Multi-Teacher Knowledge Transfer for Optimal CRN Spectrum Handoff Control with Hybrid Priority Queueing

Yeqing Wu[1], Fei Hu[1]*, John D. Matyjas[2], Yingying Zhu[3], Sunil Kumar[4]
[1] Electrical and Computer Engineering, University of Alabama, Tuscaloosa, AL
[2] Air Force Research Laboratory, Rome, NY
[3] Electrical Engineering, University of California, Riverside, CA
[4] Electrical and Computer Engineering, San Diego State University, San Diego, CA
Email: {ywu40@crimson.ua.edu, *Corresponding author:fei@eng.ua.edu, john.matyjas@us.af.mil, and skumar@mail.sdsu.edu}

***Abstract*—In this work, we target an optimal spectrum handoff design in cognitive radio networks (CRNs). It has two innovative designs: (1) *Hybrid, rule-based priority queueing model for spectrum handoff*: To overcome the drawbacks of preemptive or non-preemptive resume priority queueing model used by the secondary users (SUs), we propose a hybrid queueing model *with discretion rule* to characterize the spectrum access priority among SUs. Such a hybrid queueing model is then used to calculate the channel waiting time for spectrum handoff decision. (2) *Multi-teacher knowledge transfer for intelligent spectrum handoff*: Unlike existing CRN cognition engine designs that focus on spectrum adaptation through SU self-learning (i.e., a SU learns how to adapt to the dynamical CRN radio conditions), we propose the concept of *multi-teacher knowledge transfer*, which allows multiple SUs that already have mature spectrum adaptation strategies to transfer their learning results to an inexperienced SU. We will solve *who* and *how* issues, that is, *who* should be the teachers and *how* multiple teachers can transfer the knowledge to a student SU. Our simulation results show that the proposed new designs can improve the spectrum handoff accuracy under complex CRN radio contexts.**

***Index Terms*—Spectrum Handoff, Cognitive Radio Networks, Multi-Teacher Learning, Hybrid Queueing.**

I. Introduction

Spectrum handoff is an important issue in cognitive radio networks (CRNs) [1]. Each time a primary user (PU) reoccupies the channel, the secondary user (SU) faces two options: either switch to a new channel or wait for that channel to become available again. Using the second option may be more efficient in some cases as the channel switching takes time and introduces the computational overhead. Two types of queueing models have been proposed in the literature to analyze the handoff delay, i.e., *preemptive* and *non-preemptive* resume priority models. A preemptive resume priority M/G/1 queueing model (PRP) was proposed in [2]–[4]. This queueing model can cause frequent spectrum handoffs due to the interruptions from other SUs. In the non-preemptive resume priority M/G/1 model (NPRP) [5], [6], a high-priority SU with stringent delay requirement cannot interrupt the service of a lower priority SU. Such a feature makes this model unsuitable for delay-sensitive traffic.

In this paper, we propose a *hybrid preemptive and non-preemptive resume priority (PRP/NPRP) M/G/1 queueing model with discretion rule*, to manage and characterize the spectrum usage behaviors of PUs and SUs. In our hybrid model, we use a discretion rule to determine whether a high-priority SU can preempt the service of a low-priority SU. A high-priority SU can be serviced immediately if the discretion rule is satisfied; Otherwise, it must wait in the queue for the completion of the service of the lower priority SUs. In this way, we can not only reduce the waiting time of the high-priority SUs, but also avoid frequent spectrum handoffs. However, the handoff decision should not be just based on the queueing delay. Other factors, such as packet error rate (PER), are also important. Therefore, in this work, we propose to perform spectrum handoff based on a comprehensive channel quality metric, including channel waiting time, handoff delay, PER, packet drop rate (PDR), etc..

Existing schemes mostly select a channel for handoff in a *myopic* manner by maximizing the ‘rewards’ in the current operation round [7], [8]. Our previous results [4] showed that such a myopic handoff strategy may not achieve the optimal cumulative rewards in all the rounds of a communication session. We also showed that the reinforcement learning (RL) can achieve the optimal total reward in the long term. However, the convergence of the RL-based spectrum handoff is usually slow due to the varying and complex communication conditions in CRNs. In [9], we introduced the concept of apprenticeship learning (AL) to enhance the SU’s learning process. It enables an SU (called a student or apprentice) to learn the most suitable transmission policies from the experienced users (called teachers or experts).

The main contributions of this paper are two-fold:

1) A *Hybrid PRP/NPRP M/G/1 queueing model with discretion rule for spectrum handoff* that supports differentiated services. This hybrid model significantly improves the non-preemptive queueing model presented in our previous work [4], [9]. In this hybrid queueing model, the high-priority SUs are allowed to preempt the service of the low-priority SUs based on a discretion rule. This model can avoid frequent

spectrum handoffs, while reducing the wait time for the high-priority SUs. Although our hybrid queueing model more accurately reflects the PU/SU traffic transmission relationship, it is challenging to mathematically analyze the queueing delay. We introduce the concept of *delay cycle*, and utilize the Laplace transform of time functions to develop a closed-form handoff delay framework.

2) *Multi-teacher apprenticeship learning for QoE-driven spectrum handoff*. In RL-based spectrum handoff, when an SU enters the CRN, it could take a long time to make an optimal spectrum decision due to the difficult choice of initial RL parameters and utility function under the complex radio conditions. The proposed multi-teacher apprenticeship learning (MAL) model speeds up the learning process by enabling a newly-joined (apprentice) SU to learn from multiple neighboring expert SUs. We use the manifold learning to search the most suitable 'teachers' (or experts) with statistically similar features to the 'apprentice' SU.

A preliminary version of this scheme was described in [10]. However, the current paper not only provides the complete handoff delay calculation model for three different types of SUs, but also introduces the multi-teacher learning model to achieve an intelligent handoff.

The rest of this paper is organized as follows: The related work is briefly summarized in Section II. In Section III, we describe the proposed hybrid queueing model with discretion rule. The handoff delay analysis is presented in Section IV. Section V first describes the RL-based, QoE-driven handoff scheme, followed by the MAL-based scheme. Section VI presents the simulation results and performance analysis, followed by the conclusions in Section VII.

II. Related Work

A voluntary handoff approach was proposed in [7] to minimize the SU communication disruption time. In [8], a proactive handoff scheme was designed based on the discrete-time Markov chain. The SU can use the past channel usage statistics to decide the channel switching time. An optimal target channel sequence selection mechanism was discussed in [11] for smooth handoff.

Besides the above proactive handoff designs, some studies have studied the reactive handoff schemes. For example, Zhang et al. proposed an opportunistic spectrum handoff scheme in [1]. A fuzzy-based handoff was described in [12]; It makes sure that the aggregate interference among all SUs to the PUs does not go beyond a threshold.

Most of the existing handoff schemes assume that all SUs have the same priority. Thus their designs cannot fully support the multimedia quality of service (QoS) and quality of experience (QoE) objectives. Very few schemes have considered the SU priority issues. In [13], a prioritized SU traffic queueing model was discussed. However, it allows the higher priority SUs to preempt the lower priority SUs. This could deteriorate the QoS/QoE performance due to frequent handoffs, especially when the network traffic is high. Further, most of the existing handoff schemes use *myopic* handoff models. A RL-based handoff scheme was proposed in [4] to achieve long-term optimization. However, it does not allow the high-priority SUs to preempt the low-priority SU.

To the best of our knowledge, a *multi-teacher* learning model to enhance the CRN handoff performance has not been used earlier in the literature. Our previous work [9] only considered a single-teacher learning model. Other works mostly use Q-leaning to search a global optimization point without considering the *knowledge transfer* between SUs.

III. Network Model

We consider a CRN with M independent channels. An SU's communication session could experience multiple interruptions by PUs or other SUs. A PRP/NPRP M/G/1 queue with discretion rule is used to model the spectrum usage behavior. In this model, a discretion rule, based on the elapsed service time of the SU, is used to determine if an SU, which is currently in service (i.e., using the channel), can be interrupted by a higher-priority SU. A higher-priority SU can be given channel access immediately by interrupting the service of a lower-priority SU if the preemptive discretion rule is satisfied. Otherwise, it must wait in the queue till the current SU finishes its transmission.

Each channel maintains a separate priority queue for every user class in order to avoid the head-of-line blocking [14]. Specifically, $Q_p^{(k)}$ is the primary queue for PUs at channel k, and $Q_j^{(k)}$ is the queue for SUs with priority j, $1 \leq j \leq N$, where $N-1$ is the number of SU priority classes. Class $j=1$ has the highest priority and $j=N$ has the lowest priority.

As shown in Fig. 1, two SUs with priority j are transmitting over channels k' and k. A PU can always interrupt the SU_j. However, when a higher-priority SU arrives (e.g., SU_2), it needs to check whether the preemptive discretion rule of SU_j is satisfied or not. If the rule is satisfied, SU_j is interrupted and needs to choose to either stay at k (and wait for the current channel to become available again) or switch to another channel. If the SU chooses to stay at k, as shown in the "No" branch following the "Switch" box in Fig. 1, it is pushed to the head of $Q_j^{(k)}$. If it switches to channel $k^{''}$ (as shown by the "Yes" branch after the "Switch"), it is pushed back to the tail of $Q_j^{(k^{''})}$. If the preemptive discretion rule is not satisfied, SU_j will continue its service without being interrupted and SU_2 waits in the queue.

IV. Spectrum Handoff Delay Analysis

A. Variables and Concepts

We use the *delay cycle* concept [15] and the Laplace transform to analyze the expected waiting time of SUs. The delay cycle of an SU consists of two parts: 1) *Initial delay*: the service time for the initiating user at the target channel; 2) *Delay busy period*: the time for servicing other users before the SU under consideration.

