

Distributed Resource Management in Multihop Cognitive Radio Networks for Delay-Sensitive Transmission

Hsien-Po Shiang and Mihaela van der Schaar, *Senior Member, IEEE*

Abstract—In this paper, we investigate the problem of multiuser resource management in multihop cognitive radio networks for delay-sensitive applications. Since tolerable delay does not allow propagating global information back and forth throughout the multihop network to a centralized decision maker, the source nodes and relays need to adapt their actions (transmission frequency channel and route selections) in a distributed manner, based on *local* network information. We propose a distributed resource-management algorithm that allows network nodes to exchange information and that explicitly considers the delays and cost of exchanging the network information over multihop cognitive radio networks. In this paper, the term “cognitive” refers to both the capability of the network nodes to achieve large spectral efficiencies by dynamically exploiting available frequency channels and their ability to learn the “environment” (the actions of interfering nodes) based on the designed information exchange. Note that the node competition is due to the mutual interference of neighboring nodes using the same frequency channel. Based on this, we adopt a multiagent-learning approach, i.e., adaptive fictitious play, which uses the available interference information. We also discuss the tradeoff between the cost of the required information exchange and the learning efficiency. The results show that our distributed resource-management approach improves the peak signal-to-noise ratio (PSNR) of multiple video streams by more than 3 dB, as opposed to the state-of-the-art dynamic frequency channel/route selection approaches without learning capability, when the network resources are limited.

Index Terms—Cognitive radio networks, delay-sensitive applications, distributed resource management, multiagent learning, multihop wireless networks.

I. INTRODUCTION

THE DEMAND for wireless spectra has increased and will rapidly keep increasing 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, in 2002, the Federal Communications Commission [1] suggested 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], which enable wireless users to *sense* and *learn* the surrounding environment and correspondingly *adapt* their transmission strategies, were proposed.

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). The second challenge is how to manage the available spectrum resources among the license-exempt users to satisfy their quality-of-service (QoS) requirements while limiting the interference to the primary licensees. In this paper, we focus on the second problem, i.e., 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]–[9]. In this paper, we focus on the resource management problem in the more general setting of multihop cognitive radio networks. A key advantage of such flexible multihop infrastructures is that the same infrastructure can be reused 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 multihop 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 or spectrum opportunities [2], [6]), as in the single-hop case but the routes through the various relay nodes to the destination nodes as well. Second, the transmission strategies will need to be adapted at not only the source nodes but also the relay nodes. In cognitive radio networks, network nodes are generally capable of sensing the spectrum, modeling the behavior of the primary users (PUs), and thereby identifying the available spectrum holes. In multihop cognitive radio networks, the network nodes will also need to model the behavior of the other neighbor nodes [i.e., other secondary users (SUs)] to successfully optimize the routing decisions. In other words, network relays (NRs) also require a learning capability in the multihop cognitive radio network. Third, to learn and efficiently adapt their decisions over time, the wireless nodes need to possess accurate (timely) information about the channel conditions, interference patterns, and other node-transmission strategies. However, in a distributed setting such as a multihop cognitive radio network, the information is decentralized; thus,

Manuscript received July 22, 2007; revised January 18, 2008, March 11, 2008, and March 17, 2008. First published May 14, 2008; current version published February 17, 2009. This work was supported by grant from the Office of Naval Research and by National Science Foundation CAREER Grant CCF 0541867. The review of this paper was coordinated by Dr. E. Hossein.

The authors are with the Department of Electrical Engineering, University of California, Los Angeles, Los Angeles, CA 90095 USA (e-mail: hpshiang@ee.ucla.edu; mihaela@ee.ucla.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TVT.2008.925308

there is a certain delay associated with gathering the necessary information from the various network nodes. Hence, an effective solution for multihop cognitive radio networks will need to trade off the “value” of having information about other nodes with the transmission overheads associated with gathering this information in a timely fashion across different hops in terms of the utility impact.

