# Delay-Sensitive Resource Management in Multi-hop Cognitive Radio Networks

Hsien-Po Shiang and Mihaela van der Schaar

Department of Electrical Engineering (EE), University of California Los Angeles (UCLA) Los Angeles, CA
{hpshiang, mihaela} @ee.ucla.edu

*Abstract*–**Dynamic resource management by the various cognitive nodes fundamentally changes the *passive* way that wireless nodes are currently adapting their transmission strategies to match available wireless resources, by enabling them to *consciously* influence the wireless system dynamics based on the gathered information about other network nodes. In this paper, we discuss the main challenges of performing such dynamic resource management by emphasizing the *distributed information* in the dynamic multi-agent system. Specifically, the decisions on how to adapt the aforementioned resource management at sources and relays need to be performed in an informationally-decentralized manner, as the tolerable delay does not allow propagating information back and forth throughout the multi-hop infrastructure to a centralized decision maker. The term "cognitive" refers in our paper to both the capability of the network nodes to achieving large spectral efficiencies through exploitation and mitigation of channel and interference variability by dynamically using different frequency bands *as well as* their ability to learn the "environment" (channel conditions and source characteristic) and the actions of competing nodes through the designed information exchange. We propose our dynamic resource management algorithms performed at each network nodes integrated with multi-agent learning that explicitly consider the timeliness and the cost of such information exchange. The results show that our dynamic resource management approach improves the PSNR of multiple video streams by more than 3dB as opposed to the state-of-the-art dynamic frequency channel/route selection approaches without learning capability, when the network resources are limited.**

*Keywords –dynamic resource management, cognitive radio networks, multi-hop wireless networks, multi-agent learning, delay sensitive applications.*

## I. INTRODUCTION

The demand for wireless spectrum has increased and will keep increasing rapidly in the foreseeable future with the introduction of multimedia applications such as YouTube, peer to peer multimedia networks, and distributed gaming. However, scanning through the radio spectrum reveals its inefficient occupancy in most frequency channels. Hence, the Federal Communications Commission (FCC) suggested in 2002 [1] improvements for spectrum usage, which enable more efficient allocations of frequency channels to license-exempt users without impacting the primary licensees. Based on this, cognitive radio networks [2][3] were proposed which enable wireless users to *sense* and *learn* the surrounding environment and correspondingly *adapt* their transmission strategies.

In such cognitive wireless environments, two main challenges arise. The first challenge is how to sense the spectrum and model the behavior of the primary licensees to identify available frequency channels (spectrum holes)[1]. The second challenge is how to manage the available spectrum resources among the license-exempt users to satisfy their QoS requirements while limiting the interference to the primary licensees. In this paper, we focus on the second problem, i.e. the resource management, and rely on the existing literature for the first challenge [4][5].

The majority of the resource management research in cognitive radio networks has focused on a single-hop wireless infrastructure [6][7]. In this paper, we focus on the resource management problem in the more general setting of multi-hop cognitive radio networks. A key advantage of such flexible multi-hop infrastructures is that the same infrastructure can be re-used and reconfigured to relay the content gathered by various transmitting users (e.g. sources nodes) to their receiving users (e.g. sinks nodes). These users may have different goals (application utilities etc.) and may be located at various locations. For the multi-hop infrastructure, there are three key differences as opposed to the single-hop case. First, the users have as available network resources not only the vacant frequency channels (spectrum holes [2][7]) as in the single-hop case, but also the routes through the various network relays to the destination nodes. Second, the transmission strategies will need to be adapted not only at the source nodes, but also at the network relay nodes. In cognitive radio networks, network nodes are generally capable of sensing the spectrum and modeling the behavior of the primary users and thereby, identifying the available spectrum holes. In multi-hop cognitive radio networks, the network nodes will also need to model the behavior of the other neighbor nodes (i.e. other secondary users) in order to successfully optimize the routing decisions. In other words, network relays also require a learning capability in the multi-hop cognitive radio network. Third, to learn and

This work was sponsored by an ONR grant.

[1] In the wireless environment without primary licensees, such as the ISM band, there is no such problem. The main challenge is the resource management problem.

efficiently adapt their decisions over time, the wireless nodes need to possess accurate (timely) information about the channel conditions, interference patterns and other nodes transmission strategies. However, in a distributed setting such as a multi-hop cognitive radio network, the information is decentralized, and thus, there is a certain delay associated with gathering the necessary information from the various network nodes. Hence, an effective solution for multi-hop cognitive radio networks will need to tradeoff the "value" of having information about other nodes versus the transmission overheads associated with gathering this information in a timely fashion across different hops, in terms the utility impact.

In this paper, we will focus on delay-sensitive applications such as real-time multimedia streaming, i.e. the receiving users need to get the transmitted information within a certain delay. Due to the informationally decentralized nature of the multi-hop wireless networks, a centralized resource management solution for these delay-constrained applications is not practical [9], since the tolerable delay does not allow propagating information back and forth throughout the network to a centralized decision maker. Moreover, the complexity and the information overhead of the centralized optimization grow exponentially with the size of the network. The problem is further complicated by the dynamic competition for wireless resources (spectrum) among the various wireless nodes (i.e. source nodes/relays). The centralized optimization will require a large amount of time to process and the collected information will no longer be accurate by the time transmission decisions need to be made. Hence, a dynamic resource management solution, which explicitly considers the availability of information, the transmission overheads and incurred delays, as well as the value of this information in terms of the utility impact is necessary.

To design such a dynamic resource management in multi-hop cognitive radio networks, several main challenges need to be addressed:

- **Dynamic adaptation to a time-varying network environment**

To fulfill the requirements of delay sensitive applications given the dynamic nature of the cognitive radio networks, wireless nodes need to learn, dynamically self-organize and strategically adapt their transmission strategies to the available resources. In this paper, we will aim at modeling the behaviors of interacting cognitive radio nodes that use simple adaptation rules to sequentially adjust and optimize their transmission strategies. The outcomes of these interactions do not need to converge to an equilibrium, i.e., disequilibrium and perpetual adaptation of strategies may persist, as long as the performance of delay sensitive application is maximized. Hence, repeated information exchange across the networks is required for nodes to efficiently learn and adapt to the changing network dynamics.

- **Information availability in multi-hop infrastructures**

Due to the informationally-decentralized nature of the multi-hop wireless environment, information is only useful when it can be conveyed in time so that wireless nodes can manage the available resource in a distributed manner. The timeliness constraint of the information exchange depends on the delay deadline of the applications, the information overhead, and the condition of the network links, etc. Hence, the value of information in terms of its impact on the users' utilities will need to be quantified for the different settings of the multi-hop cognitive radio network. This information will impact the accuracy with which the wireless nodes can model the behavior of other nodes (including the primary users) and hence, respond to this environment by adequately optimizing their transmission strategies.

In summary, the paper makes the following contributions.

a) We study the required information exchange across the network nodes for dynamic resource management in multi-hop cognitive radio networks. Based on the required information, dynamic resource management algorithms are proposed. Given different types of information exchanges, learning strategies can be adopted by the wireless users. Specifically, we evaluate the cost of these different information overheads and their impacts on the application performance.

b) We introduce the notion of an "information cell" to explicitly identify the network nodes that can convey timely information to a certain center node. Importantly, we investigate the case that this information cell does not cover all the interfering neighbor nodes in the interference graph.

c) Based on the available information exchange, we propose a multi-agent learning approach which allows the various nodes to optimize their transmission strategies autonomously, in a distributed manner, in multi-hop cognitive radio networks. We discuss the tradeoffs between the cost of the required information exchange and the learning efficiency of the multi-agent learning approach in terms of the utility impact. This will provide insights for designing new algorithms for optimizing the various layers of the protocol stack in multi-hop cognitive radio networks.

The paper is organized as follows. Section II provides the multi-hop cognitive radio network settings and strategies and Section III gives problem formulation of the dynamic resource management for delay sensitive transmission in such networks. In Section IV, we determine how to quantify the rewards and costs associated with various information exchanges in the multi-hop cognitive radio networks. In Section V, we propose our dynamic resource management algorithms with information exchange. In Section VI, we introduce the multi-agent learning approach – adaptive fictitious play integrated in the proposed algorithms. Simulation results are in Section VII. Finally, Section VIII concludes the paper.

