_{\mathrm{d}})\zeta_{\mathrm{b}} \\ p_4 &= p_{\mathrm{d}}\zeta_{\mathrm{b}} \end{aligned} \tag{11.12}$$

Hence, the 4×4 transition matrix $\mathbf{R}$ can be derived as

$$\mathbf{R} = \begin{pmatrix} p_1, p_2, p_3, p_4 \\ p_1, p_2, p_3, p_4 \\ p_1, p_2, p_3, p_4 \\ p_1, p_2, p_3, p_4 \end{pmatrix} \tag{11.13}$$

11.6 MAC Scheduling Delay

To formulate the MAC frame scheduling delay for secondary users in cognitive radio networks, we first denote T_0 the duration of a time slot, and t_{s} the corresponding channel sensing time allocated for each time slot, as shown in Fig. 11.4.

In the proposed channel model shown in Fig. 11.3, if the current channel is in state s_2 or s_4, the MAC frame has to wait for the next time slot. When new time slot arrives, it has to wait again if the channel is still in state s_2 or s_4. This process repeats until a time slot becomes available or until a maximum waiting threshold in terms of the number of time slots is reached, denoted as $N_s^{\max}$. If this maximum threshold is

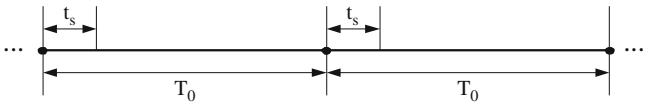

Fig. 11.4 Illustration of time slot duration and channel sensing time over cognitive radio networks

reached, the waiting packet has to be dropped from the sending queue. We call this truncated MAC scheduling. Furthermore, to ensure real-time video transmission, every video packet has to meet a delay bound. Therefore, all the video packets of the same video frame have the same delay bound. We denote $T_n^{\max}$ the delay bound of the video frame n. Then, $N_s^{\max}$ can be calculated as

$$N_s^{\max} = \left\lfloor \frac{T_n^{\max}}{T_0} \right\rfloor \tag{11.14}$$

Note that forward error control coding (FEC) and automatic repeat request (ARQ) are the two major error resilient approaches used by video encoder. However, ARQ is not always feasible for real-time video transmission, due to the excessive delay caused by retransmissions [26, 27]. Therefore, in this chapter, we do not consider ARQ. Instead, we will optimize system parameters in a holistic way to improve the overall system performance for secondary users.

As shown in Fig. 11.3, only when channel is in state s_1 can it be effectively used to transmit data for secondary users. When the channel is in state s_2, s_3, or s_4, it is not available for secondary users or it is not detected by the secondary users as available. Therefore, the probability p_w that a given MAC frame has to wait for the next time slot can be expressed as

$$p_w = 1 - (1 - p_{\mathrm{f}})(1 - \zeta_{\mathrm{b}}) \tag{11.15}$$

Thus, assuming the availability of time slots are independent, the average scheduling time at hop h can be represented as

$$\begin{aligned} t_{\mathrm{sched}}^h &= \left(p_w + p_w^2 + \cdots + p_w^{N_s^{\max}} \right) * T_0 \\ &= \left(\frac{p_w - p_w^{N_s^{\max}}}{1 - p_w} \right) * T_0 \end{aligned} \tag{11.16}$$

11.7 Transmission Delay

In this chapter, adaptive modulation and coding (AMC) technique at the physical layer is adopted on a packet-by-packet basis to enhance the throughput. With AMC, the optimal combination of different modulation constellations and different rates of error-control codes is selected based on the time-varying channel quality. For example, in good channel conditions, AMC schemes with larger constellation sizes and higher channel coding rate will guarantee the required packet error rate for QoS provisioning [28, 29]. Usually, bit error rate (BER) $\varepsilon(\xi)$ can be calculated from the following approximated expression:

$$\varepsilon(\xi) = a_m \mathrm{e}^{-b_m \xi} \tag{11.17}$$

Table 11.2 AMC schemes at the physical layer

AMC mode (m)	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$
Modulation Scheme	BPSK	QPSK	QPSK	16-QAM	16-QAM	64-QAM
Coding Rate (c_m)	1/2	1/2	3/4	9/16	3/4	3/4
r_m (bits/sym.)	0.50	1.00	1.50	2.25	3.00	4.50
a_m	1.1369	0.3351	0.2197	0.2081	0.1936	0.1887
b_m	7.5556	3.2543	1.5244	0.6250	0.3484	0.0871

where coefficients a_m and b_m can be obtained by fitting (11.17) to the exact BER, as shown in Table 11.2. Therefore, the frame error rate $\rho(\xi)$ can be expressed as

$$\rho(\xi) = 1 - (1 - \varepsilon(\xi))^{\ell} \tag{11.18}$$

where ℓ is the MAC frame size and ξ is the received signal-to-noise ratio (SNR).