Based on the impact of PUs and other SUs on the waiting time of SU_j, we group all the related users into three classes: type-α, type-j and type-β. Type-α users include all PUs and SUs with a higher priority over j. Type-β users include SUs with a lower priority $j+1$ through N. Type-j users include

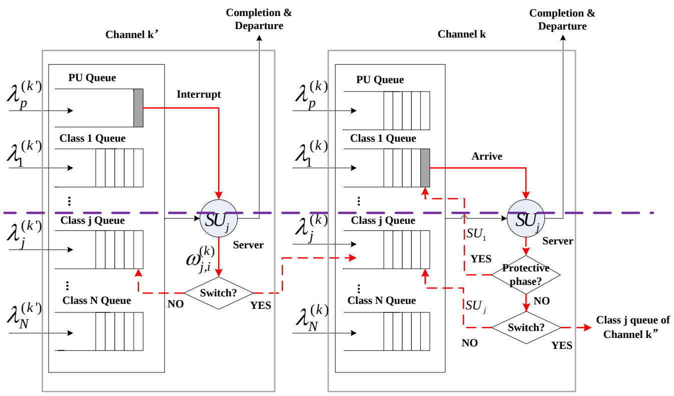


Fig. 1. The hybrid PRP/NPRP M/G/1 queueing model (adapted from [4], [9]

all SUs with a priority j. A newly-arrived SU_j needs to wait in the queue if there exist any user whose service is in the protective, non-preemptive duration. Otherwise, SU_j can be served immediately if the channel is in an idle state and there is no PUs or SUs with priority higher than or equal to j, or the type-β SU in service is in its preemptive phase.

Accordingly, in the analysis of waiting time of SU_j, it involves three types of delay cycles as in [15]: Type-α delay cycle (initiated by a type-α PU or SU with a priority higher than j), type-j delay cycle (initiated by a type-j SU in service) and type-β delay cycle (initiated by a type-β SU in its protective non-preemptive phase). All delay cycles will end when there is no type α and type j user at the channel.

We assume that the user arrival process follows a Poisson distribution as in [2]. The SU connection with priority j that is experiencing its i^{th} interruption is denoted as type-(j,i) connection, where $i \geq 0$. The main variables used in the analysis are listed in Table I. The relationship among these variables is shown in Fig. 2 through an example of SU_j operating at channel k, during which I interruptions occur. From the figure, we can see that the effective service time $S_j^{(k)}$ is the sum of all interrupted service times $S_{j,i}^{(k)}$ and the final successful service time $S_{j,f}^{(k)}$ at channel k. Note that all symbols use the statistical expectation values.

B. Discretion Rule

The discretion rule is developed based on the elapsed service time of the SU (i.e., $SU_{j'}$) in service (i.e., currently accessing the channel). If the elapsed service time of $SU_{j'}$ is less than a predefined threshold $\tau_{j'}$, the preemptive discretion rule is satisfied. Then, a higher-priority SU (e.g., SU_j $(j < j')$) can interrupt its service and be served immediately. Otherwise, SU_j must wait in the queue for $SU_{j'}$ to complete its service.

We use two variables S_{A_j} (and S_{B_j}) to denote the preemptive period (and non-preemptive period) of SU_j by a higher-priority SU. Then the service time of SU_j is given by,

$$S_j = S_{A_j} + S_{B_j}, \tag{1}$$

where

$$S_{A_j} = min\{S_j, \tau_j\}.$$
$$S_{B_j} = max\{0, S_j - \tau_j\}. \tag{2}$$

TABLE I
MAIN PARAMETERS USED IN QUEUEING ANALYSIS

Symbol	Meaning
$\lambda_j^{(k)}$	Arrival rate of a user with priority j at channel k.
$\mu_j^{(k)}$	Service rate of a user with priority j at channel k.
$E[X_j^{(k)}]$	First moment of service time for a user with priority j at channel k.
$E[N_j^{(k)}]$	Average number of priority j users in queue $Q_j^{(k)}$ at channel k.
$\rho_j^{(k)}$	Normalized load of channel k due to SU_j at the channel, where $\rho_j^{(k)} = \lambda_j^{(k)} E[X_j^{(k)}]$.
$\omega_{j,i}^{(k)}$	Arrival rate of a type-(j,i) SU connection at channel k. $\omega_{j,0}^{(k)} = \lambda_j^{(k)}$.
$\rho_{j,i}^{(k)}$	Normalized load of channel k due to a type-(j,i) SU at channel k.
$W_j^{(k)}$	Waiting time of a SU_j connection before it is serviced at channel k.
$D_j^{(k)}$	Breakdown time of SU_j elapsed before it is served at channel k again.
$R_j^{(k)}$	Residence time elapsed from the time SU_j starts its service until it competes its service.
$C_j^{(k)}$	Completion time elapsed from the time SU_j starts its service until channel k becomes free to serve the next SU_j.
$T_j^{(k)}$	Response time of SU_j actually spends at channel k.
$S_j^{(k)}$	Effective service time of SU_j at channel k.
$S_{j,i}^{(k)}$	Effective service time of SU_j after the $(i-1)^{th}$ and before the i^{th} interruption at channel k. $E[S_{j,1}^{(k)}] = E[X_j^{(k)}]$
$S_{j,f}^{(k)}$	Final successful service time of SU_j at channel k.
$F_x(.)$	Probability distribution function of a random variable X.
$X^*(s)$	Laplace transform associated with a random variable x.
$\Lambda_j^{(k)} = \sum_{l=1}^{j} \sum_{i=1}^{I_{max}} \omega_{l,i}^{(k)}$	Sum of arrival rates of type-α and type-j users at channel k. I_{max} is the maximum number of interruptions.
$\gamma_j^{(k)} = \sum_{i=1}^{I_{max}} \omega_{j,i}^{(k)}$	Sum of arrival rates of type-j users at channel k.
$\sigma_j^{(k)} = \sum_{l=1}^{j} \sum_{i=1}^{I_{max}} \rho_{l,i}^{(k)}$	Sum of normalized load of type-α and type-j users at channel k.

Note that PUs have the highest priority (priority 1) and are non-preemptive, thus we have $S_{A1} = 0$ and $S_{B1} = S_1$.

C. Residence Time and Completion Time

Let $B_j^{(k)}$ denote a busy time period elapsed from the SU_j arriving at channel k until the channel becomes empty for higher priority SUs. It can also be considered as a busy period during which a SU_j arrives with the service time $C_j^{(k)}$ at channel k. According to [15], we have

$$B_j^{*(k)}(s) = C_j^{*(k)}(s + \gamma_j^{(k)} - \gamma_j^{(k)} B_j^{*(k)}(s)). \tag{3}$$

The length of the breakdown time $D_j^{(k)}$ initiated by a SU_{j-1} is equivalent to its busy period. Also, the breakdown time initiated by an SU with a higher priority than $(j-1)$ or a

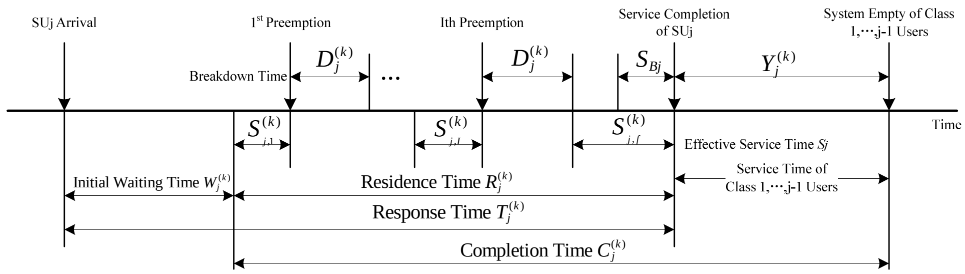


Fig. 2. Relationship among the random variables during service of SU_j. The superscript k indicates the variable is associated with channel k.

PU may be regarded as a delay cycle with the initial delay of $D_{j-1}^{(k)}$, during which each accumulated SU_{j-1} generates a sub-busy period of $B_{j-1}^{(k)}$. These two types of breakdown times occur with the probability of $\gamma_{j-1}^{(k)}/\Lambda_{j-1}^{(k)}$ and $(\Lambda_{j-1}^{(k)} - \gamma_{j-1}^{(k)})/\Lambda_{j-1}^{(k)}$, respectively. Hence, we can represent $D_j^{*(k)}(s)$ in a recursive form as [15]

$$D_j^{*(k)}(s) = \frac{\gamma_{j-1}^{(k)}}{\Lambda_{j-1}^{(k)}} B_{j-1}^{*(k)}(s) + \frac{\Lambda_{j-1}^{(k)} - \gamma_{j-1}^{(k)}}{\Lambda_{j-1}^{(k)}} D_{j-1}^{*(k)}(s + \gamma_{j-1}^{(k)} - \gamma_{j-1}^{(k)} B_{j-1}^{*(k)}(s)). \tag{4}$$

where $j \geq 2$ and $D_1^*(s) = 1$. Only class 1 through class $j-1$ users can preempt the service of SU_j. Therefore, according to [15], we can get the first two moments of $D_j^{(k)}$ in (5) and (6) by taking differentiation at $s = 0$ on (4).