## II. Multi-hop Cognitive Radio Networks - Settings and Strategies

### A. Network entities

In this paper, we assume that a multi-hop cognitive radio network involves the following network entities and their interactions:

- **Primary Users (PUs)** are the incumbent devices that possess transmission licenses for specific frequency bands (channels). Without loss of generality, we assume that there are $M$ frequency channels in the considered cognitive radio network. We also assume that the maximum number of primary users that can be present in the network equals $M$. Note that these primary users can only occupy their assigned (licensed) frequency channels and not other primary users' channels. Since the primary users are licensed users, they will be guaranteed an interference-free environment [2][4]. When a primary user is not transmitting data using its assigned frequency channel, a spectrum hole is formed at the corresponding frequency channel.
- **Secondary Users (SUs)** are the autonomous wireless stations that perform channel sensing and access the existing spectrum holes in order to transmit their data. The secondary users can occupy the spectrum holes available in the various frequency channels. In this paper, the secondary users are deploying delay sensitive applications. Specifically, we assume that there are $V$ delay sensitive applications simultaneously sharing the cognitive radio network infrastructure, having unique source and destination nodes. These secondary users are able to deploy their applications across various frequency channels and routes.
- **Network Relays (NRs)** are autonomous wireless nodes that perform channel sensing and access the existing spectrum holes in order to relay the received data to one of its neighboring nodes or SUs. Hence, unlike in the SUs case, there is no source or destination present at the NRs. Note that multiple applications can use the same NR using different frequency channels.

### B. Source traffic characteristics

Let $V_i$ denote the delay sensitive application of the $i$-th SU. Assume that the application $V_i$ consists of packets in $K_i$ priority classes. The total number of applications is $V$. We assume that there are a total of $K = \sum_{i=1}^{V} K_i + 1$ priority classes (i.e., $\mathbf{C} = \{C_1, \dots, C_K\}$). The reason for adding an additional priority class is because the highest priority class $C_1$ is reserved for the traffic of the primary users. The rest of the classes $C_k, k > 1$ can be characterized by:

- $\lambda_k$, the impact factor of a class $C_k$. For example, this factor can be obtained based on the money paid by a user (different service levels can be assigned for different SUs by the cognitive radio network), based on the distortion impact experienced by the application of each SU or based on the tolerated delay assigned by the applications. The classes of the delay sensitive applications are then prioritized based on this impact factor, such that $\lambda_k \geq \lambda_{k'}$ if $k < k', k = 2, \dots, K$. The impact factor is encapsulated in the header (e.g. RTP header) of each packet.
- $D_k$, the delay deadline of the packets in a class $C_k$. In this paper, a packet is regarded useful for the delay sensitive applications only when it is received before its delay deadline.
- $L_k$, the average packet length in the class $C_k$.

A variety of delay sensitive applications can use the cognitive radio set-up discussed in this paper. Multimedia transmission such as video streaming or video conferencing can be examples of such applications [9]. We assume in this paper that an application layer scheduler is implemented at each network node to send the most important packet first based on the impact factor encapsulated in the packet header.

### C. Multi-hop cognitive radio network specification

We consider a multi-hop cognitive radio network, which is characterized by a general topology graph $\mathcal{G}(\mathbf{M}, \mathbf{N}, \mathbf{E})$ that has a set of primary users $\mathbf{M} = \{m_1, \dots, m_M\}$, a set of network nodes $\mathbf{N} = \{n_1, \dots, n_N\}$ (include SUs and NRs) and a set of network edges (links) $\mathbf{E} = \{e_1, \dots, e_L\}$ (connecting the SUs and NRs). There are a total of $N$ nodes and $L$ links in this network. Each of these $N$ network nodes is either a secondary user (as a source or a destination node) or a network relay.

We assume that $\mathbf{F} = \{f_1, \dots, f_M\}$ is the set of frequency channels in the network, where $M$ is the total number of the frequency channels. To avoid interference to the primary users, the network nodes can only use spectrum holes for transmission. Hence, to establish a link with its neighbor nodes, each network node $n \in \mathbf{N}$ can only use the available frequency channels in a set $\mathbf{F}_n \subseteq \mathbf{F}$. Note that these wireless nodes in a cognitive radio network will continuously sense the environment and exchange information and hence, $\mathbf{F}_n$ may change over time depending on whether the primary users are transmitting in their assigned frequency channels.

The network resource for a network node $n \in \mathbf{N}$ of the multi-hop cognitive radio network includes the routes composed by the various links and frequency channels. We define the resource matrix $\mathbf{R}_n = [R_{ij}] \in \{0,1\}^{L \times M}$ for the network node $n$ as follows:

$$R_{ij} = \begin{cases} 1, & \text{if link } e_i \text{ is connected to the node } n \\ & \text{and the frequency channel } f_j \text{ is available.} \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Whether or not the resource $R_{ij}$ is available to node $n \in \mathbf{N}$ depends not only on the topology connectivity, but also on the interference from other traffic using the same frequency channel. Next, we discuss the interference from other users (including the primary users).

### D. Interference characterization

Recall that the highest priority class $C_1$ is always reserved in each frequency channel for the traffic of the primary users. The traffic of the SUs can be categorized into $K-1$ priority classes ($C_2,...,C_K$) for accessing frequency channels. The traffic priority determines its ability of accessing the frequency channel. Primary users in the highest priority class $C_1$ can always access their corresponding channels at any time. The traffic of the SUs can only access the spectrum holes for transmission. Hence, we define two types of interference behavior in the considered multi-hop cognitive radio network:

*1) Interference from primary users.*

The primary users' interference depends on the location of the $M$ primary users. A primary user will block all the neighbor links using its frequency channel. After identifying the spectrum holes, each network node is able to obtain the interference matrix of the primary users $\mathbf{I}_1 = [I_{ij}] \in \{0,1\}^{L \times M}$:

$$I_{ij} = \begin{cases} 1, \text{ if the primary user is occupying frequency channel } f_j \\ \quad \text{and the link } e_i \text{ can interfere with the primary user.} \\ 0, \text{ otherwise.} \end{cases} \quad (2)$$

*2) Interference from competing secondary users.*

We define $\mathbf{I}_k = [I'_{ij}] \in \{0,1\}^{L \times M}$ as the interference matrix for the traffic in priority class $C_k$.

$$I'_{ij} = \begin{cases} 1, \text{ if link } e_i \text{ using frequency channel } f_j \text{ can be} \\ \quad \text{interfered by the traffic of priority class } C_k. \\ 0, \text{ otherwise.} \end{cases} \quad (3)$$

The interference caused by the traffic in priority class $C_k$ can be determined based on the interference graph of the nodes that transmit the traffic. The interference graph is defined as the corresponding links that are interfered by the transmission of the class $C_k$ traffic[2].

A simple example is illustrated in Figure 1, which indicates the interference matrix of the primary users and the resource matrix of each network node in the multi-hop cognitive radio network. The interference matrix can be computed by the information exchange among the neighbor nodes which can interfere with each other. The available resource matrix can be masked out by the interference matrix of the higher priority classes, i.e. $\mathbf{R}_{nk}^{(I)} = \mathbf{R}_n \otimes \overline{\mathbf{I}}_{k-1} \otimes ... \otimes \overline{\mathbf{I}}_1$, where the notation $\otimes$ represents element-wise multiplication of the matrixes and $\overline{\mathbf{I}}$ denotes the inverse operation, which turns 1 into 0 and 0 into 1. The resulting resource matrix $\mathbf{R}_{nk}^{(I)}$ represents the *available resource* around network node $n$ for the class $C_k$ traffic under the interference of other higher priority traffic (classes). Next, we define the actions available to the network nodes in a multi-hop cognitive radio network.