Assume that the xth packet of video frame n is fragmented into Z MAC frames, and these MAC frames are transmitted along a path with H hops. Thus, the packet loss rate for a given packet x can be expressed as

$$\rho_x = 1 - \prod_{z=1}^{Z} \left(1 - \rho_x^z(\xi)\right) \tag{11.19}$$

where $\rho_x^z(\xi)$ is the frame error rate of MAC frame z of packet x.

Denote BW_h the bandwidth (symbols/second) of hop h, $\ell_x^z(m,h)$ the bit number of the zth physical frame of packet x at hop h. Therefore, the transmission delay can be represented as

$$t_{\text{trans}}^h = \left\lceil \frac{\frac{\ell_x^z(m,h)}{r_{m,h} \times \mathrm{BW}_h}}{T_0 - t_s} \right\rceil * T_0 \tag{11.20}$$

where $r_{m,h}$ is the rate (bits/symbol) of AMC mode m at hop h as shown in Table 11.2.

11.8 Problem Formulation and Optimal Solution

To achieve the best video quality at the receiver side, the expected end-to-end video distortion under the constraint of video packet delay should be minimized. With the proposed distortion-delay framework, the source coding, MAC scheduling, transmission, and modulation and coding are jointly optimized in a cross-layer fashion.

In wireless video transmission, all packets of a given frame f_n are constrained by a frame delay bound $T_n^{\max}$. Therefore, all packets of the video frame f_n have the same delay constraint $T_n^{\max}$. Denote $\mathbf{Q}$ as all possible operating points of source coding parameter (such as quantization step size $q_{n,x}$) of packet x of frame n, $\mathbf{M}$

as all possible modulation, and channel coding schemes $m_{n,x}$. Thus, the proposed problem can be formulated as

$$\min_{\{q_{n,x}\in\mathbf{Q},m_{n,x}\in\mathbf{M}\}} \sum_{n=1}^{N}\sum_{x=1}^{X_n} E[D_{n,x}]$$
$$s.t.: \max_{1\le x\le X_n} t_{n,x} \le T_n^{\max},\ \forall x \tag{11.21}$$

where N is the total number of video frames of the given video sequence, and X_n is the total number of packets generated from the nth video frame. In addition,

$$t_{n,x} = \sum_{z=1}^{Z}\sum_{h=1}^{H}\left(t_{\text{sched}}^{z,h} + t_{\text{trans}}^{z,h}\right) \tag{11.22}$$

where $t_{\text{sched}}^{z,h}$ and $t_{\text{trans}}^{z,h}$ have already been derived by using (11.16) and (11.20), respectively. Z and H are the MAC frame number and the hop number, respectively. In other words, the xth packet of video frame n is fragmented into Z MAC frames, and these MAC frames are transmitted on a path with H hops. Therefore, the proposed problem has been formulated into a MIN-MAX problem [30]. Denote α as the index of video packet x over the entire video clip. Thus, the parameter vector of packet x of video frame n can be represented as

$$\begin{aligned} \mathcal{V}_\alpha &:= [q_{n,x}, m_{n,x}] \\ \text{where } & q_{n,x}\in\mathbf{Q},\ \ m_{n,x}\in\mathbf{M} \\ \text{and } & \ 1\le\alpha\le N\times X_n \end{aligned} \tag{11.23}$$

where $q_{n,x}$ and $m_{n,x}$ are the quantization step size, the AMC mode of packet x of the nth video frame, respectively. In addition, $\mathbf{Q}$ and $\mathbf{M}$ are the sets of all the possible values of $q_{n,x}$ and $m_{n,x}$, respectively.