A simpler expression can be found in [16] by using inductive analysis as

$$E[D_j^{(k)}] = \frac{\sigma_{j-1}^{(k)}}{\Lambda_{j-1}^{(k)}(1 - \sigma_{j-1}^{(k)})}, \text{where} \quad E[D_1] = 0. \tag{7}$$

Assuming SU_j encounters I interruptions before completing its service, it receives service $I+1$ times in total from the network. It's residence time $R_j^{(k)}$ can be represented by the sum of I breakdowns plus interrupted service times and one final successful service time:

$$R_j^{(k)} = \sum_{i=1}^{I_{max}} (D_j^{(k)} + S_{j,i}^{(k)}) + S_{j,f}^{(k)}. \tag{8}$$

When the residence time of the SU_j ends, there may be SUs with higher priority accumulated during the non-preemptive interval S_{B_j} of the final successful service time $S_{j,f}^{(k)}$. Each of those should be served before the SU_j gets the channel access again. Therefore, as shown in Fig. 2, the completion time $C_j^{(k)}$ consists of the residence time $R_j^{(k)}$ and a delayed busy period Y_j, which is generated by these higher-priority SUs that arrived during S_{B_j}.

The mean duration of the breakdown time is denoted as $E[D_j^{(k)}]$. Assuming the arrival rate follows Poisson process, we have

$$\begin{aligned} E[C_j^{(k)}] &= E[S_j^{(k)}] + \Lambda_{j-1}^{(k)} E[S_j^{(k)}] E[D_j^{(k)}] \\ &= (1 + \Lambda_{j-1}^{(k)} E[D_j^{(k)}]) E[S_j^{(k)}]. \end{aligned} \tag{9}$$

From (7) and (9), we have

$$E[C_{j-1}^{(k)}] = \frac{E[S_{j-1}^{(k)}]}{1 - \sigma_{j-2}^{(k)}}. \tag{10}$$

Substituting $E[D_{j-1}^{(k)}]$ and $E[C_{j-1}^{(k)}]$ into (6), we can obtain the simpler expression of the two moments of $E[D_j^{(k)}]$ as

$$E[(D_j^{(k)})^2] = \frac{1}{\Lambda_{j-1}^{(k)}(1 - \sigma_{j-1}^{(k)})^2} \sum_{l=1}^{j-1} \lambda_l^{(k)} \frac{(1 - \sigma_{l-1}^{(k)})^2}{1 - \sigma_l^{(k)}} E[C_l^2]. \tag{11}$$

After obtaining the mean duration of the completion time and breakdown time of class j, we can use them to analyze the handoff delay as follows.

D. Analysis of Expected Handoff Delay

When an SU is interrupted, it can either stay at the current channel or switch to another available channel. We call the first case as the *staying* case and the second one as the *switching* case. To choose the optimum handoff behavior for the interrupted SU, the expected mean opinion score (MOS) of the target channel and the expected handoff delay of choosing each available channel need to be estimated. We now provide a mathematical model to analyze the expected handoff delay for different cases.

Let the handoff delay $E[\widetilde{W}_{j,i}'^{(k)}]$ be the time duration from the instant the i^{th} interruption occurs to the instant the interrupted transmission is resumed, assuming channel k is chosen for spectrum handoff. We have [17]:

$$E[\widetilde{W}_{j,i}'^{(k)}] = \begin{cases} E[W_j'^{(k)}], & \text{if } c_{i-1} = c_i = k \\ E[W_j^{(k)}] + t_s, & \text{if } (c_{i-1} = k') \neq (c_i = k) \end{cases}. \tag{12}$$

Here c_i denotes the target channel for spectrum handoff at i^{th} interruption. $E[W_j'^{(k)}]$ (or $E[W_j^{(k)}]$) denotes the average waiting time of the i^{th} interruption if the SU_j chooses to stay at the current channel (or switches to channel k). Since the switching time t_s is already known (depending on the channel switching hardware architecture), we now describe how to calculate $E[W_j'^{(k)}]$ and $E[W_j^{(k)}]$. We use the busy period described before to analyze the waiting time of a SU_j when its service is interrupted. A busy period can be considered as a sequence of delay cycles. Fig. 3 shows these three types of delay cycles involved in delay analysis.

$$E[D_j^{(k)}] = \frac{\Lambda_{j-2}^{(k)} E[D_{j-1}^{(k)}] + \gamma_{j-1}^{(k)} E[C_{j-1}^{(k)}]}{\Lambda_{j-1}^{(k)}(1-\gamma_{j-1}^{(k)} E[C_{j-1}^{(k)}])} \tag{5}$$

$$E[(D_j^{(k)})^2] = \frac{\Lambda_{j-2}^{(k)} E[(D_{j-1}^{(k)})^2](1-\gamma_{j-1}^{(k)} E[C_{j-1}^{(k)}]) + \gamma_{j-1}^{(k)} E[(C_{j-1}^{(k)})^2](1+\Lambda_{j-2}^{(k)} E[D_{j-1}^{(k)}])}{\Lambda_{j-1}^{(k)}(1-\gamma_{j-1}^{(k)} E[C_{j-1}^{(k)}])^3} \tag{6}$$

a) Switching case: In this case, the interrupted SU connection of type j chooses to switch from channel $c_{i-1} = k'$ to another channel (e.g., $c_i = k$). After switching, the interrupted class j SU may find that the channel c_i is in one of the four states: 0, α, j, β, corresponding to virtually idle, type-α delay cycle, type-j delay cycle, type-β delay cycle, in that order. Let $\pi_l^{(k)}$ denote the steady-state probability that the channel is in state l, where $l \in \{0, \alpha, j, \beta\}$. $\pi_l^{(k)}$ can be obtained as [15]

$$\begin{aligned}
&\pi_0^{(k)} = 1 - \rho^{(k)}, \\
&\pi_\alpha^{(k)} = \rho_\alpha^{(k)}(1-\rho^{(k)})/(1-\rho_\alpha^{(k)} - \rho_j^{(k)}), \\
&\pi_j^{(k)} = \rho_j^{(k)}(1-\rho^{(k)})/(1-\rho_\alpha^{(k)} - \rho_j^{(k)}), \\
&\pi_\beta^{(k)} = \rho_\beta^{(k)}/(1-\rho_\alpha^{(k)} - \rho_j^{(k)}).
\end{aligned} \tag{13}$$

where

$$\rho_\alpha^{(k)} = \sum_{l=1}^{j-1} \sum_{i=0}^{I_{max}} \rho_{\alpha_l,i}^{(k)},$$

$$\rho_\beta^{(k)} = \sum_{l=j+1}^{N} \rho_{\beta_l}^{(k)} = \sum_{l=j+1}^{N} \lambda_{\beta_l}^{(k)} E[S_{B_l}],$$

and $\rho^{(k)} = \rho_\alpha^{(k)} + \rho_j^{(k)} + \rho_\beta^{(k)}$.

The Laplace transform of the conditional waiting time of type-j SU can be represented as [15]

$$\begin{aligned}
W_j^{*(k)}(s) &= \pi_0^{(k)} + \pi_\alpha^{(k)} W_{j|\alpha}^{*(k)}(s) + \pi_j^{(k)} W_{j|j}^{*(k)}(s) \\
&\quad + \sum_{l=j+1}^{N} \pi_{\beta_l}^{(k)} W_{j|\beta_l}^{*(k)}(s).
\end{aligned} \tag{14}$$

where $W_{j|l}^{*(k)}(s) = E[e^{-sW_j^{(k)}}|l]$, $l \in \alpha, j, \beta$ is the Laplace transform of the conditional waiting time that type-j SU needs to wait for when it arrives at channel k which is in state l.

As shown in Fig. 3, the delay cycle consists of the initial delay and the delay busy period. Let $\Psi_{j,i|l}^{(k)}$ denote the Laplace transform of the initial delay of type-l delay cycle that the arrived SU_j encounters, $l \in \{\alpha, j, \beta\}$. With similar derivative process as in [15], we can obtain

$$W_{j|l}^{*(k)}(s) = \frac{(1-\Psi_{j|l}^{*(k)}(s))}{E[T_{j|l}^{(k)}](s - \gamma_j^{(k)} + \gamma_j^{(k)} C_j^{*(k)}(s))}. \tag{15}$$

where $E[T_{j|l}^{(k)}]$ represents the mean duration of type-l delay cycle. Meanwhile, as shown in the Fig. 3, we can represent the Laplace transform of the initial delay of type-α and type-j delay cycle as

$$\begin{aligned}
\Psi_{j|\alpha}^{*(k)}(s) &= D_j^{*(k)}(s), \\
\Psi_{j|j}^{*(k)}(s) &= C_j^{*(k)}(s).
\end{aligned} \tag{16}$$

Moreover, we use $\gamma_l^{*(k)}$ to denote the rate at which type-l delay cycles are encountered by SU_j. Type-α or type-j delay cycle commences only when type α or type j user enters the channel in its virtually idle state, thus we can obtain $\lambda_\alpha^{*(k)} = \Lambda_{j-1}^{(k)} \pi_0^{(k)}$ and $\lambda_j^{*(k)} = \gamma_j^{(k)} \pi_0^{(k)}$. Meanwhile, we have $\pi_\alpha^{(k)} = \lambda_\alpha^{*(k)} E[T_{j|\lambda}^{(k)}]$ and $\pi_j^{(k)} = \lambda_j^{*(k)} E[T_{j|j}^{(k)}]$. Substituting these variables and (15), (16) into (14), we can get $W_j^{*(k)}(s)$ as shown in (17).

In this paper, we compare the elapsed service time of SU_l, $l \in \{j+1, ..., N\}$, with the predefined threshold τ_l for our preemptive discretion rule. If the elapsed service time is less than τ_l, a high-priority user can interrupt the service of class l user. Let $F_l(t) = Pr(S_l \leq t)$ denote the cumulative distribution function (CDF) of the service time of type-l user. When the type l user cannot be preempted by the higher-priority users, it initiates a type-β delay cycle. Thus, we have $\lambda_{\beta_l}^{*(k)} = \lambda_l^{(k)}(1 - F_l(\tau_l))$.