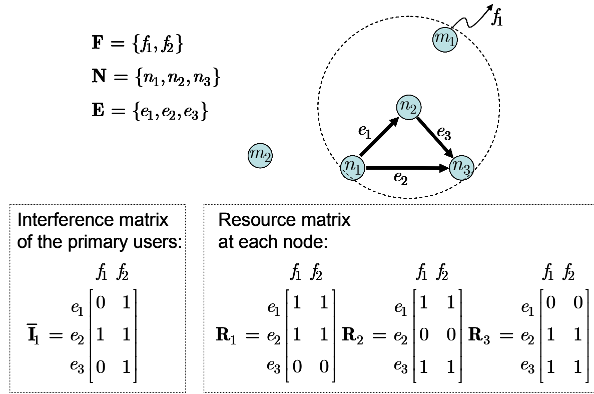


Fig. 1. A simple multi-hop cognitive radio network with three nodes and two frequency channels

## E. Nodes' actions

We define the action of the network node $n$ in order to relay the delay sensitive application $V_i$ as $A_n = (e \in \mathbf{E}_n, f \in \mathbf{F}_n)$. We assume that a network relay $n$ can select a set of links to its neighbor nodes (links connected to node $n$) $\mathbf{E}_n \subseteq \mathbf{E}$. Corresponding to the actions, we define the transmission strategy vector of the network node $n$ as $\mathrm{s}_n = [s_A \mid A = (e \in \mathbf{E}_n, f \in \mathbf{F}_n)]$, where $s_A$ represent the probability that the network node $n$ will choose an action $A$. We refer to an action at a node $n$ as a "*feasible action*" for transmitting a class $C_k$ traffic, if $A = (e,f)$ is an "*available resource*" in $\mathbf{R}_{nk}^{(I)}$ (i.e. element $R_{ef} = 1$ in $\mathbf{R}_{nk}^{(I)}$), since in this case the selected link and frequency channel do not interfere with the traffic in the higher priority classes. That is,

$$\hat{\mathbf{A}}_n(k) = \{A = (e,f) \mid \mathbf{R}_{nk}^{(I)} = [R_{ef}]^{L \times M}, R_{ef} = 1\}. \quad (4)$$

We denote the set of all the feasible actions for node $n$ as $\hat{\mathbf{A}}_n(k)$ for class $C_k$ traffic. We next determine the corresponding delay based on different actions, which considers the deployed cross-layer transmission strategies in order to compute the Effective Transmission Time (ETT) [15] over the transmission links.

Each network node $n$ computes the ETT $ETT_{nk}(e,f)$, with $e \in \mathbf{E}_n, f \in \mathbf{F}_n$ for transmitting delay sensitive applications in priority class $C_k$:

$$ETT_{nk}(e,f) = \frac{L_k}{T(e,f) \times (1 - p(e,f))}. \quad (5)$$

$T(e,f)$ and $p(e,f)$ represent the transmission rate and the packet error rate of using the frequency channel $f$ over the

[2] In a wireless environment, the transmission of neighbor links can interfere with each other and significantly impact their effective transmission time. Based on the interference graph, the action of a node can impact and be impact by the action of the other relay nodes. Hence, the definition of the interference graph plays an important role for characterizing the mutual interaction of the network nodes. For simplicity, we only consider the binary "1" and "0" when constructing the interference matrix.

link $e$. $T(e,f)$ and $p(e,f)$ are estimated by the MAC/PHY layer link adaptation [16], which can be modeled as sigmoid functions of the sensed Signal-to-Interference-Noise Ratio (SINR) $s(e,f)$ by the receiver of link $e$ using frequency channel $f$ [16]:

$$p(e,f) = \frac{1}{1+\exp(\zeta(s(e,f)-\delta))}, \quad (6)$$

$$T(e,f) \times (1-p(e,f)) = \frac{T(e,f)}{1+\exp(-\zeta(s(e,f)-\delta))}, (7)$$

where $\zeta$ and $\delta$ are empirical constants corresponding to the modulation and coding schemes used for transmitting a given packet length $L_k$ of class $C_k$ over link $e$ using frequency channel $f$. We assume that the channel condition of each link-frequency channel pair can be modeled using a continuous-time Markov chain [12] with a finite number of states $\mathbf{S}_{(e,f)}^n$. The time a channel condition spends in state $i \in \mathbf{S}_{(e,f)}^n$ is exponentially distributed with parameter $\nu_i$ (rate of transition at state $i$ in transitions/sec). We denote the maximum[3] transition rate of the network as $\nu$ and hence, the variation of the channel conditions in a time interval $\tau \le 1/\nu$ can be regarded to be negligible.

Define the action vector $\mathbf{A}_i = [A_n \mid n \in \boldsymbol{\sigma}_i]$ as the vector of the actions of all the network relay nodes for transmitting $V_i$. Assume that the $i$ th delay sensitive application $V_i$ are transmitted from the source node $n_i^s \in \mathbf{N}$ to the destination node $n_i^d \in \mathbf{N}$ with a total of $q_i$ packets. The routes of $V_i$ are denoted as $\boldsymbol{\sigma}_i = \{\sigma_{ij} \mid j = 1,...,q_i\}$, where $\sigma_{ij}$ is the route of the $j$ th packet in $V_i$. A route $\sigma_{ij}$ is a set of link-frequency pairs that the packets flow through, i.e.

$$\sigma_{ij} = \{(e,f) \mid \text{ the } j\text{th packet of } V_i \text{ flows} \\ \text{through link } e \text{ using frequency channel } f\}. \quad (8)$$

Note that if the action of a certain relay node changes, the corresponding route $\sigma_{ij}(\mathbf{A}_i)$ of relaying $V_i$ also changes. We denote the end-to-end delay of the packets transmitted using the route $\sigma_{ij}(\mathbf{A}_i)$ as $d_{ij}(\sigma_{ij}(\mathbf{A}_i))$. Based on the topology, each network relay node receiving a packet can decide where to relay the packet to and using which frequency channel, in order to minimize its end-to-end delay $d_{ij}(\sigma_{ij}(\mathbf{A}_i))$. Finally, to calculate $d_{ij}(\sigma_{ij}(\mathbf{A}_i))$, the source node need to obtain the delay information from other nodes according to the actions taken by the relay nodes, i.e.

$$d_{ij}(\sigma_{ij}(\mathbf{A}_i)) = \sum_{n\in\sigma_{ij}} ETT_{nk}(\mathbf{A}_i), \text{ for } V_i \in C_k. \quad (9)$$

# III. Resource Management Problem Formulation over Multi-hop Cognitive Radio Networks

By examining the cumulated ETT values, the objective of a delay sensitive application is to minimize its own end-to-end packet delay. The centralized and proposed distributed problem formulations are subsequently provided.

**- Centralized problem formulation with global information available at the sources**

If we assume that the global[4] information $\mathcal{G}_i$ is available to the source node $n_i^s$ for the delay sensitive application $V_i$, the route $\sigma_{ij}(\mathbf{A}_i,\mathcal{G}_i)$ can be determined for each packet $j$ of $V_i$. We will discuss the required information in more details in Section IV. The centralized optimization can be performed at every source node in order to maximize the utility $u_i$. Hence, for application $V_i$ we have:

$$u_i(\mathbf{A}_i,\mathcal{G}_i) = \sum_{j=1}^{q_i} w(\lambda_{ij}) \cdot \text{Prob}\{d_{ij}(\sigma_{ij}(\mathbf{A}_i,\mathcal{G}_i)) \le D_{ij}\}, \\ D_{ij} = D_k \text{ and } \lambda_{ij} = \lambda_k \text{ if } j \in C_k, \quad (10)$$

$$\text{with } \begin{array}{l} \mathbf{A}_i^{opt} = \arg\max u_i(\mathbf{A}_i,\mathcal{G}_i) \\ \qquad \text{subject to } A \in \hat{\mathbf{A}}_n \text{ for all } A \in \mathbf{A}_i \end{array}. \quad (11)$$

$w(\lambda_{ij})$ represents a weighting factor which depends on the impact factor $\lambda_{ij}$ of the packet $j$. However, due to the limited wireless network resource, the end-to-end delay constraint $d_{ij}(\sigma_{ij}(\mathbf{A}_i,\mathcal{G}_i)) \le D_k$ can make the optimization solution infeasible. Hence, greedy algorithms that perform optimizations sequentially from the highest priority class to the lowest priority class are commonly adopted [21][9]. Specifically, for class $C_k$, the following optimization is considered:

$$\mathbf{A}_{ik}^{opt} = \arg\min \sum_{j\in C_k} d_{ij}(\sigma_{ij}(\mathbf{A}_{ik},\mathcal{G}_i)) \\ \text{subject to } d_{ij}(\sigma_{ij}(\mathbf{A}_{ik},\mathcal{G}_i)) \le D_k, \quad , \quad (12) \\ A \in \hat{\mathbf{A}}_n \text{ for all } A \in \mathbf{A}_{ik}.$$

[3] In case that some of the channel conditions change severely in the network, a threshold $\nu_{th}$ can be set by protocols to avoid these fast-changing nodes and the $\nu$ is hence selected as the maximum transition rate below this threshold so that a time interval $\tau \le 1/\nu$ can still be used for the rest of the slow-changing nodes to adapt.