To solve the MIN-MAX problem as shown in (11.21), we first convert it into an unconstrained optimization problem. According to the formulation (11.21), any parameter vector $\mathcal{V}_\alpha$ resulting in the expected packet delay greater than the constraint $T_n^{\max}$ cannot be the optimal parameter vector $\mathcal{V}_\alpha^*$, defined as $\mathcal{V}_\alpha^* := \left[q_{n,x}^*, m_{n,x}^*\right]$. Therefore, the objective function can be re-written as

$$E[D_{n,x}] = \begin{cases} \infty: & t_{n,x} > T_n^{\max} \\ E[D_{n,x}]: & t_{n,x} \le T_n^{\max} \end{cases} \tag{11.24}$$

where the average distortion of a packet with expected delay greater than the delay bound $T_n^{\max}$ is set to infinity, meaning that the corresponding parameter vector of the possible solution will not satisfy the packet delay bound $T_n^{\max}$. In this way, the minimum distortion problem with delay constraint is transformed into an unconstrained optimization problem. Note that most modern source codecs such

as H.264 [17] adopt error concealment strategies to improve the visual quality, which also introduces dependencies among slices/packets. Therefore, if the error concealment algorithm uses the motion vector of the previous slice to recover the lost slice, it will cause the calculation of the expected distortion of the current slice to depend on its previous slice. As mentioned earlier, a packet is generated from one slice. Therefore, without losing generality, we assume that the current packet depends on its previous θ packets ($\theta \geq 0$) in error concealment. To solve the formulated optimization problem in (11.21), we define a cost function $f_{\mathbb{k}}(\mathcal{V}_{\mathbb{k}-\theta}, ..., \mathcal{V}_{\mathbb{k}})$ to represent the minimum average distortion up to and including the $\mathbb{k}$th packet, where $\mathcal{V}_{\mathbb{k}-\theta}, ..., \mathcal{V}_{\mathbb{k}}$ are decision vectors of the $(\mathbb{k}-\theta)$th to $\mathbb{k}$th packets. Denote τ the total packet number of the video sequence, where $\tau := N \times X_n$. Therefore, $f_\tau(\mathcal{V}_{\tau-\theta}, ..., \mathcal{V}_\tau)$ represents the minimum distortion incurred by all packets of the given video sequence. Thus, solving (11.21) is essentially to solve the following equation

$$\min_{\{\mathcal{V}_{\tau-\theta},...,\mathcal{V}_\tau\}} f_\tau(\mathcal{V}_{\tau-\theta}, ..., \mathcal{V}_\tau) \tag{11.25}$$

Therefore, given the cost function $f_{\mathbb{k}-1}(\mathcal{V}_{\mathbb{k}-\theta-1}, ..., \mathcal{V}_{\mathbb{k}-1})$ and the $\theta+1$ decision vectors $\mathcal{V}_{\mathbb{k}-\theta-1}, ...\mathcal{V}_{\mathbb{k}-1}$ for the $(\mathbb{k}-\theta-1)$th to the $(\mathbb{k}-1)$th packets, the selection of the next decision vector $\mathcal{V}_{\mathbb{k}}$ is independent of the selection of the previous decision vectors $\mathcal{V}_1, \mathcal{V}_2, ..., \mathcal{V}_{\mathbb{k}-\theta-2}$. This means that the cost function can be expressed recursively as

$$\begin{aligned} f_k(\mathcal{V}_{\mathbb{k}-\theta}, .., \mathcal{V}_{\mathbb{k}}) = \min_{\{\mathcal{V}_{\mathbb{k}-\theta-1},...,\mathcal{V}_{\mathbb{k}-1}\}} \{ & f_{\mathbb{k}-1}(\mathcal{V}_{\mathbb{k}-\theta-1}, ..., \mathcal{V}_{\mathbb{k}-1}) \\ & + E[D_{\mathbb{k}}]\} \end{aligned} \tag{11.26}$$

which implies that the next step of the optimization process for the cost function is independent of its past steps, forming the foundation of dynamic programming.

Essentially, (11.26) can be further converted into and solved as a graph theory problem of finding the shortest path in a directed acyclic graph (DAG) [31]. By using dynamic programming to solve this shortest path problem, the computational complexity of the algorithm is decreased to $\mathbf{O}(\tau \times |\mathcal{V}|^{\theta+1})$ (where $|\mathcal{V}|$ is the cardinality of $\mathcal{V}$), depending directly on the value of θ. For most cases, θ is a small number, so the computational complexity of the algorithm is effectively decreased, compared with the exponential computational complexity of exhaustive search algorithm [32, 33].