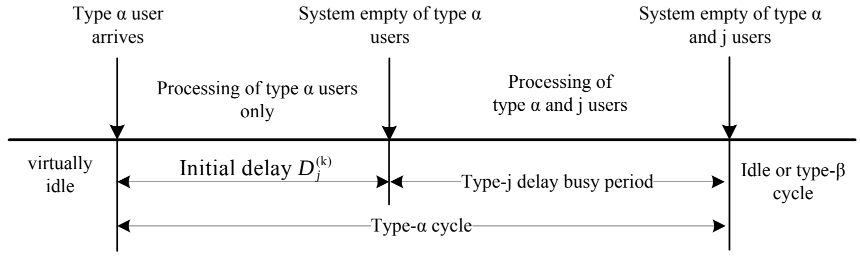
Moreover, in a type-β delay cycle, other SUs cannot interrupt the service of current low-priority SU. However, a PU can still interrupt the service of the SU. A type-β delay cycle can be considered as completed only when all type-α and type-j users arrive during the protective nonpreemptive duration of the serving SU. The busy period of type-α and type-j users which arrive during the protective nonpreemptive region of SU_j, can be denoted as the breakdown time $D_j^{(k)}$. Thus, we obtain

$$\Psi_{j|\beta_l}^{*(k)}(s) = \Phi_{S_{B_l}}(s + \Lambda_{j-1}^{(k)} - \Lambda_{j-1}^{(k)} D_j^{*(k)}(s)). \tag{18}$$

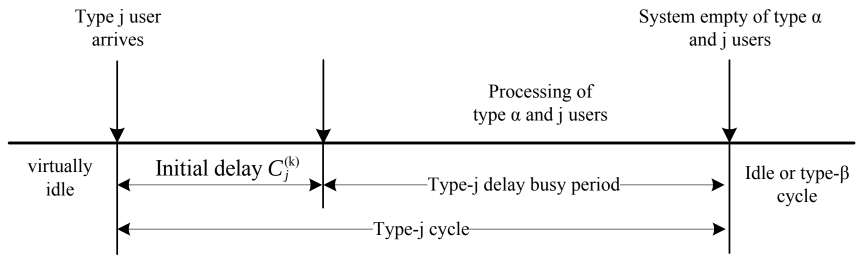
Substituting (18) into (17), we can obtain the final Laplace transform associated with the waiting time of class j user as shown in (19).

By taking the differentiation on (19) at $s = 0$, we get the expected waiting time of SU_j at channel k as in (20).

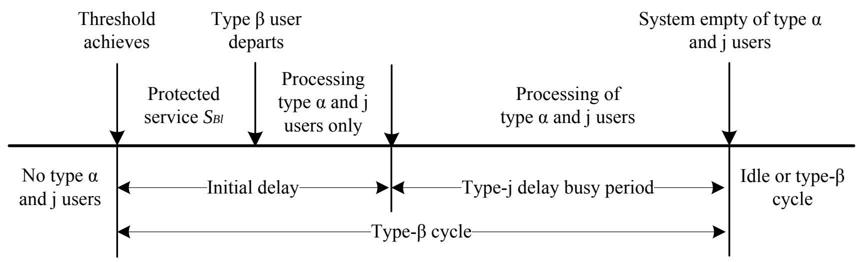
b) staying case: In this case, class k SU chooses to stay at the current operating channel, e.g. $c_i = c_{i-1} = k$. Fig. 1 already shows that after being interrupted it will be pushed to the head of the queue $Q_j^{(k)}$. It must wait until the completion of the service of all type-α users in the queue or the newly arrived type-α users. Therefore, we can consider the waiting time of SU_j in *staying* case as a type-α delay cycle, as shown in Fig. 4. The interrupted SU_j will be served when the system has served all class α users. Thus, we can obtain the Laplace



(a) Type-α delay cycle



(b) Type-j delay cycle



(c) Type-β delay cycle

Fig. 3. Three types of delay cycles for the channel switching case.

$$W_j^{*(k)}(s) = \frac{(1-\rho_\alpha^{(k)}-\rho_j^{(k)}-\sum_{l=j+1}^{N}\lambda_{\beta_l}^{(k)}E[S_{B_l}])(s+\Lambda_{j-1}^{(k)}-\Lambda_{j-1}^{(k)}D_j^{*(k)}(s))+\sum_{l=j+1}^{N}\lambda_{\beta_l}^{*(k)}(1-\Psi_{\beta_l}^{*(k)}(s))}{s+\gamma_j^{(k)}-\gamma_j^{(k)}C_j^{*(k)}(s)} \quad (17)$$

$$\begin{aligned} W_j^{*(k)}(s) = & \frac{(1-\rho_\alpha^{(k)}-\rho_j^{(k)}-\sum_{l=j+1}^{N}\lambda_{\beta_l}^{(k)}E[S_{B_l}])(s+\Lambda_{j-1}^{(k)}-\gamma_{j-1}^{(k)}D_j^{*(k)}(s))}{s+\gamma_j^{(k)}-\gamma_j^{(k)}C_j^{*(k)}(s)} \\ & + \frac{\sum_{l=j+1}^{N}\lambda_{\beta_l}^{*(k)}(1-\Phi_{S_{B_l}}(s+\Lambda_{j-1}^{(k)}-\Lambda_{j-1}^{(k)}D_j^{*(k)}(s)))}{s+\gamma_j^{(k)}-\gamma_j^{(k)}C_j^{*(k)}(s)} \end{aligned} \quad (19)$$

$$E[W_j^{(k)}] = \frac{(1-\rho_\alpha^{(k)}-\rho_j^{(k)})(1-\rho_\alpha^{(k)})^2\Lambda_{j-1}^{(k)}E[(D_j^{(k)})^2]+(1-\rho_\alpha^{(k)})^2\gamma_j^{(k)}E[(C_j^{(k)})^2]+\sum_{l=j+1}^{N}\lambda_{\beta_l}^{*(k)}E[S_{B\beta_l}^2]}{2(1-\rho_\alpha^{(k)})(1-\rho_\alpha^{(k)}-\rho_j^{(k)})} \quad (20)$$

transform of the waiting time of SU_j in *staying* case as

$$\begin{aligned} W_j^{'*(k)}(s) &= W_{j-1|\alpha}^{*(k)}(s) \\ &= \frac{(1-\Psi_{j-1|\alpha}^{*(k)}(s))}{E[T_{j-1|\alpha}^{(k)}](s-\gamma_{j-1}^{(k)}+\gamma_{j-1}^{(k)}C_{j-1}^{*(k)}(s))} \\ &= \frac{(1-\gamma_{j-1}^{(k)}E[C_{j-1}^{(k)}])(1-D_{j-1|\alpha}^{*(k)}(s))}{E[D_{j-1}^{(k)}](s-\gamma_{j-1}^{(k)}+\gamma_{j-1}^{(k)}C_{j-1}^{*(k)}(s))}. \end{aligned} \quad (21)$$

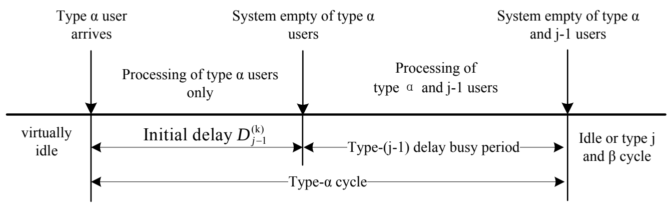


Fig. 4. The delay cycle when SU stays at the current channel.

E. Analysis of Expected Delivery Time

The expected delivery time of an SU connection, which experiences n interruptions during transmission, consists of the expected delays caused by interruptions and its service time. Since the service time of SU connections is assumed to be known, we only need to estimate its expected delay.

When $i > I_{max}$, we drop the packet. This results in $\widetilde{W}_{j,i}^{\prime(k)} = 0$. The probability that the type-(j,i) SU connection will be interrupted is $P_{j,i}^{(k)} = \lambda_p^{(k)} E[S_{j,i+1}^{(k)}]$ [5]. Thus, the expected delay of an SU connection with priority j can be derived as

$$E[Delay_j] = \sum_{i=0}^{I_{max}} i P_{j,i}^{(k)} E[\widetilde{W}_{j,i}^{\prime(k)}]. \quad (22)$$

where $E[\widetilde{W}_{j,i}^{\prime(k)}]$ is the handoff delay as described in (12).

V. MULTI-TEACHER KNOWLEDGE TRANSFER FOR INTELLIGENT SPECTRUM HANDOFF

A. QoE-Driven, Reinforcement Learning (RL) based Spectrum Handoff

Before describing our multi-teacher apprenticeship learning (MAL) model, we discuss the following two aspects which serve as the basis of the MAL-based handoff control: (1) the QoE-driven handoff strategy, which is more important to multimedia transmission than delay-based handoff (i.e., only use queueing delay to determine handoff). (2) The RL-based handoff, which will be extended to MAL-based handoff. The RL-based handoff is also a necessary step for an SU to become a 'teacher', since any new SU that cannot find a suitable teacher should learn to adapt to the complex CRN environment by itself. Our previous work [4], [9] has covered those aspects. We briefly summarize them below.

We denote the packet error rate (PER) of channel k for SU_j as $PER_j^{(k)}$. Let $Delay_{j,i}$ be the delay of the SU_j due to the first $(i-1)$ interruptions. The SU_j packet will be dropped when its delay exceeds the delay deadline d_j. Let $PDR_j^{(k)}$ (packet dropping rate) be the probability of packet being dropped during the i^{th} interruption. It equals the probability of $E[D_{j,i}^k]$ being larger than $d_j - Delay_{j,i}$. Both PER and PDR have been derived in our previous work [4].