[4] The word "global information" means the information gathered from every node throughout the network.

where $\mathbf{A}_{ik} = [A_n \mid n \in \boldsymbol{\sigma}_{ij}, j \in C_k]$.

Due to the informationally decentralized nature of the multi-hop wireless networks, the centralized solution is not practical for the multi-user delay sensitive applications, as the tolerable delay does not allow propagating the global information $\mathcal{G}_i$ back and forth throughout the network to a centralized decision maker. For instance, the optimal solution depends on the delay $d_{ij}$ incurred by the various packets across the hops, which cannot be timely relayed to a source node. Especially when the network environment is time-varying, the gathered global information $\mathcal{G}_i$ can be inaccurate due to the propagation delay for this information. Moreover, the complexity of the centralized optimization grows exponentially with the number of classes and nodes in the network. The problem is further complicated by the dynamic adaptation of the transmission strategies deployed by the wireless nodes, which impacts their spectrum access and hence, implicitly, the performance of their neighbor nodes. The optimization will require a large amount of time to process and the collected information might no longer be accurate by the time transmission decisions need to be made.

In summary, in the studied dynamic cognitive radio network, the decisions on how to adapt the aforementioned actions at sources and relays need to be performed in a distributed manner due to these informational constraints. Hence, a "decomposition" of the optimization problem into distributed strategic adaptation based on the available local information is necessary.

**- Proposed distributed problem formulation with local information at each node:**

Instead of gathering the entire global information $\mathcal{G}_i$ at each source, we propose a distributed suboptimal solution that collects the local information $\mathcal{L}_n$ at node $n$ to minimize the expected delay of the various applications sharing the same multi-hop wireless infrastructure. Note that at each node $n$, the end-to-end delay for sending a packet $j \in C_k$ in equation (12) can be decomposed as:

$$d_{ij}(\sigma_{ij}) = d_n^P(\sigma_{ij}) + E[\tilde{d}_n(k,\sigma_{ij})], \tag{13}$$

where $d_n^P(\sigma_{ij})$ represents the past delay that packet $j$ has experienced before it arrives at node $n$ and $E[\tilde{d}_n(k,\sigma_{ij})]$ represents the expected delay from the node $n$ to the destination of the packet $j \in C_k$. The sending packet $j \in C_k$ is determined by the application layer scheduler according to the impact factor $\lambda_k$. The information about $\lambda_k$ can be encapsulated in the packet header and $d_n^P(\sigma_{ij})$ can be calculated based on the timestamp available in the packet header. The priority scheduler at each node ensures that the higher priority classes are not influenced by the lower priority classes (see equation (12)). Since at the node $n$ the value of $d_n^P(\sigma_{ij})$ is fixed, the optimization problem at the node $n$ becomes:

$$\begin{aligned} A_n^{opt} &= \arg\min E[\tilde{d}_n(k,\sigma_{ij}(A_n,\mathcal{L}_n))] \\ &\text{subject to } E[\tilde{d}_n(k,\sigma_{ij}(A_n,\mathcal{L}_n))] \le \\ &\qquad D_k - d_n^P(\sigma_{ij}) - \rho,\ j \in C_k, A_n \in \hat{\mathbf{A}}_n \end{aligned} \tag{14}$$

where $E[\tilde{d}_n(k,\sigma_{ij}(A_n,\mathcal{L}_n))]$ represents the expected delay from the relay node $n$ to the destination of the packets in class $C_k$. $\rho$ represents a guard interval such that the probability $\text{Prob}\{E[\tilde{d}_n(k,\sigma_{ij}(A_n,\mathcal{L}_n))] + d_n^P(\sigma_{ij}) > D_k\}$ is small (as in [18]). To estimate the expected delay $E[\tilde{d}_n(k,\sigma_{ij}(A_n,\mathcal{L}_n))]$ in equation (14), each network node $n$ maintains an estimated transmission delay $E[\tilde{d}_n(k)]$ from itself to the destination for each class of traffic using the Bellman-Ford shortest-delay routing algorithm [12]. We assume that each node $n$ maintains and updates a delay vector $\mathbf{d}_n = [E[\tilde{d}_n(2)],\ldots,E[\tilde{d}_n(K)]]$ (note that the first priority class is reserved for the primary users) with elements for each class with a predetermined destination. The network node exchanges such information to its neighbor nodes (we will discuss the delay information update in details in Section VI.). Two key remaining questions are what other information should be exchanged for learning the routing strategy of the other nodes and, more importantly, how much information can be timely conveyed to a network node for the distributed optimization. Next, we are going to discuss the dynamic resource management with information constraints in more details.

## IV. Dynamic Resource Management with Information Constraints

### A. *Benefit of acquiring information and information constraints*

For the network node $n$, the local information $\mathcal{L}_n$ gathered from different network nodes has different impact on decreasing the objective function $E[\tilde{d}_n(k,\sigma_{ij}(A_n,\mathcal{L}_n))]$ in equation (14). Let $\mathcal{I}_n(x) = \{\mathbf{I}_k(n_x, A_{n_x}), A_{n_x}, \mathbf{d}_{n_x} \mid n_x \in \mathbf{N}_x^n\}$ denote the set of local information gathered from the neighbor nodes, which is $x$ hops away from node $n$, where $\mathbf{N}_x^n$ represents a set of nodes that is $x$ hops away from node $n$. We define $\mathcal{L}_n(x) = \{\mathcal{I}_n(l) \mid l = 1,\ldots,x\}$ as the local information gathered from all of these neighbor nodes. Given the local information $\mathcal{L}_n(x)$, we define the optimal expected delay as $K_n(k,x) = E[\tilde{d}_n(k,\sigma_{ij}(A_n^{opt},\mathcal{L}_n(x)))]$.

We assume that the network nodes within the $h$ hops form an *information cell*. Only the local information $\mathcal{L}_n(h)$ *within* the information cell is useful to the node $n$, since the reward of information is zero, i.e. $J_n(k,\mathcal{I}_n(x)) = 0$ for $\forall x > h(\nu)$. In the dynamic network, network node $n$ determines its action at time slot $t$ based on the acquired information at the previous time slot $t-1$. The optimization problem in equation (14) can be written as:

$$A_n^{opt}(t) = \arg\min E[\tilde{d}_n(k, \sigma_{ij}(A_n, \mathcal{L}_n(h, t-1)))]$$
$$\text{subject to } E[\tilde{d}_n(k, \sigma_{ij}(A_n, \mathcal{L}_n(h, t-1)))] \leq \quad .$$
$$D_k - d_n^P(\sigma_{ij}) - \rho,\ j \in C_k, A_n \in \hat{\mathbf{A}}_n(t-1) \tag{15}$$

Recall that the neighbor nodes of the node $n$ are defined as the nodes that can interfere or can be interfered by the node $n$ (within $H_n^I$ hops), which may not align with the range of the information cell (within $h$ hops). If all neighbor nodes are within the $h$-hop information cell, all necessary information are timely conveyed to the node $n$. Otherwise, the neighbor nodes that are too far away cannot convey the interference information to the node $n$ in time. Since the required information cannot be acquired in time, the solution in equation (15) becomes suboptimal. We refer to this problem as "information exchange mismatch" problem.