In this paper, we aim at learning the behaviors of interacting cognitive radio nodes that use a simple interference graph (similar to the spectrum holes used in [6] and [7]) to sequentially adjust and optimize their transmission strategies. We apply a multiagent learning algorithm, i.e., the fictitious play (FP) [14], to model the behavior of neighbor nodes based on the information exchange among the network nodes. We 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 multihop wireless networks, a centralized resource-management solution for these delay-constrained applications is not practical [13] 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 exponentially grow 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 distributed resource-management solution, which explicitly considers the availability of information, transmission overheads and incurred delays, and the value of this information in terms of the utility impact, is necessary.

This paper is organized as follows: In Section II, we discuss the related works and the contributions of this paper. Section III provides the multihop cognitive radio network settings and strategies. Section IV gives the problem formulation of the distributed resource management for delay-sensitive transmission in such networks. In Section V, we determine how to quantify the rewards and costs associated with various information exchanges in the multihop cognitive radio networks. In Section VI, we propose our distributed resource management algorithms with the information exchange and introduce the adopted multiagent learning approach, i.e., adaptive FP (AFP), in the proposed algorithms. Simulation results are presented in Section VII. Finally, Section VIII concludes this paper.

II. RELATED WORK

Distributed dynamic spectrum allocation is an important issue in cognitive radio networks. Various approaches have been proposed in recent years. In [7], decentralized cognitive medium-access control (MAC) protocols were proposed based on the theory of the partially observable Markov decision process, where an SU is able to model the PUs through Markovian state transition probabilities. In [8], the authors investigated a game-theoretic spectrum-sharing approach, where

the PUs are willing to share the spectrum and provide a determined pricing function to the SUs. In [9], a no-regret learning approach was proposed for dynamic spectrum access in cognitive radio networks. However, these studies focused on dynamic spectrum management for the single-hop network case.

Exploiting frequency diversity in wireless multihop networks has attracted enormous interests in recent years. In [10], Lee and Leung proposed distributed allocation scheme of subcarriers and power levels in wireless mesh networks that are based on orthogonal frequency-division multiple access. They proposed a fair scheduling scheme that hierarchically decouples the subcarrier and power-allocation problem based on the limited local information that is available at each node. In [11], Wu *et al.* focused on the distributed channel and routing assignment in heterogeneous multiradio multichannel multihop wireless networks. The proposed protocol coordinates the channel and route selection at each node, based on the information exchanged among two-hop neighbor nodes. However, these studies are not suitable for cognitive radio networks since they ignore the dynamic nature of spectrum opportunities and users (network nodes) need to estimate the behavior of the PUs for coexistence. To the best of our knowledge, the dynamic resource management problem in multihop cognitive radio networks has not been addressed in the literature.

In summary, this paper makes three contributions.

- 1) a dynamic resource management scheme in multihop cognitive radio network settings based on periodic information exchange among network nodes. Our approach allows each network node (SUs and relays) to exchange their spectrum opportunity information and select the optimal channel and next relay to transmit delay-sensitive packets.
- 2) We investigate the impact of the information exchange collected from various hops on the performance of the distributed resource management scheme. We introduce the notion of an “information cell” to explicitly identify the network nodes that can convey timely information. Importantly, we investigate the case where the information cell does not cover all the interfering neighbor nodes in the interference graph.
- 3) The proposed dynamic resource-management algorithm applies FP [14], which allows various nodes to learn their spectrum opportunity from the information exchange and adapt their transmission strategies autonomously in a distributed manner. Moreover, we discuss the tradeoffs between the cost of the required information exchange and the learning efficiency of the multiagent learning approach in terms of the utility impact.

III. MULTIHOP COGNITIVE RADIO NETWORKS: SETTINGS AND STRATEGIES

A. Multihop Cognitive Radio Network Specification

In this paper, we assume that a multihop cognitive radio network involves three network entities.