Let $TPER_j^{(k)}$ (total packet error rate) be the estimated total PER of channel k for the SU_j connection at its i^{th} interruption. Assuming PER and PDR are independent of each other, we have $TPER_j^{(k)} = PER_j^{(k)} + PDR_j^{(k)} - PER_j^{(k)} \cdot PDR_j^{(k)}$. Using the QoE model derived in [18], the expected MOS (mean opinion score) for an SU_j choosing channel k for its i^{th} interruption, $MOS_{j,i}^{(k)}$, can be represented as a function of the sender bitrate (SBR), frame rate (FR) and the $TPER_j^{(k)}$.

$$MOS_{j,i}^{(k)} = \frac{\tau_1 + \tau_2 FR + \tau_3 ln(SBR)}{1 + \tau_4 (TPER_j^{(k)}) + \tau_5 (TPER_j^{(k)})^2}. \quad (23)$$

The coefficients $\tau_1, \tau_2, \tau_3, \tau_4, \tau_5$ can be obtained by a linear regression analysis [18].

Next, we can use RL-based model, with MOS-based reward function, to define an optimal spectrum handoff strategy. The RL model could be based on our previous work [4], except that we have to consider the queueing delay differences for three types of SUs described in the previous section.

B. Multi-Teacher Knowledge Transfer for Intelligent Spectrum Handoff

Due to the complex and dynamic nature of the spectrum conditions in CRNs, the learning process of RL could be slow, especially in the startup phase or when an SU goes through a new radio environment. AL (Apprenticeship Learning) can be used to make the apprentice SU perform as well as the expert SU [19], [21]. Unlike our previous work that used a single-teacher model [9], here we use *multi-teacher apprenticeship learning* (MAL) framework, which allows an apprentice SU to learn from multiple teachers simultaneously. Choosing multiple teachers helps to avoid the biased knowledge of a particular teacher.

Multiple teachers could teach the apprentice SU from different aspects. For example, a SU may find a teacher with similar *link channel statistics* as itself to learn how to choose proper sending rate based on the channel statistics. It may find another teacher with similar QoS statistics to itself to learn how to choose a QoS-oriented routing path from it. To find a suitable teacher, we need to define the similarity index (i.e., feature distance) between any two SUs. As shown in Fig. 5, we consider a high-dimensional signal consisting of multiple types of network parameters such as: (1) channel statistics: This includes the channel quality factors such as BER, SNR, etc. (2) SU statistics: It includes node mobility, modulation mode, etc. (3) QoS statistics: This includes end-to-end delay, jitter, etc. If a SU has the biggest similarity index in any of those parameters, it can be selected as a teacher.

Such a similarity index should be based on the statistical distance between two manifold signals. Typically the Geodesic distance is used to measure the shortest length between any two manifold points (Fig. 5). However, such distance is difficult to calculate. Therefore, we use the Bregman divergence (BD) [20] to replace the Geodesic distance. The BD between two manifold points p and q, denoted as D(p, q), is associated

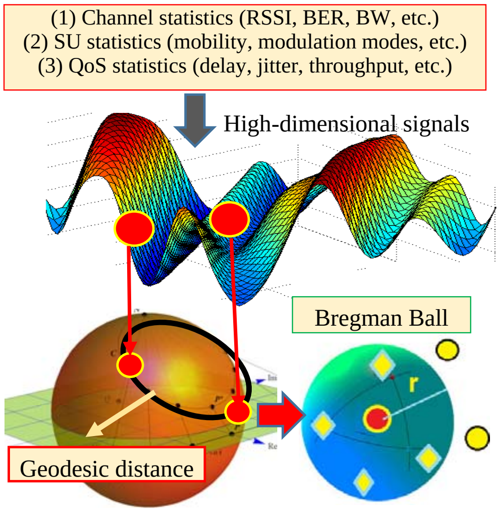


Fig. 5. Information Geometry based node-to-node teaching.

with a strict convex generator function $\Phi()$,

$$D_\Phi(p,q) = \Phi(p) - \Phi(q) - \langle \nabla\Phi(q), p-q\rangle. \tag{24}$$

where $\nabla\Phi = [\frac{\partial\Phi}{\partial x_1}, \frac{\partial\Phi}{\partial x_2}, \ldots]$ is the inner product operation. We then define the concept of Bregman ball, which has a center μ_k and radius R_k. For any manifold point X_t at time t, if it is inside this ball, it will have a strong statistical similarity between itself and the center μ_k. That is:

$$B(\mu_k, R_k) = \{X_t \in X : D_\Phi(X_t, mu_k) \le R_k\}. \tag{25}$$

In the multi-teacher model, we denote the apprentice SU as SU_a, and its state feature vector over states as $\phi(s)$. We assume that n teachers $(SU_{e_1}, \cdots, SU_{e_n})$ are selected. We assign the weight to the knowledge from the teachers by using the normalized similarity indices between the apprentice and the teacher in the feature space, $sim(SU_a, SU_{e_1})$, $\cdots$, $sim(SU_a, SU_{e_n})$, which are derived from their symmetric Bregman divergences. The similarity scores are normalized such that their sum is 1. The expected MOS of the apprentice SU, $MOS^{(k)}_{j,i}(\phi(s_i))$, is a combination of the experts' MOS, $MOS^{(k)}_{SU_{e_l},j,i}(\phi(s_i))$, and self-estimated MOS, $MOS^{(k)}_{SU_a,j,i}(\phi(s_i))$, with a decreasing effect from the teachers' MOS along time as

$$\begin{aligned} &MOS^{(k)}_{j,i}(\phi(s_i)) \\ &= \gamma^i [\textstyle\sum_{l=1}^n sim(SU_a, SU_{e_l}) \cdot MOS^{(k)}_{SU_{e_l},j,i}(\phi(s_i))] \\ &+ (1-\gamma^i) MOS^{(k)}_{SU_a,j,i}(\phi(s_i)). \end{aligned} \tag{26}$$

where $\gamma < 1$ is the multiplication ratio that decreases sequentially to indicate the weaker and weaker effect of experts' MOS expectation. The Q-values are updated for a given connection during its multiple interruptions:

$$\begin{aligned} Q(s,a) &= (1-\alpha)Q(s,a) + \alpha\{E(MOS_{j,i+1}) \\ &\quad + \gamma \max_{a' \in \mathcal{A}} Q(s', a')\}; \\ Q(s,a) &= \eta^i \sum_{l=1}^{n} Q_{SU_{e_l}}(s,a) + (1-\eta^i) Q_{SU_a}(s,a). \end{aligned} \tag{27}$$

Algorithm 1 The MAL-based spectrum handoff scheme.

Part One:
Input: Channel Statistics, Node Statistics, Application Statistics
Output: The best policy $\pi(s,a)$ of the SU
1) **if** SU is a new SU and one or multiple expert SUs can be found.
2) Perform MAL-based QoE-driven handoff.
3) **else**
4) Perform RL-based QoE-driven handoff.

Part Two:
1) Exchange info. among the SU and its neighbors.
2) Using manifold learning to find the expert SUs.
3) Transfer the MOS functions and Q-value functions from expert SUs.
4) Initialize $Q(s,a)$ with weighted Q values from the expert SUs.
5) **Repeat**
6) Part three with MOS and Q updates in (26) and (27).

Part Three:
1) Calculate channel PER.
2) Calculate the expected queueing waiting time $E[W^{(k)}_j]$ using (15).
3) Calculate the average delay $E[D^{(k)}_j]$.
4) Calculate the PDR.
5) Calculate the expected MOS using (23).
6) **if** the expected MOS is less than a predefined threshold
7) //The performance of SU_j is worse
8) Perform MAL-based QoE-driven handoff.
9) SU_j perform RL by itself.

where $\eta < 1$ is the multiplication ratio.

The system diagram of the proposed MAL-based handoff is shown in Fig. 6. If the expected MOS is less than the predefined threshold, it will perform MAL-based QoE-driven spectrum handoff to learn from its expert SUs. The handoff operations are described in Algorithm 1.

VI. Experimental Results

In this section, we evaluate the proposed spectrum handoff scheme through simulations. As suggested by IEEE 802.22 standard [22], 10 msec per time slot is used in our experiments. Similar to the three queueing models in [2], [4], [5], we assume that the service time of PUs and SUs follow the exponential distribution. According to the property of the exponential distribution, the service time $E[X] = \frac{1}{\mu}$ and the remaining transmission time of SUs follow the same exponential distribution after being interrupted by PUs and SUs [6].

A. QoS-aware Spectrum Handoff Scheme

In this section, we compare our proposed hybrid priority queueing model with three recent queueing models in [2], [4] and [5]. The number of channels is $M = 3$, and the number of priority classes of SU connections is $N = 4$, and class j has a higher priority than $j+1$. Here we consider the traffic delay, packet error rate, and user priority.

1) Effect of PU Traffic Load: It is expected that high traffic load of PUs will cause longer delay of SU connections. For high-priority users (SU1 and SU2), the average data delivery time gets longer with the increase of the normalized PU traffic load in Fig. 7(a). Our hybrid queueing model achieves much lower average delivery time for SU1 and SU2 than the queueing models in [2], [4], [5]. In Fig. 7(b), the overall average delivery time across SUs of all priorities achieved by our hybrid queueing model is much lower than the model in [2], and is comparable to the model in [4]. Although the proposed hybrid model has about 10% performance degradation in the average delivery time compared to the model in [5], the improvement of the delay performance for higher priority SUs is significant (about 2.5x improvement for the highest priority SU, i.e., SU1). Our scheme can guarantee the delay performance of higher priority SUs (such as the SUs that deliver real-time videos) at the cost of the slight performance degradation for lower priority SUs.