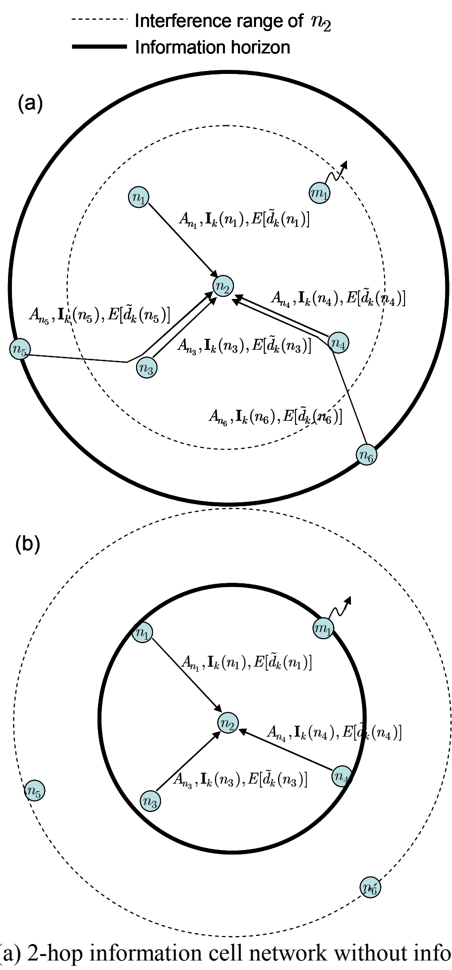


Fig 2. (a) 2-hop information cell network without information exchange mismatch problem.
(b) 1-hop information cell network with information exchange mismatch problem.

Figure 2 illustrates two simple network examples with and without the mismatch problem. In fact, due to the nature of the multi-hop wireless environment, the network nodes that are far away from the node $n$ have limited interference impact on node $n$. Hence, even though the information horizon $h$ does not match the interference range as in Figure 2, the performance degradation of the optimization problem in equation (15) using the local information $\mathcal{L}_n(h)$ is limited.

### B. Cost of information exchange

In the previous subsection, we discuss the reward of information in an $h$-hop information cell while ignoring the negative impact of the information exchange. In this section, we discuss the cost (increase of the expected delay) due to this information exchange. Denote the duration of the time slot as $t_I(\nu)$, which is also the interval between the repeated information exchanges in the network (the concept is similar to the time slots in the service interval in IEEE 802.11e [14]). We define there are $c$ time slots in $\tau$ seconds, i.e.

$$t_I(\nu) = \frac{\tau(\nu)}{c}. \tag{16}$$

$c$ defines the frequency of the decision making as well as the learning process, which will be discussed in detail in Section VI. Note that decisions can be made every $t_I$ and this time slot duration is short enough compared to $\tau$. Hence, the network changes in $t_I$ is also negligible.

We denote the duration of each information exchange at the network node $n$ as $d_I(\mathcal{L}_n(h))$. Assume the information unit for the required information is $U^{(I)}$, $U^{(A)}$, and $U^{(d)}$ per class, respectively. Assume the average number of nodes in an $h$-hop information cell is $\bar{N}(h)$. The information time overhead of $\mathcal{L}_n(h)$ is on average $d_I(\mathcal{L}_n(h)) = \bar{N}(h)[(K-1)(U^{(d)} + U^{(I)}) + U^{(A)}]$.

Note that even though the information exchange may be implemented in a designated coordination channel [8], a network node with a single antenna cannot transmit both the data and the control signals at the same time[5]. This information exchange time overhead decreases the effective transmission rate at node $n$ using the line $e$ and frequency channel $f$:

$$T'(e,f) = \frac{t_I(\nu) - d_I(\mathcal{L}_n(h))}{t_I(\nu)} \times T(e,f). \tag{17}$$

Figure 3 provides an illustration of the time line for transmission at the node $n$. Hence, the effective transmission time at a node $n$ using the line $e$ and frequency channel $f$ to transmit a packet in class $C_k$ becomes:

[5] For the case with multiple antennas that the data and control signals can be received at the same time, there will be no such information exchange time overhead.

$$ETT'_{nk}(e,f) = \frac{t_I(\nu)}{t_I(\nu) - d_I(\mathcal{L}_n(h))} \times ETT_{nk}(e,f)\,. \quad (18)$$

In conclusion, the increase of the effective transmission time degrades the performance of the delay sensitive applications. The degradation depends on the content of the local information exchange $\mathcal{L}_n(h)$.

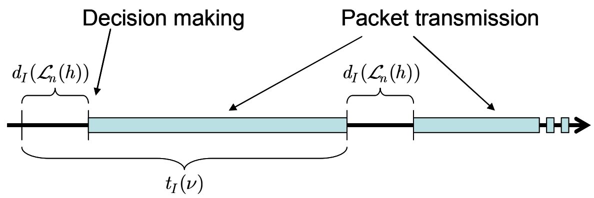


Fig. 3. Transmission time line at the node $n$ with network variation speed $\nu$ information horizon $h$ and local information $\mathcal{L}_n(h)$.

Next, we discuss the proposed dynamic resource management algorithm based on the information exchanges and learning capabilities to tackle the optimization problem in equation (15).

## V. DYNAMIC RESOURCE MANAGEMENT ALGORITHMS

Figure 4 provides a system diagram of the proposed dynamic resource management. First, a packet $j \in C_k$ is selected from the application scheduler at the node $n$ based on the impact factor $\lambda_k$ of the packet and an action $A_n$ is taken for that packet. The application layer information including $C_k, L_k, D_k$ is conveyed to the network layer for this action decision. Network conditions $T(e,f), p(e,f)$ are then conveyed from the MAC/PHY layer for computing the ETT values using equation (5).

In addition to the $T(e,f), p(e,f)$, the action selection is impacted by the interference induced from the action of these neighbor nodes and hence, the information received from the neighbor nodes in the information cell. Recall that $\mathcal{L}_n(h) = \{\mathcal{I}_n(l) \mid l = 1,...,h\}$. We use the notation $-n(h)$ to represent the set of the neighbor nodes of the network node $n$ in the $h$-hop information cell. Hence, the local information exchanged $\mathcal{L}_n(h) = \{\mathbf{I}_k(-n(h), A_{-n(h)}), A_{-n(h)}, \mathbf{d}_{-n(h)}\}$ across the network nodes is required. Hence, the node $n$ knows the estimated delay $\mathbf{d}_{-n(h)}$ from its neighbor nodes to the destinations, so as the actions $A_{-n(h)}$ of its neighbor nodes and their interference matrixes $\mathbf{I}_k(-n(h), A_{-n(h)})$ [6]. Based on the delay information from the neighbor nodes, a network node can update its own estimated delay to the various destinations using Bellman-Ford algorithm [12] by adding the ETT value of the feasible actions and determine the minimum-delay action based on it.

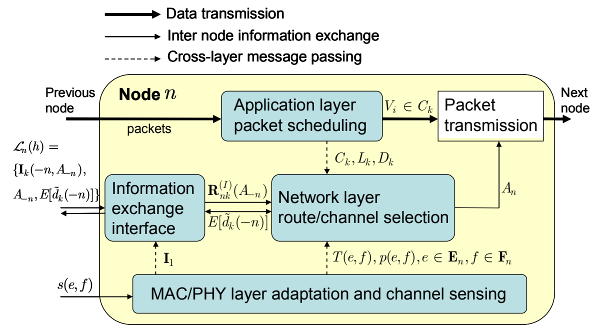


Fig. 4. System diagram of the proposed dynamic resource management.

We separate the dynamic resource management into two blocks at the node $n$ as in Figure 4 – the information exchange interface block that regularly collects required local information and the route/channel selection block to determine the optimal action. We now discuss the role of the exchanged information and the two algorithms implemented in these blocks, respectively.