11.9 Experimental Analysis

11.9.1 Experimental Environment

In this chapter, video coding is performed by H.264/AVC JM 15.1 codec. The video sequence "Foreman" is adopted for performance analysis. The first 100 frames of

the QCIF (176×144) video clip are coded at frame rate of 30 frames/s, and each I frame is followed by 9 P frames. Assume the whole packet/slice is lost if one of the MAC frames of the packet is lost, which is reasonable since usually intra-prediction is derived from the decoded samples of the same decoded slice. To avoid prediction error propagation, a 10% macroblock level intra-refreshment is used during the experiments. When a packet is lost during transmission, the temporal-replacement error concealment strategy will be used. The motion vector of a missing MB can be estimated as the median of motion vectors of the nearest three MBs in the preceding row. If that row is also lost, the estimated motion vector is set to zero. The pixels in the previous frame, pointed by the estimated motion vector, are used to replace the missing pixels in the current frame.

In the experiments, quantization step size (QP) q and AMC mode m of each packet are considered as the parameters to be optimized. The possible values of QP are chosen from 1 to 50, while the available AMC schemes are 1–6 as shown in Table 11.2. According to [25], the channel is assumed to be busy with an average probability of $\zeta_b = 0.1$. Also, since performance enhancement of secondary users in cognitive radio networks is the main focus of this chapter, we set the detection threshold $t_d = 1.35$, so that false alarm probabilities are effectively decreased. Therefore, under this setting, the channel sensing is reliable, and the interference to primary users is minimal.

Given an average SNR $\bar{\xi}$, the instantaneous link quality ξ can be randomly produced from (11.3). In this chapter, the link bandwidth is set to $100k$ symbols/s. Moreover, without losing generality, a single hop scenario is considered in the experiments to verify the performance of the proposed framework. Similar conclusions derived from the single hop scenarios may straightly apply to the multi-hop scenarios when the channel state information (CSI) of each hop is available.

Also, the delay bound is set in accordance with the frame rate, as adopted in literature [34–36]. As the most important performance metric for video applications [37, 38], peak signal-to-noise ratio (PSNR) of the received video frames of secondary users is used as the performance metric to compare the proposed system with the existing system, which has fixed AMC schemes.

11.9.2 Performance Evaluation

The performance enhancement of the proposed system for secondary users under various packet delay bounds is verified in Fig. 11.5, where time slot duration T_0 is set to 5 ms, average SNR ξ is set to 15 dB, and channel sensing time t_s is set to 0.5 ms and 1 ms, respectively. From the figure, we can observe that by jointly optimizing the system parameters residing in different network layers under the proposed system, significant performance improvement can be achieved. Another observation is that as the delay bound becomes more and more stringent, the

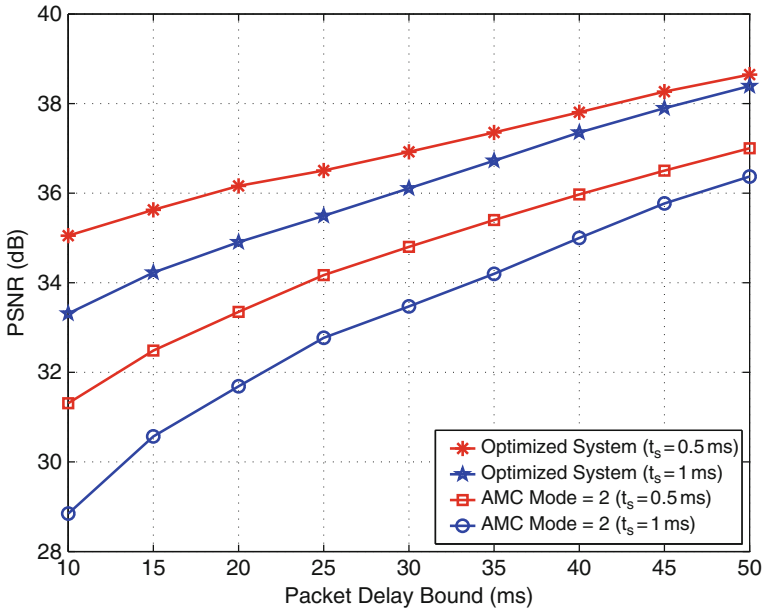

Fig. 11.5 Average PSNR comparison between the proposed quality-driven cross-layer optimized system and the existing system for real-time wireless video transmission over cognitive radio networks with various packet delay bounds $T_n^{\max}$