2) Effect of SU Traffic Load: In this set of experiments, we evaluate the delay performance of different queueing models for different SU traffic loads. In the experiment, as in [4], the SU connections are assumed to have the same service time, and the PU has the same settings: $\lambda_p^{(k)} = \lambda_p = 0.05(arrivals/slot)$ and $E[X_p^{(k)}] = E[X_p] = 6(slots/arrival)$.

The average data delivery time of high-priority users (SU1 and SU2) is shown in Fig. 8(a), and the result for all priorities of users is shown in Fig. 8(b). For both cases, the delivery time increases with the increase in the arrival rate, as more connections need to access the channel at the same time. With the same SU traffic load, our hybrid queueing model has lower delay for high-priority SUs than the other three queueing models. In Fig. 9(a) (for high-priority users SU1 and SU2) and (b) (for all users), the performance gain of our model over others increases with the service time of the SU connections. With longer service time, the high-priority SUs in [4], [5] need to wait for a longer time for the completion of the services of low-priority SUs, since in these models, a newly arrived SU cannot preempt the low-priority SUs.

B. QoE-driven Spectrum Handoff Scheme

In this section, we compare the performance of our QoE-driven spectrum handoff scheme with other two schemes - QoE-driven handoff based on the queueing model in [4], and QoS-based handoff adopted in [5], where the effect of handoff delay alone was considered when choosing a channel, without considering the effect of channel quality. The performance would suffer when choosing a channel with little delay if it has a high PDR. For fair comparison, we apply our proposed hybrid priority queueing model to the QoS-based handoff scheme, instead of using queueing model in [5].

The video sequences were encoded using H.264/AVC JM reference software [23] for a GOP length of 30 frames at 30 frames/sec. The FR (30 frames/sec) and SBR (200Kbps) are fixed. The number of channels is $M = 3$ and the number of priority classes is $N = 4$, where class j has higher priority than class $j + 1$.

As in [2], we assume that SUs may experience different channel conditions. Table II shows different PERs of each channel and the delay deadline of each SU. However, the total PER of each channel is almost the same for all SUs. Here we use the same type of traffic for all SUs, but assign different priorities (i.e., different delay deadlines) to them. The same channel status can be used for different video sequences.

TABLE II
SIMULATION PARAMETERS.

SU	CH1	CH2	CH3	d_j(sec)	Applications
SU1	16%	3%	11%	0.5	Low delay Video
SU2	2%	18%	10%	2	Live Stream
SU3	17%	9%	4%	4	Video on Demand
SU4	10%	16%	4%	NULL	File Download

Note: NULL denotes that the application has no delay deadline.

The total packet error rate (TPER) achieved by the different handoff schemes for the different normalized loads of PU is shown in Table III. Both QoE-driven spectrum handoff schemes have a lower TPER than the delay-driven scheme because the delay-driven scheme does not consider the effect of channel errors, whereas the QoE-driven schemes consider both the transmission delay and channel errors when choosing the

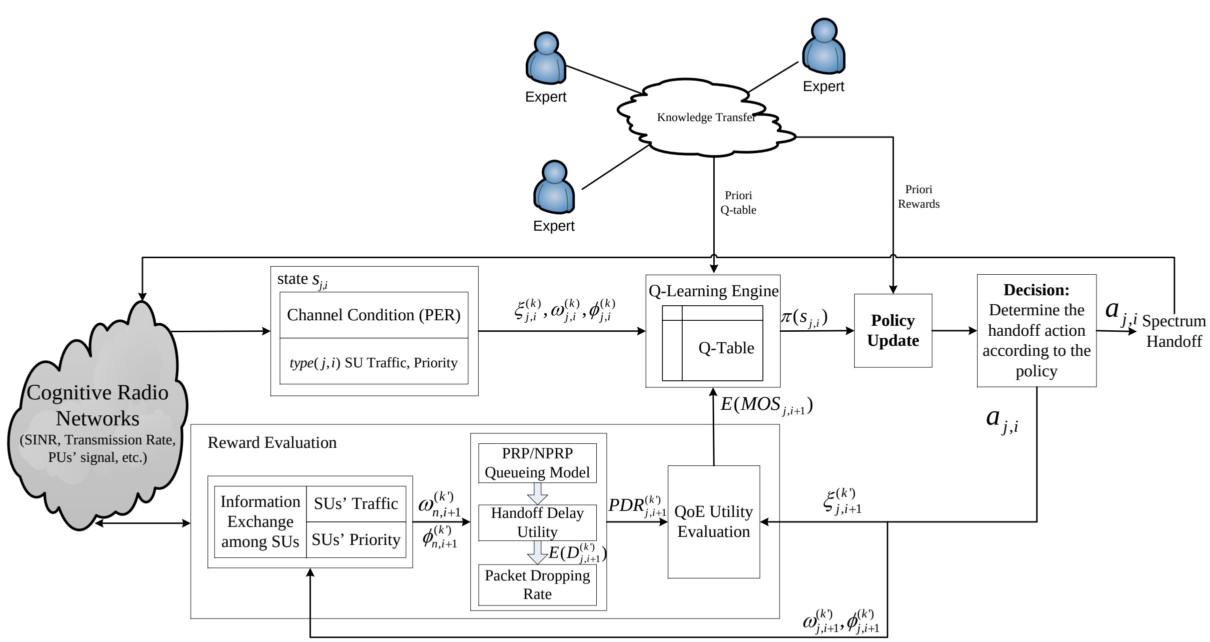
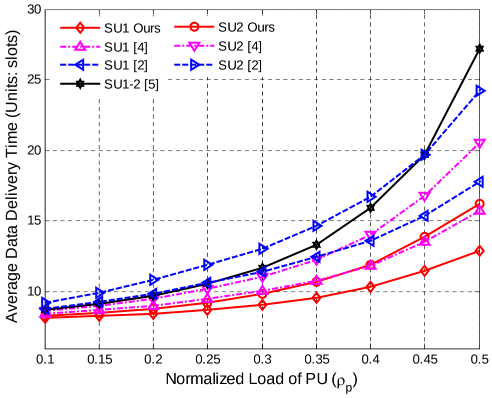
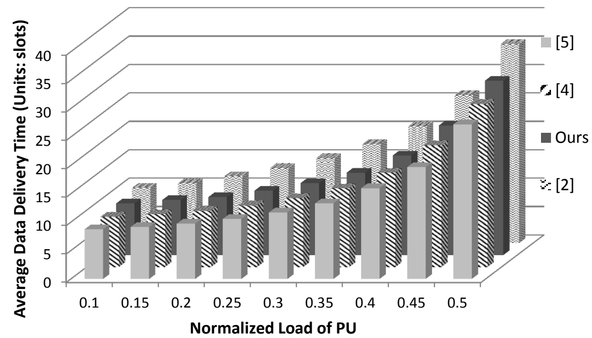


Fig. 6. Multi-teacher based spectrum handoff. (adapted from our single-teacher model in [9]

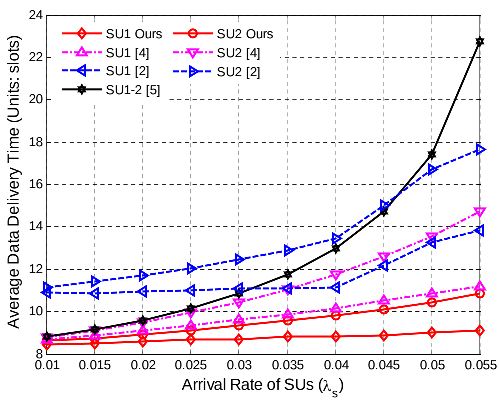


(a) SU1 and SU2

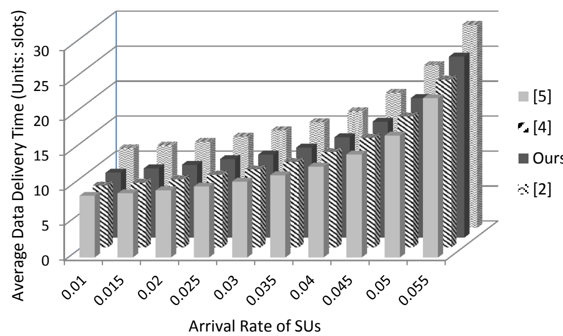


(b) Average delay of SU1 through SU4

Fig. 7. Effect of the normalized PU traffic load (ρ_p) on the average delivery time for E[X_p] = 10(slots/arrival), λ_s = 0.03(arrivals/slot), and E[X_s] = 8(slots/arrival). For clarity, we show only the average data delivery time of priority 1 and 2 SU for each handoff scheme in (a).

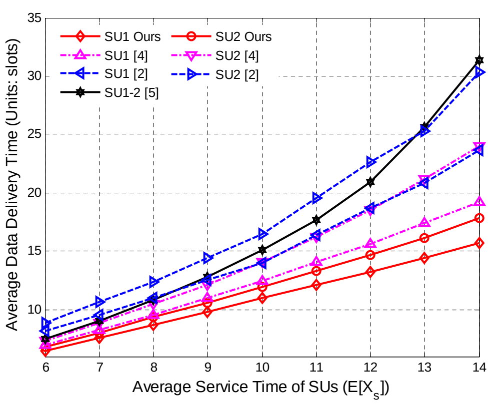


(a) SU1 and SU2

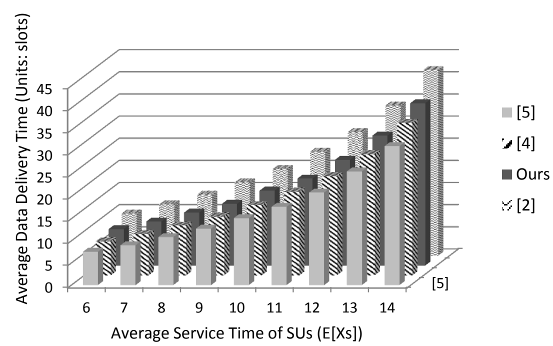


(b) Average delay of SU1 through SU4

Fig. 8. The effect of SU arrival rate on the average delivery time for λ_p=0.05(arrivals/slot), E[X_p]=6(slots/arrival), and E[X_s] = 8(slots/arrival). For clarity, we show only the average data delivery time of priority 1 and 2 SUs for each handoff scheme in (a).