The next algorithm is performed at network node $n$ at the information exchange interface in Figure 4.

*Algorithm 1. Periodic information exchange algorithm:*

**Step 1. Collect the required information** – the node $n$ first collects the required information $\mathcal{L}_n(h) = \{\mathbf{I}_k(-n(h), A_{-n(h)}), A_{-n(h)}, \mathbf{d}_{-n(h)}\}$ from the neighbor nodes in the information cell.

**Step 2. Learn the behavior of the neighbor nodes** – by continuously monitoring the actions of the neighbor nodes, node $n$ can model the behavior of the neighbor nodes or learn a better transmission strategy using strategy vectors $\mathbf{s}(n') = [s_A(n') \mid A = (e \in \mathbf{E}_{n'}, f \in \mathbf{F}_{n'})]$, $n' \in -n(h)$, where $s_A(n')$ represents the probability (strategy) of selecting an action $A$ by the node $n'$, which will be discussed in Section VI.

**Step 3. Estimate the resource matrix** – from the interference matrix $\mathbf{I}_k(n', A_{n'})$ gathered from the neighbor node $n'$, the resource matrix can be obtained for each class of traffic by $\mathbf{R}^{(I)}_{nk} = \mathbf{R}_n \otimes \overline{\mathbf{I}}_{k-1} \otimes ... \otimes \overline{\mathbf{I}}_1$, which will be explained in Section V.*A* in more details. Then the available resource $\mathbf{R}^{(I)}_{nk}(A_{-n})$ are provided to the network layer route/channel selection block stated in the Algorithm 2.

[6] In the case that the network topology is fixed (where node mobility is not considered), if the topology information and primary users' information are known for all the nodes in the network, the interference matrix $\mathbf{I}_k(-n(h), A_{-n(h)})$ in $\mathcal{L}_n(h)$ can be computed from $A_{-n(h)}$ at any node in the network. Hence, only $\mathcal{L}_n(h) = \{A_{-n(h)}, \mathbf{d}_{-n(h)}\}$ is required instead of $\mathcal{L}_n(h) = \{\mathbf{I}_k(-n(h), A_{-n(h)}), A_{-n(h)}, \mathbf{d}_{-n(h)}\}$.

**Step 4. Update information** $\{\mathbf{I}_k(n,A_n), A_n, \mathbf{d}_n\}$ – based on the recently selected action $A_n$, the latest delay vector $\mathbf{d}_n$, and the interference matrix $\mathbf{I}_k(n,A_n)$. Two different interference models are considered in this paper when constructing the interference matrix $\mathbf{I}_k(n,A_n)$ from equation (3):

(1) A network node can transmit and receive packets at the same time – Note that a node cannot reuse a frequency channel $f \in \mathbf{F}_n$ used by its neighbor nodes. If a frequency channel is used by its neighbor nodes, all the elements in the column of the interference matrix $\mathbf{I}_k(n,A_n)$ that is associated with the frequency channel are set to 1.

(2) A network node cannot transmit and receive packets at the same time – In this case, if the frequency channel $f \in \mathbf{F}_n$ is used, all the elements in the column of the interference matrix $\mathbf{I}_k(n,A_n)$ associated with the frequency channel are set to 1. In addition, if a network link $e \in \mathbf{E}_n$ is used by its neighbor nodes, all the elements of the interference matrix $\mathbf{I}_k(n,A_n)$ that is associated with the node $n$ are also set to 1, no matter what frequency channel it uses.

**Step 5. Broadcast the information** $\{\mathbf{I}_k(n,A_n), A_n, \mathbf{d}_n\}$ **and repeat the algorithm periodically in every** $t_I(\nu)$ **seconds.**

The next algorithm is performed at the network node $n$ at the network layer route/channel selection block in Figure 4.

*Algorithm 2. Minimum-delay route/channel selection algorithm:*

**Step 1. Determine the packet to transmit** – Based on the impact factor, one packet $j$ in the buffer at the node $n$ is scheduled to be transmitted. Assume the packet $j \in C_k$, and the information of $C_k$, $L_k$, $D_k - d_n^P$ are extracted or computed from the application layer.

**Step 2. Estimate the channel condition** – The transmission rate $T(e,f)$ and packet error rate $p(e,f)$ for each link-frequency channel pair $(e \in \mathbf{E}_n, f \in \mathbf{F}_n)$ are provided from the PHY/MAC layer through link adaptation [16] (see equation (6) and (7)).

**Step 3. Construct the feasible action set** – Construct the feasible action set $\hat{\mathbf{A}}_n(k)$ from the resource matrix $\mathbf{R}_{nk}^{(I)}$ given from the information exchange interface for the priority class $C_k$ at the node $n$ (see equation (4)).

**Step 4. Calculate the expected delay toward the destination** – for each action $A_n \in \hat{\mathbf{A}}_n(k)$ of the traffic class $C_k$:

$$E[\tilde{d}_n(k,A_n)] = ETT_{nk}(A_n) + E[\tilde{d}_{n'(A_n)}(k)], \quad \text{for } \forall A_n \in \hat{\mathbf{A}}_n(k), \tag{19}$$

where $E[\tilde{d}_{n'(A_n)}(k)]$ represents the corresponding element for the class $C_k$ in the delay vector $\mathbf{d}_{-n}$ from the neighbor node $n'(A_n)$. $ETT_{nk}(A_n)$ can be calculated based on $L_k$, $T(e,f)$, and $p(e,f)$ using equation (5).

**Step 5. Check the delay deadline** – if $E[\tilde{d}_n(k)] \geq D_k - d_n^P - \rho$, drop the packet.

**Step 6. Select the minimum delay action** – if $E[\tilde{d}_n(k)] < D_k - d_n^P - \rho$, find the minimum-delay route and frequency channel selection, i.e. determine the optimal action $A_n^{opt}$ from the feasible action set $\hat{\mathbf{A}}_n(k)$. In other words, the goal here is to solve equation (15) at node $n$:

$$A_n^{opt} = \arg \min_{A_n \in \hat{\mathbf{A}}_n(k)} E[\tilde{d}_n(k,A_n)]. \tag{20}$$

Note that the feasible action set $\hat{\mathbf{A}}_n(k)$ in equation (20) depends on the actions of other neighbor nodes $A_{-n}$. It is important for the network nodes to adopt learning approaches for modeling the behaviors of these network nodes to decrease the complexity of the dynamic adaptation. This will be discussed in details in Section VI.

**Step 7. Update the delay and the current action information** – after selecting the optimal action, update the estimated delay $E[\tilde{d}_n(k)]$ using

$$E[\tilde{d}_n(k)] = E[\tilde{d}_n(k,A_n^{opt})], \tag{21}$$

and provide the updated delay vector $\mathbf{d}_n = [E[\tilde{d}_n(2)],...,E[\tilde{d}_n(K)]]$ to Algorithm 1 at the information exchange interface.

## VI. Adaptive Fictitious Play

In this section, we provide a learning approach for the SUs to learn the feasible action set $\hat{\mathbf{A}}_n(k)$ in equation (20) for our dynamic resource management algorithms introduced in Section V. Specifically, based on the information exchange $\mathcal{L}_n(h)$, the behaviors of the neighbor nodes in the information cell can be learned (Step 2 of Algorithm 1) and based on the behaviors, the feasible action set $\hat{\mathbf{A}}_n(k)$ is determined. The learning approach is similar to the fictitious play in [10], applied when the SUs are willing to reveal their current action information and thereby, they are able to model the behaviors (strategies) of other SUs (a model-based learning [13]). However, due to the information constraint discussed in Section IV, only the information from the neighbor nodes in the information cell is useful. Hence, we adapt the fictitious play learning approach to our considered network setting.