performance gain becomes higher and higher. This implies that the proposed system is especially suitable for real-time video transmission over cognitive radio networks with stringent delay bound. Furthermore, we can also observe that under the same network conditions, if the channel sensing time becomes longer, the performance gain actually becomes higher. This indicates that the proposed system might be useful for cognitive networks, due to the fact that when more time is allocated to perform channel sensing, MAC scheduling, and ARQ management, the negative impact on the overall system performance is minimized.

In Fig. 11.6, the visual comparison of one video frame (the 40th frame of the Foreman video clip) is presented with the packet delay bound being set to 30 ms and t_s 0.5 ms. Other environment settings remain the same with those of the above figure. Thus, it is observable that the user-perceived video quality has been greatly improved.

We also evaluate the relationship between the channel quality SNR ξ and the perceived video quality under the proposed system as shown in Fig. 11.7. In this experiment, time slot duration T_0 is set to 5 ms and the channel sensing time t_s is 0.5 ms. The packet delay bound is set to 20 ms and 30 ms, respectively. From the figure, we can observe that the proposed system can significantly improve PSNR performance, especially when the channel quality is not good or/and the delay bound is more stringent.

Fig. 11.6 The visual comparison between the proposed quality-driven cross-layer optimized system and the existing system for real-time wireless video transmission over cognitive radio networks. (**a**) Original frame; (**b**) Without optimization; (**c**) Optimized

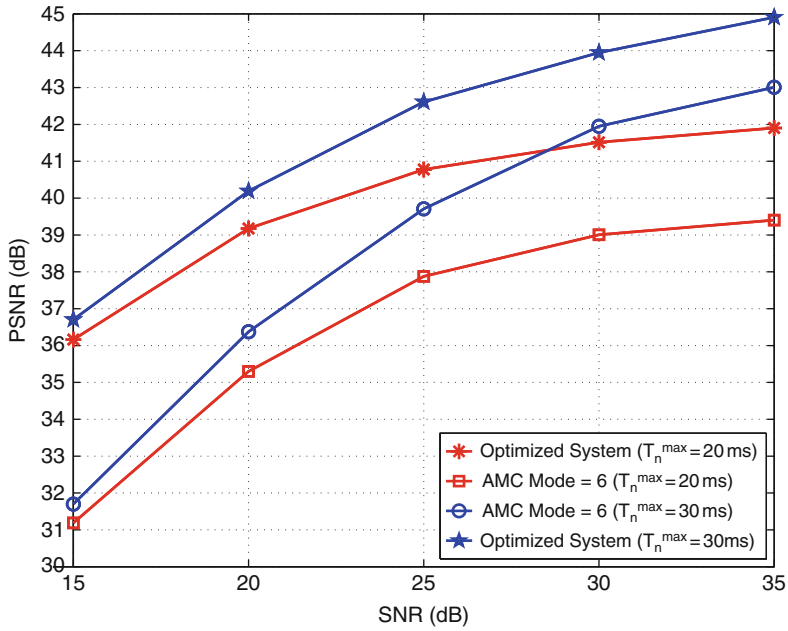

Fig. 11.7 Average PSNR comparison between the proposed quality-driven cross-layer optimized system and the existing system for real-time wireless video transmission over cognitive radio networks with various SNRs

Furthermore, the impact of time slot duration T_0 on secondary users is shown in Fig. 11.8. In this case, the channel sensing time t_s is set to 0.5 ms and average SNR ξ is 15 dB. The packet delay bound $T_n^{\max}$ is set to 20 ms and 30 ms, respectively. From the figure, we can observe that the proposed system achieves higher performance