(a) SU1 and SU2



(b) Average delay of SU1 through SU4

Fig. 9. The effect of SU service time on the average delivery time for λ_p=0.05(arrivals/slot), E[X_p]=6(slots/arrival), and λ_s=0.03(arrivals/slot). For clarity, we show only the average data delivery time averaged across all users and those of priority 1 and 2 SUs for each handoff scheme in (a).

target channel for handoff. Further, the QoE-driven spectrum handoff based on our hybrid queueing model has lower TPER than the QoE-driven handoff based on the queueing model in [4], especially for higher priority SUs under heavy traffic load. This is because the queueing model in [4] does not allow the higher priority SUs to interrupt the lower priority SU. As a result, the increased wait time for higher priority (i.e., lower delay deadline) SU may cause its packets to be dropped when its delay exceeds the deadline for a heavy traffic load. On the other hand, our proposed hybrid queueing model allows a higher-priority SU to preempt the service of a lower-priority SU if the discretion rule is satisfied.

The corresponding PSNR result for Foreman video sequence is shown in Fig. 10. For conciseness, only the PSNR results of high priority SU1 and SU2 connections are shown in the figure. The QoE-driven handoff based on our queueing

TABLE III
COMPARISONS OF TPER FOR DIFFERENT NORMALIZED LOADS OF PU (ρ_p).

	ρ_p	0.1	0.2	0.3	0.4	0.5	0.6	0.7
SU1	**QoE Ours**	3.5%	3.6%	3.8%	4.0%	4.3%	4.7%	5.3%
	QoE [4]	3.6%	4.0%	4.6%	5.4%	6.3%	7.9%	9.2%
	Delay	8.7%	9.0%	9.5%	10.3%	11.5%	12.9%	14.4%
SU2	**QoE Ours**	2.5%	2.7%	3.0%	3.5%	4.4%	5.5%	7.0%
	QoE [4]	2.5%	3.0%	3.7%	4.8%	6.4%	8.3%	10.7%
	Delay	7.6%	8.0%	8.7%	10.1%	11.9%	13.7%	15.6%
SU3	**QoE Ours**	4.6%	4.8%	5.2%	6.1%	7.2%	8.8%	11.1%
	QoE [4]	4.7%	5.2%	6.1%	7.4%	8.6%	10.4%	13.3%
	Delay	8.7%	8.9%	9.4%	10.2%	11.3%	13.4%	16.4%
SU4	**QoE Ours**	4.0%	4.0%	4.0%	4.0%	4.0%	4.0%	4.0%
	QoE [4]	4.0%	4.0%	4.0%	4.0%	4.0%	4.0%	4.0%
	Delay	7.7%	7.7%	7.7%	7.7%	7.7%	7.7%	7.7%

model has obvious improvement in terms of video PSNR compared to the QoE-driven handof [4] and the delay-driven handoff schemes.

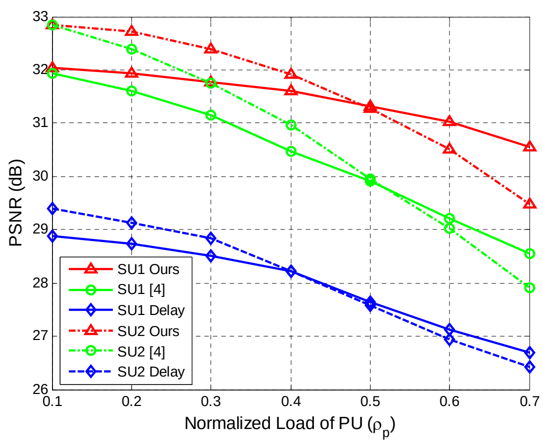


Fig. 10. Average PSNR (in SU1 and SU2) of Foreman video sequence (bit rate: 200kps) vs. the normalized load of PU (ρ_p) for λ_s = 0.05 (arrivals/slot), E[X_s] = 10 (slots/arrival), and E[X_p] = 20 (slots/arrival). The lossless PSNR of Foreman sequence is 36.81dB.

C. MAL-based Spectrum Handoff Scheme

In this section, we compare the performance of the proposed MAL-based spectrum handoff scheme against the general AL-based and the RL-based handoff schemes. The 480p $Whale\_show$ video sequence is used in the experiments, which is encoded at the bit rate of 1Mbps, by using the H.264/AVC JM reference software [23] for a GOP length of 30 frames at 30 frames/sec. The number of channels is $M = 3$, and the number of priority classes of SU connections is $N = 4$. The PER of a channel is picked randomly from 2% to 10%. The arrival rate and the service time of the PU and SU connections are set as $\lambda_p = 0.05(arrivals/slot)$, $E[X_p] = 6(slots/arrival)$, and $\lambda_s = 0.05(arrivals/slot)$, $E[X_s] = 8(slots/arrival)$, respectively. The discount rate (γ) of RL-based handoff scheme is set to 0.6 and the temperature υ in the softmax policy is set to 0.3.

1) Effect of the AL-enhanced QoE-driven Handoff: Figure 11 shows the expected MOS achieved by the AL- and RL-based handoff schemes. The AL-enhanced handoff outperforms the RL-based handoff, especially in the early stage of the transmission. Particularly, the AL-based handoff scheme converges much faster than the RL-based handoff scheme. For the time slot ¿ $3{*}10^4$ in Fig. 11(a), the performance of the RL-based scheme is almost the same as the AL-based scheme as both schemes converge to the stable state. This is because the newly-joined SU is almost static and thus does not experience much variation in channel conditions.

2) Effect of Multiple Teachers on the Performance of Handoff: Here we evaluate the expected MOS performance of the proposed MAL-based handoff scheme. From Fig. 11, we observe that the MAL based spectrum handoff scheme achieves better MOS performance than the AL and RL-based handoff scheme for the time-varying environment.

3) Effect of Dynamic Environment on the Performance of Handoff Scheme: Figure 11(b) shows the MOS performance when the newly-joined SU experiences time-varying channel conditions. The AL-based handoff is triggered when the predefined performance threshold is not satisfied. We observe see that our proposed MAL-based scheme outperforms the AL-based and the RL-based spectrum handoff schemes.

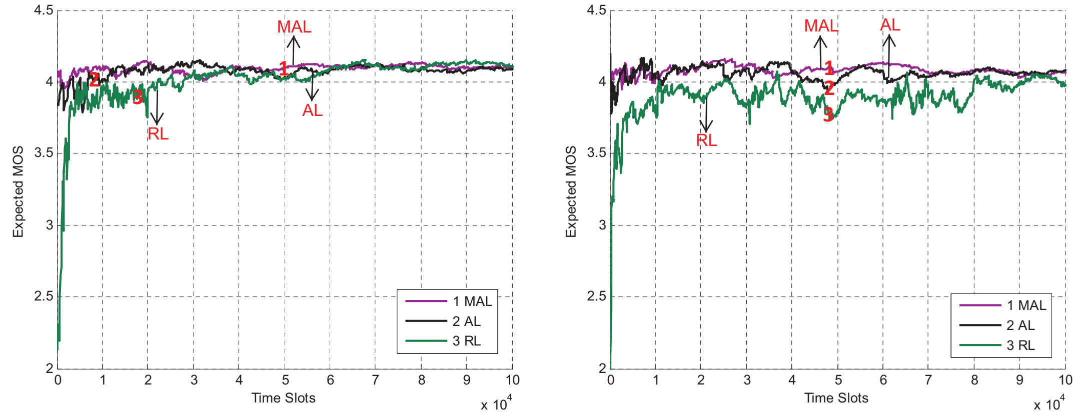
In Fig. 12, we compare the video transmission result of the MAL-based, the AL-based and the RL-based handoff schemes for SU_1 in dynamic radio environment. These video results match well with the result of Fig. 11. We have zoomed-in a part of the video frame to highlight the difference between the MAL-based and the RL-based schemes. The MAL scheme has lower PER and thus better image quality.

VII. CONCLUSIONS

In this paper, we have proposed a hybrid PRP/NPRP M/G/1 queueing model with discretion rule to manage spectrum handoff for multimedia applications in CRNs. The queueing model is designed to meet the prioritized transmission requirements while avoiding the excessive delay caused by frequent spectrum handoffs. Based on the queueing model, MAL is integrated into QoE-driven spectrum handoff scheme to allow a SU to learn from its neighbors, and performs spectrum handoff intelligently. Simulation results show that our proposed approaches improve the end-user satisfaction in terms of delivery time and video PSNR levels.

VIII. ACKNOWLEDGMENT OF SUPPORT AND DISCLAIMER

The authors and contractors acknowledge U.S. government's support in the publication of this paper. This material is based upon work funded by U.S. AFRL (Air Force Research Lab), under U.S. DoD Contract No. 8750-13-1-046. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of U.S. AFRL.