### *A. Adaptive fictitious play (AFP)*

Fictitious play is a learning approach to model the behaviors (strategies) of the other users [10]. However, in our cognitive radio network setting, the gathered information is limited within the information cell instead of all the other network nodes. Hence, only part of the SUs can be modeled via the learning approach depending on the information horizon. Specifically, a node $n$ maintains a strategy vector over time $\mathbf{s}(n',t) = [s_A(n',t) \mid A = (e \in \mathbf{E}_{n'}, f \in \mathbf{F}_{n'})]$ for each of its neighbor nodes $n' \in -n(h)$ in the information cell. $s_A(n',t)$ represents the frequency selection strategy of the node $n'$ making action $A$ at time $t$, which is obtained using:

$$s_A(n',t) = \frac{r_A(n',t)}{\sum_{A \in (\mathbf{E}_{n'}, \mathbf{F}_{n'})} r_A(n',t)}, \tag{22}$$

where $r_A(n',t)$ is the *propensity* [11] of node $n'$ for taking action $A$ at time $t$, which can be computed by:

$$r_A(n',t) = \alpha \times r_A(n',t-1) + I(A_{n'}(t) = A), \tag{23}$$

where $\alpha < 1$ is a discount factor quantifying the importance of the history value. $I(A_{n'}(t) = A)$ represents an indicator function such that,

$$I(A_{n'}(t) = A) = \begin{cases} 1, \text{ if the action of the node } n' \\ \qquad \text{at time } t \text{ is } A \\ 0, \text{ otherwise} \end{cases} . \tag{24}$$

As stated in Section II.*E*, $s_A(n',t)$ represent the probability that the network node $n'$ will choose an action $A$. Hence, the probability $s_A(n',t)$ for modeling the node $n'$ making an action $A$ at time $t$ will increase with the actual times that the action $A$ is selected. Based on the strategy $s_A(n',t)$, the adaptive fictitious play provide the estimated interference matrix $\mathbf{I}_k$, and then the feasible action set $\hat{\mathbf{A}}_n(k)$ can be computed.

From the gathered interference matrix $\mathbf{I}_k(n', A_{n'})$ from the neighbor node $n' \in -n(h)$, the node $n$ can compute the expected interference matrix from

$$\mathbf{I}_k^e = [I_{ij}^e] = \sum_{n' \in -n(k)} \mathbf{I}_k(n') = \sum_{n' \in -n(k)} \sum_A s_A(n') \mathbf{I}_k(n', A). \tag{25}$$

Then, the node $n$ can estimate the interference matrix $\mathbf{I}_k$ for the traffic in class $C_k$:

$$\mathbf{I}_k = [I_{ij} \mid I_{ij} = \begin{cases} 1, \text{ if } I_{ij}^e \geq \mu \\ 0, \text{ if } I_{ij}^e < \mu \end{cases}], \tag{26}$$

where $\mu$ represents a threshold value that determines whether or not a link-frequency-channel pair $(e,f)$ is considered to be occupied. Feasible action set $\hat{\mathbf{A}}_n(k)$ can hence be learned based on the resource matrix $\mathbf{R}_{nk}^{(I)} = \mathbf{R}_n \otimes \bar{\mathbf{I}}_{k-1} \otimes \ldots \otimes \bar{\mathbf{I}}_1$ using equation (4). By learning the feasible action set $\hat{\mathbf{A}}_n(k)$, the best response actions are computed using equation (20).

### *B. Information exchange overhead reduction*

The fictitious play suffers from the large information overhead, since it requires all the local information $\mathcal{L}_n(h) = \{\mathbf{I}_k(-n(h), A_{-n(h)}), A_{-n(h)}, \mathbf{d}_{-n(h)}\}$ in the $h$–hop information cell. We know that the overhead can increase the expected delay, especially when the network change slowly (i.e. with a large information cell). Hence, the overhead reduction is required to mitigate the performance degradation.

**(1) Reducing the information horizon.**

Recall that the information overhead of $\mathcal{L}_n(h)$ is $\bar{N}(h)[(K-1)(U^{(d)} + U^{(I)}) + U^{(A)}]$ in average ($\bar{N}(h)$ is the average number of nodes in an $h$-hop information cell). With an information horizon $h' < h$, the overhead becomes $\bar{N}(h')[(K-1)(U^{(d)} + U^{(I)}) + U^{(A)}]$, where $\bar{N}(h') < \bar{N}(h)$. Alternatively, the reward of information $J_n^d(k, \mathcal{I}_n(x))$, $x < h$ provides a metric to select the most valuable information from the nodes within the information cell.

**(2) Reducing the number of classes.**

From equation (14), we know that the higher priority classes will not be influenced by the lower priority classes. Hence, the information overhead can be reduced by ignoring the information exchange of the lower priority classes. The overhead becomes $\bar{N}(h)[(k'-1)(U^{(d)} + U^{(I)}) + U^{(A)}]$, $k' < K$.

**(3) Reducing the frequency of learning.**

Although we divide $c$ time slots in $\tau$ seconds, a network node $n$ does not have to learn in all these $c$ time slots. In other words, the periodic learning process of the node $n$ does not have to be aligned with the information exchange (decision making). In order to avoid simultaneous learning among network neighbors in a distributed manner, at each time slot, the network node $n$ updates the strategy vector $s_A(n',t)$ with probability $\varepsilon_n = b_n / c$ ($b_n \leq c$), and keeps the same strategy vector with probability $1 - \varepsilon_n$. In other words, the network node $n$ chooses $b_n$ time slots out of $c$ time slots in $\tau$ seconds to model the behavior of other neighbor nodes. Note that the parameter $b_n$ characterize the speed of learning at different network node $n$. The larger $b_n$ gives the network node $n$ faster learning capability. The information overhead of $\mathcal{L}_n(h)$ becomes $b_n / c \times \bar{N}(h)[(K-1)(U^{(d)} + U^{(I)}) + U^{(A)}]$.

## VII. SIMULATION RESULTS

We simulate two video streaming applications that are transmitting videos $V_1$ "Coastguard" and $V_2$ "Mobile" (16 frames per GOP, frame rate of 30Hz, CIF format) over the same multi-hop cognitive radio network. Each video sequence is divided into four priority classes ($K_i = 4, K = 9$) with average packet length $L_k$ = 1000 bytes and delay deadline $D_k$ = 500 millisecond. Although the first priority class $C_1$ is reserved for the primary users, let us first consider the case when there are no primary users, i.e. only the SUs and NRs are transmitting. We assume that there are two frequency channels ($M$=2). The wireless network topology is shown in Figure 5 in a 100x100 meters region with $N$ = 15 nodes and $L$ = 22 links similar to the network settings in [17]. A link is established as long as the channel condition (described in the paper by the link SINR) is acceptable within the transmission distance (approximately 36 meters). Note that this transmission distance is not aligned with the interference range $H_n^I$. Neighbor nodes that are beyond the transmission distance can still interfere with each other.

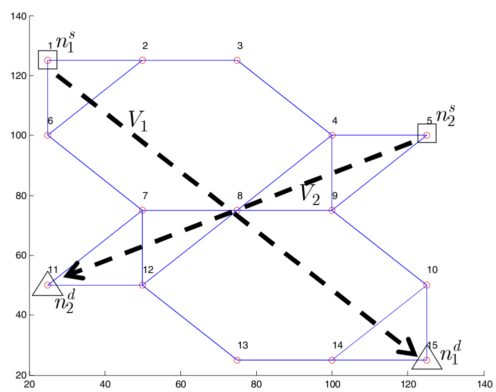


Fig. 5. Wireless network settings for the simulation of two video streams.

### A. *Application layer performance with different information horizons and interference ranges*

We next compare the proposed dynamic resource management algorithm using adaptive fictitious play (AFP) with two other resource management methods – AODV [19] with load balancing over the two available frequency channels (AODV/LB) and the Dynamic Least Interference Channel Selection [20] (DCS) extended to a network setting. Table I and II show the results of the Y-PSNR of the two video sequences using different approaches. The results show that the proposed algorithm using learning from the nodes within the information cell outperforms the alternative approaches. Especially, when the interference range is large ($H_n^I$ = 80 meters), the proposed AFP approach significantly improves the video quality (X represents PSNR below 26 dB, which is unacceptable for a viewer).