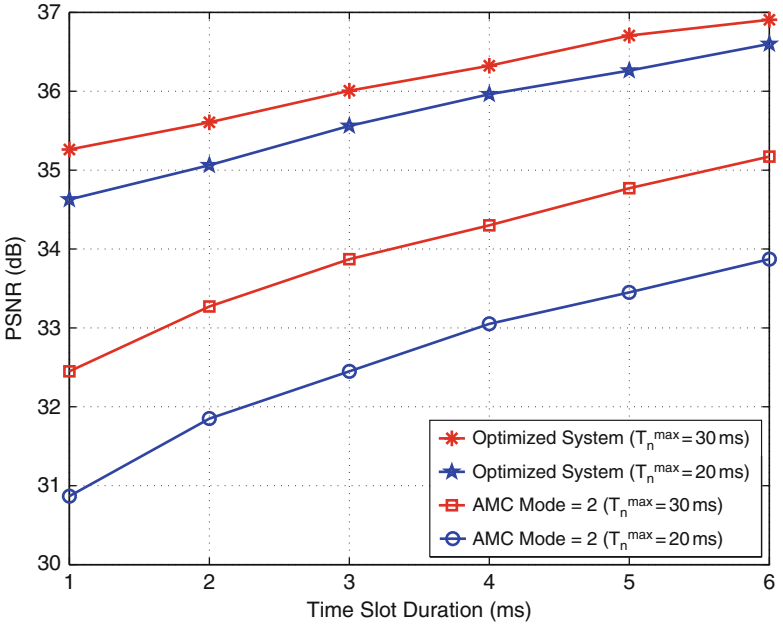

Fig. 11.8 Average PSNR comparison between the proposed quality-driven cross-layer optimized system and the existing system for real-time wireless video transmission over cognitive networks with various time slot durations T_0

gain when the time slot duration T_0 is smaller under the same packet delay bound $T_n^{\max}$. Another observation is that when time slot durations T_0 are the same, higher performance gain can be achieved when channel sensing time t_s is set to 1 ms than that achieved by 0.5 ms. This indicates that under more stringent delay bound, the proposed system can adapt the time-varying channel and choose the optimal system parameters to achieve the best user-perceived video quality.

The effect of channel sensing time t_s on secondary users is also studied. Here, the time slot duration T_0 is set to 5 ms and SNR is 15 dB. The packet delay bound $T_n^{\max}$ is set to 20 ms and 30 ms, respectively. As shown in Fig. 11.9, by increasing t_s, the overall performance decreases. This is reasonable because when the time spent on channel sensing increases, the time spent on transmission $(T_0 - t_s)$ decreases accordingly. However, the proposed system is able to achieve higher performance gain when more time is spent on channel sensing, thus minimizing the negative impact on the overall system performance.

In summary, all the experimental results demonstrate the significant performance enhancement of the proposed system for secondary users in cognitive radio networks. The experimental results also indicate that the performance gain is usually higher when the wireless channel experiences bad quality in a more stringent delay-bounded video application.

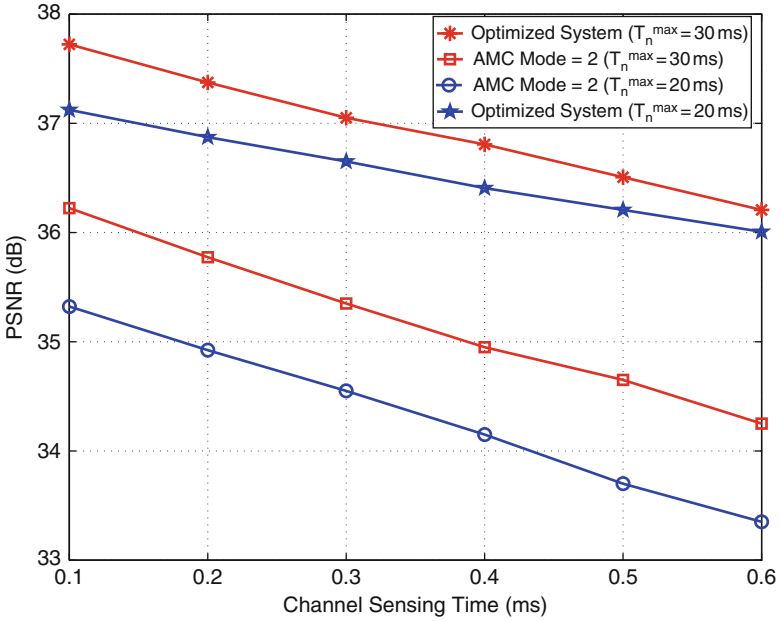

Fig. 11.9 Average PSNR comparison between the proposed quality-driven cross-layer optimized system and the existing system for real-time wireless video transmission over cognitive networks for various channel sensing times t_s