(a) Static Environment

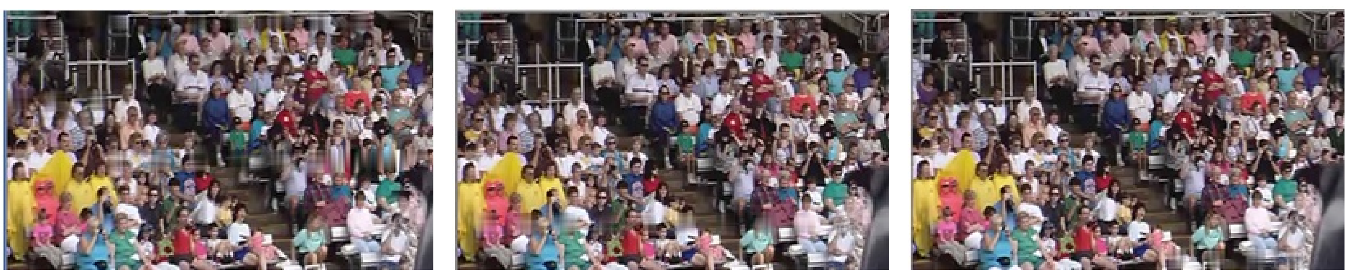
(b) Dynamic Environment

Fig. 11. Comparison of the RL-based, AL-based and our proposed MAL-based spectrum handoff schemes.

(a) RL

(b) AL

(c) MAL



(d) RL (Local Zoom)

(e) AL (Local Zoom)

(f) MAL (Local Zoom)

Fig. 12. Visual comparison (Frame 201) of different spectrum handoff schemes for the $Whale\_Show$ video sequence.

References

[1] Y. Zhang, "Spectrum Handoff in Cognitive Radio Networks: Opportunistic and Negotiated Situations," IEEE Int. Communications Conf., Dresden, 2009, pp. 1-6.

[2] H.-P. Shiang and M. van der Schaar, "Queuing-based Dynamic Channel Selection for Heterogeneous Multimedia Applications over Cognitive Radio Networks," IEEE Trans. Multimedia, vol. 10, Aug. 2008, pp. 896-909.

[3] L. Zhang, T. Song, M. Wu, J. Guo, D. Sun, and B. Gu, "Modeling for Spectrum Handoff based on Secondary Users with Different Priorities in Cognitive Radio Networks," Int. Conf. Wireless Commun. and Signal Proc. (WCSP), Oct. 2012, pp. 1-6.

[4] Y. Wu, F. Hu, S. Kumar, A. Talari, N. Rahnavard, and John D. Matyjas, "A Learning-based QoE-Driven Spectrum Handoff Scheme for Multimedia Transmissions over Cognitive Radio Networks," IEEE J. Selected Areas in Commun., vol. 32, no. 11, pp. 2134-2148, November 2014.

[5] L.-C. Wang, C.-W. Wang, and C.-J. Chang, "Modeling and Analysis for Spectrum Handoffs in Cognitive Radio Networks," IEEE Trans. Mobile Computing, vol. 11, Sept. 2012, pp. 1499-1513.

[6] C.-W. Wang and L.-C. Wang, "Modeling and Analysis for Proactive Decision Spectrum Handoff in Cognitive Radio Networks," Proc. IEEE ICC, Jun. 2009, pp. 1-6.

[7] S. Yoon and E. Ekici, "Voluntary Spectrum Handoff: A Novel Approach to Spectrum Management in CRNs," Proc. IEEE ICC, May 2010, pp. 1-5.

[8] Y. Song and J. Xie, "ProSpect: A Proactive Spectrum Handoff Framework for Cognitive Radio Ad Hoc Networks without Common Control Channel," IEEE Trans. Mobile Computing, vol. 11, Jul. 2012, pp. 1127-1139.

[9] Y. Wu, F. Hu, S. Kumar, and John D. Matyjas, "Apprenticeship Learning based Spectrum Decision in Multi-Channel Wireless Mesh Networks with Multi-Beam Antennas, IEEE Trans. Mobile Computing, (to appear).

[10] Y. Wu, F. Hu, S. Kumar, and E. Bentley, "Multimedia-Priority-Aware, Globally Optimal Spectrum Handoff based on Discreet Queueing in Cognitive Radio Networks," Int. Conf. Computer Communications and Networks (ICCCN), August 1-4, 2016, Waikoloa, Hawaii, USA.

[11] S. Zheng, X. Yang, S. Chen, and C. Lou, "Target Channel Sequence Selection Scheme for Proactive-Decision Spectrum Handoff," IEEE Commun. Letters, vol. 15, Dec. 2011, pp. 1332-1334.

[12] L. Giupponi and A. Perez-Neira, "Fuzzy-based Spectrum Handoff in Cognitive Radio Networks," Int. Conf. Cognitive Radio Oriented Wireless Netw. and Commun., May 2008, pp. 1-6.

[13] H.-P. Shiang and M. van der Schaar, "Online Learning in Autonomic Multi-hop Wireless Networks for Transmitting Mission-Critical Applications," IEEE J. Selected Areas Comm., vol. 28, Jun. 2010, pp. 728-741.

[14] J. Wang, H. Zhai, and Y. Fang, "Opportunistic Packet Scheduling and Media Access Control for Wireless LANs and Multi-hop Ad Hoc Networks," IEEE Wireless Commun. Net. Conf. (WCNC), vol. 2, Mar. 2004, pp. 1234-1239.

[15] R. W. Conway, W. L. Maxwell, and L. W. Miller, Theory of Scheduling. Addison-Wesley, NY, 1967.

[16] Y. Cho and C. Un, "Analysis of the M/G/1 Queue under a Combined Preemptive/Nonpreemptive Priority Discipline," IEEE Trans. Comm., vol. 41, Nov. 1993, pp. 132-141.

[17] A. Lertsinsrubtavee, N. Malouch, and S. Fdida, "Controlling Spectrum Handoff with a Delay Requirement in Cognitive Radio Networks," Proc. of ICCCN, Aug. 2012.

[18] A. Khan, L. Sun, E. Jammeh, and E. Ifeachor, "Quality of Experience-Driven Adaptation Scheme for Video Applications over Wireless Networks," IET Communications, vol. 4, Jul. 2010, pp. 1337-1347.

[19] P. Abbeel and A. Ng, "Apprenticeship Learning via Inverse Reinforcement Learning," Proc. 21st ACM Int. Conf. Machine Learning (ICML), 2004, pp. 1-8.

[20] A. Banerjee, S. Merugu, I. S. Dhillon, and J. Ghosh, "Clustering with Bregman Divergences," J. Machine Learning Research, vol. 6, Jan. 2005, pp. 1705-1749.

[21] A. Galindo-Serrano, L. Giupponi, P. Blasco, and M. Dohler, "Learning from Experts in Cognitive Radio Networks: The Docitive Paradigm," Proc. Int. Conf. CrownCom, Jun. 2010, pp. 1-6.

[22] IEEE Std 802.22-2011, Wireless Regional Area Network (WRAN) Specific Requirements Part 22: CognitiveWireless RAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Policies and Procedures for Operation in the TV Bands, Jul. 2011.

[23] JVT, H.264/AVC Reference Software JM14.2, ISO/IEC Std. [Online]. Available: http://iphome.hhi.de/suehring/tml/download/.

Yeqing Wu Dr. Wu received his Ph.D. degree from the Department of Electrical and Computer Engineering, University of Alabama, Tuscaloosa, USA, in 2015. He received his M.S. degree from the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China, in 2008. His research interests include cognitive radio networks, video codec, image processing, machine learning, rateless codes, and multimedia transmissions.

Fei Hu Dr. Hu is currently a professor in the Department of Electrical and Computer Engineering at the University of Alabama, Tuscaloosa, USA. He obtained his Ph.D. degrees at Tongji University (Shanghai, China) in Signal Processing (in 1999), and at Clarkson University (New York, USA) in Electrical and Computer Engineering (in 2002), respectively. He has published over 200 journal/conference papers, books, and book chapters. Dr. Hu's research has been supported by NSF, Department of Defense (DoD), Cisco, Sprint, etc. His research interests are in 3S (Security, Signals, and Sensors) and wireless networks.

John D. Matyjas Dr. Matyjas received his Ph.D. in Electrical Engineering from State University of New York at Buffalo in 2004. Currently, he is serving as the Technical Advisor for the Computing and Communications Division at the Air Force Research Laboratory (AFRL) in Rome, NY. His research interests include dynamic multiple-access communications and networking, spectrum mutability, statistical signal processing and optimization, and neural networks. He serves on the IEEE Transactions on Wireless Communications Editorial Advisory Board. Dr. Matyjas is the recipient of the 2012 IEEE R1 Technology Innovation Award, 2012 AFRL Harry Davis Award for "Excellence in Basic Research," and the 2010 IEEE Int'l Communications Conf. Best Paper Award. He is an IEEE Senior Member, chair of the IEEE Mohawk Valley Signal Processing Society, and member of Tau Beta Pi and Eta Kappa Nu.

Yingying Zhu Dr. Zhu received her Ph.D. degree from the Department of Electrical Engineering, University of California, Riverside, in 2015. She received her M.S. degrees in Engineering in 2007 and 2010 from Shanghai Jiao Tong University and Washington State University, respectively. Her research interests include computer vision, pattern recognition and machine learning, image/video processing and communication.

Sunil Kumar Dr. Kumar received his Ph.D. from Birla Institute of Technology and Science, Pilani (India) in 1997. Currently, he is a Professor and Thomas G. Pine Faculty Fellow in the Electrical and Computer Engineering Department at San Diego State University, CA, USA. His research interests include the QoS-aware and cross-layer protocols for wireless networks, and robust video compression (including H.264 and HEVC). He has published over 135 peer-reviewed journal/conference papers and four books. His research has been funded by NSF, DOD, DOE, California Energy Commission, and industry.