TABLE I. Y-PSNR OF THE TWO VIDEO SEQUENCES USING VARIOUS APPROACHES ($H_n^I$ = 40 METERS)

| Network Bandwidth | | Y-PSNR (dB) AODV/LB | DCS | AFP (1-hop information cell) |
|---|---|---|---|---|
| Average $T$ =5.5 Mbps | $V_1$ | 32.47 | 35.21 | 35.61 |
| | $V_2$ | 31.70 | 33.32 | 33.32 |

TABLE II. Y-PSNR OF THE TWO VIDEO SEQUENCES USING VARIOUS APPROACHES ($H_n^I$ = 80 METERS)

| Network Bandwidth | | Y-PSNR (dB) AODV/LB | DCS | AFP (1-hop information cell) | AFP (2-hop information cell) |
|---|---|---|---|---|---|
| Average $T$ =5.5 Mbps | $V_1$ | X | X | 28.19 | 29.80 |
| | $V_2$ | X | X | 31.26 | 31.70 |
| Average $T$ =10 Mbps | $V_1$ | 30.47 | 34.46 | 35.61 | 35.61 |
| | $V_2$ | 31.92 | 33.08 | 33.32 | 33.32 |

### B. *Impact of the primary users*

The simulation implies that the reward of information is also impacted by the existence of the primary users. Next, we consider the impact of the primary users, which always have higher priority to access the pre-assigned frequency channels than the network nodes in Figure 5. Assume that the frequency channel $F_1$ is occupied by the primary users with time fraction $\rho$=0%, 20%, 40%, 60%, and 80% around a certain congestion region (network nodes $n$ = 7, 11, 12) in Figure 5. Figure 6 shows the packet loss rate for the two video streams using the AFP with various information horizons. The average transmission rate is set to 5.5 Mbps, and the interference rage is 80 meters.

The results show that as the time fraction $\rho$ increases, the packet loss rates of both applications increase, since fewer resources are available for the secondary users to transmit the packets. As the simulation in the previous subsection, when the interference rage is 80 meters, AFP with 2-hop information cell still performs better than 1-hop information cell case. Interestingly, for application $V_1$, AFP with 3-hop information cell performs even better in a large $\rho$ case, even though more cost of information is needed. This is because the congestion region are more likely to be discovered at the source node $n$ =1 and detour the packets through other routes. However, such advantage is not exploited by the application $V_2$, since its destination node is affected by the primary users and there is no way to detour the packets. Note that when there is no primary user ($\rho$ = 0), AFP with 3-hop information cell performs worse than 2-hop case due to the larger cost of information exchange.

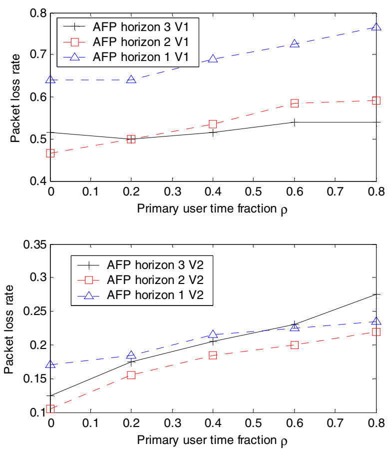


Fig. 6. Packet loss rate vs. time fraction $\rho$ of the primary users occupying frequency channel $F_1$ AROUND NETWORK NODE $n$ = 7, 11, 12 (average $T$ =5.5Mbps, $H_n^I$ = 80 meters).

## VIII. CONCLUSIONS

In this paper, we show that the dynamic resource management integrated with multi-agent learning significantly improves the performance of delay sensitive applications in a multi-hop cognitive radio network. The network nodes need to learn about the wireless environment as well as the behavior of the neighbor nodes based on the information exchange. In the dynamic multi-hop cognitive radio networks, different information has different values for the network nodes in terms of their utility impact. Based on the reward of the obtained information (i.e. the impact on decreasing the expected end-to-end delay), we define the information horizon in our adaptive fictitious play. The results show that our delay-sensitive resource management approach improves the performance of multiple video streams significantly as opposed to the state-of-the-art dynamic frequency channel/route selection approaches without learning capability, when the network resources are limited. In addition to the reward, the cost of the information exchange is also considered in terms of transmission time overheads. Various approaches of decreasing this time overhead are discussed and their performance impact is quantified.

## REFERENCES


[1] Federal Communications Commission (FCC), "Spectrum Policy Task Force," *ET Docket* no. 02-135, Nov 15, 2002.

[2] S. Haykin, "Cognitive Radio: Brain-Empowered Wireless Communications," in *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, Feb 2005.

[3] J. Mitola etal., "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13-18, Aug. 1999.

[4] I. F. Akyildiz, W. –Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey," *Computer Networks: The International Journal of Computer and Telecommunication Networking*, vol. 50, issue 13, Sep 2006.

[5] T. X. Brown, "An analysis of unlicensed device operation in licensed broadcast service bands," in *Proc. IEEE DySPAN 2005*, Nov 2005, pp. 11-29.

[6] N. Devroye, P. Mitran, and V. Tarokh, "Limits on communications in a cognitive radio channels," *IEEE Communications Mag.*, vol. 44, pp. 44-49, June 2006.

[7] C. Cordeiro, K. Challapali, D. Birru and S. Shankar N, "IEEE 802.22: An Introduction to the First Wireless Standard based on Cognitive Radios," *Journal of Communications, Academy Publishers*, vol. 1, issue 1, Apr 2006.

[8] V. Brik, E. Rozner, S. Banarjee, P. Bahl, "DSAP: a protocol for coordinated spectrum access," in *Proc. IEEE DySPAN 2005, Nov. 2005*, pp. 611-614.

[9] H. Shiang and M. van der Schaar, "Informationally Decentralized Video Streaming over Multi-hop Wireless Networkss," *IEEE Trans. Multimedia*, vol. 9, issue 6, Sep 2007.

[10] D. Fudenberg and D.K. Levine, *The Theory of Learning in Games*, MIT Press, Cambridge, MA, 1998.

[11] H. P. Young, *Strategic Learning and its Limits*, Oxford University Press, NY, 2004.

[12] D. Bertsekas, R. Gallager, *Data Networks*, Prentice Hall, Inc. Upper Saddle River, NJ, 1987.

[13] Y. Shoham, R. Powers, and T. Grenager, "If multi-agent learning is the answer, what is the question?" *Artificial Intelligence*, 2007. to appear.

[14] IEEE 802.11e/D5.0, DRMft Supplement to Part 11: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS), Jun 2003.

[15] R. Draves, J. Padhye, and B. Zill, "Routing in multi-radio, multi-hop wireless mesh networks," in *Proc. ACM Internat. Conf. on Mob. Computing and Networking (MOBICOM)*, 2004, pp. 114-128.

[16] D. Krishnaswamy, "Network-assisted Link Adaptation with Power Control and Channel Reassignment in Wireless Networks," *3G Wireless Conference*, pp. 165-170, 2002.

[17] D. Krishnaswamy, et al. "A cross-layer cross-overlay architecture for proactive adaptive processing in mesh networks," *2nd IEEE Workshop on Wireless Mesh Networks (WiMesh 2006)*, 2006. pp. 74-82.

[18] D. Jurca and P. Frossard, "Media Streaming with Conservative Delay on Variable Rate Channels," in *Proceedings of IEEE international Conference on Multimedia and Expo (ICME, 2006)*, 2006.

[19] C. E. Perkins, E. M. Royer, "Ad hoc on-demand distance vector routing," *in Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications*, pp. 90-100, Feb 1999.

[20] G. D. Kondylis and G. J. Pottie, "Dynamic Channel Allocation Strategies for Wireless Packet Access," *IEEE VTC*, Amsterdam, Sep 1999.

[21] J. Chakareski and P. Frossard, "Rate-Distortion Optimized Distributed Packet Scheduling of Multiple Video Streams Over Shared Communication Resource," *IEEE Transactions on Multimedia*, vol. 8, no. 2, Apr, 2006.