11.10 Conclusions

In this chapter, a cross-layer optimized system for real-time video transmission over cognitive radio networks has been studied. The proposed system can achieve the best possible video quality for secondary users, significantly improving the user experience of secondary users in cognitive radio networks, leading to the possibility of wide deployment of CR technologies in video applications. The design problem is presented to minimize the expected video distortion under the constraint of packet delay bound, which has been formulated as a MIN-MAX problem and solved by using dynamic programming. Experimental results have validated the effectiveness of the proposed system.

References

1. D. Niyato and E. Hossain, "Cognitive radio for next-generation wireless networks: An approach to opportunistic channel selection in IEEE 802.11-based wireless mesh," *IEEE Wireless Commun.*, vol. 16, pp. 46–54, Feb 2009.
2. S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 201–220, Feb 2005.
3. R. Tandra, S. Mishra, and A. Sahai, "What is a spectrum hole and what does it take to recognize one?" *Proceedings of the IEEE*, Apr 2009.

4. H. Luo, D. Wu, S. Ci, A. Argyriou, and H. Wang, "Quality-driven TCP friendly rate control for real-time video streaming," *IEEE GlobeCom*, Dec. 2008.
5. H. Luo, A. Argyriou, D. Wu, and S. Ci, "Joint source coding and network-supported distributed error control for video streaming in wireless multi-hop networks," *IEEE Transactions on Multimedia*, vol. 11, no. 7, pp. 1362–1372, Nov 2009.
6. D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA: MIT Press, 1991.
7. R. Gibbons, *A Primer in Game Theory*. Upper Saddle River, NJ: Prentice Hall, 1992.
8. M. Felegyhazi and J.-P. Hubaux, "Game theory in wireless networks: A tutorial," *EPFL technical report, LCA-REPORT-2006-002*, Feb 2006.
9. H. Shiang and M. van der Schaar, "Queuing-based dynamic channel selection for heterogeneous multimedia applications over cognitive radio networks," *IEEE Transactions on Multimedia*, vol. 10, no. 5, pp. 896–909, Aug 2008.
10. Y. Chen, Y. Wu, B. Wang, and K. Liu, "An auction-based framework for multimedia streaming over cognitive radio networks," *ICASSP 2010*, vol. 48, no. 2, pp. 2350–2353, Mar 2010.
11. Y. Yu, L. Wang, and Q. Yu, "Cross-layer architecture in cognitive ad hoc networks," *Commun. Mobile Comput*., vol. 2, pp. 47–51, Jan 2009.
12. S. Ci and J. Sonnenberg, "A cognitive cross-layer architecture for next-generation tactical networks," *IEEE MILCOM*, vol. 77, pp. 1–6, Oct 2007.
13. C. Ghosh and D. P. Agrawal, "ROPAS: Cross-layer cognitive architecture for wireless mobile adhoc networks," *Cognitive Radio Oriented Wireless Networks Commun*., pp. 514 – 518, Aug 2007.
14. C. Ghosh, B. Xie, and D. P. Agrawal, "ROPAS: Cross-layer cognitive architecture for mobile UWB networks," *IEEE International Conference on Mobile Adhoc and Sensor Systems,* pp. 1–7, Pisa, Italy, Oct. 2007.
15. H. Su and X. Zhang, "Cross-layer based opportunistic MAC protocols for QoS provisionings over cognitive radio wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 118–129, Jan. 2008.
16. P. Mahonen, M. Petrova, J. Riihijarvi, and M. Wellens, "Cognitive wireless networks: Your network just became a teenager," *In Proceedings of IEEE INFOCOM*, Barcelona, 2006.
17. "Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264ąISO/IEC 14496-10 AVC)," *ftp://ftp.imtc-files.org/jvt-experts/2003-03-Pattaya/JVT-G50r1.zip*, May 2003.
18. R. Zhang, S. L. Regunathan, and K. Rose, "Video coding with optimal inter/intra-mode switching for packet loss resilience," *IEEE J. Select. Areas Commun.*, vol. 18, no. 6, pp. 966–976, Jun 2000.
19. D. Wu, T. Hou, W. Zhu, H.-J. Lee, T. Chiang, Y.-Q. Zhang, and H. J. Chao, "On end-to-end architecture for transporting MPEG-4 video over the Internet," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, pp. 923–941, Sep. 2000.
20. G. Cote, S. Shirani, and F. Kossentini, "Optimal mode selection and synchronization for robust video communications over error prone networks," *IEEE J. Sel. Areas Commun.*, vol. 18, pp. 952–965, Jun 2000.
21. A. Argyriou, "Real-time and rate-distortion optimized video streaming with TCP," *Elsevier Signal Process. Image Commun*., vol. 22, pp. 374–388, Apr 2007.
22. Q. Liu, S. Zhou, and G. Giannakis, "Cross-layer combining of adaptive modulation and coding with truncated arq over wireless links," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1746–1755, Sept 2004.
23. H. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. Springer, New York, NY, 1994.
24. Y. Sung, L. Tong, and H. V. Poor, "Neyman-Pearson detection of Gauss-Markov signals in noise: Closed-form error exponent and properties," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1354–1365, Apr 2006.
25. S. Akin and M. C. Gursoy, "Effective capacity analysis of cognitive radio channels for quality of service provisioning," *CoRR*, vol. abs/0906.3888, 2009.

26. H. Wang, S. Tsaftaris, and A. K. Katsaggelos, "Joint source-channel coding for wireless object-based video communications utilizing data hiding," *IEEE Trans. Image Process.*, vol. 15, no. 8, pp. 2158–2169, Sept 2008.
27. T. Stockhammer, M. M. Hannuksela, and T. Wiegand, "H.264/AVC in wireless environments," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 657–673, Jul 2003.
28. M. S. Alouini and A. J. Goldsmith, "Adaptive modulation over Nakagami fading channels," *Kluwer J. Wireless Commun.*, vol. 13, pp. 119–143, May 2000.
29. A. Doufexi, S. Armour, M. Butler, A. Nix, D. Bull, J. McGeehan, and P. Karlsson, "A Comparison of the HIPERLAN/2 and IEEE 802.11a Wireless LAN Standards," *IEEE Commun. Magazine*, vol. 40, pp. 172–180, May 2002.
30. Z. Li, G. Schuster, and A. Katsaggelos, "MINMAX optimal video summarization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, pp. 1245–1256, Oct 2005.
31. G. M. Schuster and A. K. Katsaggelos, *Rate-Distortion Based Video Compression: Optimal Video Frame Compression and Object Boundary Encoding*. Norwell, MA: Kluwer, 1997.
32. A. Ortega and K. Ramchandran, "Rate-distortion methods for image and video compression," *Signal Process.*, vol. 15, pp. 23–50, Nov 1998.
33. D. Wu, S. Ci, and H. Wang, "Cross-layer optimization for video summary transmission over wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 4, pp. 841–850, May 2007.
34. Y. Andreopoulos, N. Mastronade, and M. van der Schaar, "Cross-layer optimized video streaming over wireless multi-hop mesh networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 11, pp. 2104–2115, Nov 2006.
35. P. Pahalawatta, R. Berry, T. Pappas, and A. Katsaggelos, "Content-aware resource allocation and packet scheduling for video transmission over wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 4, pp. 749–759, May 2007.
36. E. Maani, P. Pahalawatta, R. Berry, T. Pappas, and A. Katsaggelos, "Resource allocation for downlink multiuser video transmission over wireless lossy networks," *IEEE Trans. Image Process.*, vol. 17, no. 9, pp. 1663–1671, Sept 2008.
37. H. Luo, S. Ci, D. Wu, and H. Tang, "Cross-layer design for real-time video transmission in cognitive wireless networks," *IEEE INFOCOM 2010 Workshop on Cognitive Wireless Communications and Networking*, San Diego, Mar 2010.
38. H. Luo, S. Ci, D. Wu, and H. Tang "Quality-driven cross-layer optimized video delivery over lte," *IEEE Commun.*, vol. 48, no. 2, pp. 102–109, 2010.