- 1) PUs: licensed users that will be guaranteed an interference-free environment [2], [4];

- 2) SUs: autonomous wireless stations that perform channel sensing and access the existing spectrum holes for their applications;
- 3) NRs: Autonomous wireless nodes that also perform channel sensing and access the spectrum holes for relaying applications. Note that multiple applications can use the same NR using different frequency channels.

We consider a multihop cognitive radio network, which is characterized by a general topology graph $\mathcal{G}(\mathbf{M}, \mathbf{N}, \mathbf{E})$ that has a set of PUs $\mathbf{M} = \{m_1, \dots, m_M\}$, 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 an SU (as a source or a destination node) or an NR.

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 frequency channels. To avoid interference to the PUs, 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; hence, \mathbf{F}_n may change over time, depending on whether the PUs are transmitting in their assigned frequency channels.

The network resource for a network node $n \in \mathbf{N}$ of the multihop 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 network node n as follows:

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

Whether or not 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. We will discuss the interference from other users (including the PUs) in Section III-C.

B. Source Traffic Characteristics

Let V_i denote the delay-sensitive application of the i th SU. Assume that 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 PUs. The rest of the classes C_k ($k > 1$) can be characterized by three components.

- 1) λ_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., real-time protocol header) of each packet.
- 2) 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.
- 3) L_k : the average packet length in class C_k .

A variety of delay-sensitive applications can use the cognitive radio setup discussed in this paper. Multimedia transmission such as video streaming or video conferencing can be examples of such applications [13]. 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. Interference Characterization

Recall that the highest priority class C_1 is always reserved in each frequency channel for the traffic of the PUs. The traffic of the SUs can be categorized into $K - 1$ priority classes (C_2, \dots, C_K) for accessing frequency channels. The traffic priority determines its ability of accessing the frequency channel. The PUs 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 to the SUs in the considered multihop cognitive radio network.

1) *Interference From PUs*: In practical cognitive networks, even though PUs have the highest priority, SUs will cause some level of interference to the PUs due to their imperfect awareness (sensing) of the PUs. The PUs' interference depends on the location of the M PUs. We rely on methods such as those in [5] that consider the power and location of the SUs to ensure that the SUs do not exceed some critical interference level to the PUs. We also assume that the spectrum opportunity map is available to the SUs, as in [6] and [9]. Since the PUs will block all the neighbor links using its frequency channel, a network node n will sense the channel and obtain the spectrum opportunity matrix (SOM) of the PUs, i.e., as in (2), as shown at the bottom of the page.

$$\mathbf{Z}_n = [Z_{ij}] \in \{0, 1\}^{L \times M}, \quad \text{with} \\ Z_{ij} = \begin{cases} 1, & \text{if the PU is occupying frequency channel } f_j \\ & \text{and link } e_i \text{ can interfere with the PU} \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

2) *Interference From Competing SUs*: We define $\mathbf{I}_k = [I_{ij}] \in \{0, 1\}^{L \times M}$ as the interference matrix (IM) for the traffic in priority class C_k ($k \geq 2$), i.e.,

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

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 (as in [9]). The interference graph is defined as the corresponding links that are interfered by the transmission of class C_k traffic.¹ The IM can be computed by the information exchange among the neighbor nodes.

The available resource matrix can be masked out by the SOM and IM of the higher priority classes, i.e., $\mathbf{R}_{nk}^{(I)} = \mathbf{R}_n \otimes \bar{\mathbf{I}}_{k-1} \otimes \cdots \otimes \bar{\mathbf{Z}}_n$, where the notation \otimes represents elementwise multiplication of the matrices and $\bar{\mathbf{I}}$ 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 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 multihop cognitive radio network.

D. Nodes' Actions

We define the action of network node n to relay the delay-sensitive application V_i as $A_n = (e \in \mathbf{E}_n, f \in \mathbf{F}_n)$. We assume that an NR 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 network node n as $s_n = [s_A | A = (e \in \mathbf{E}_n, f \in \mathbf{F}_n)]$, where s_A represents the probability that 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, i.e.,

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

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 to compute the effective transmission time (ETT) [17] over the transmission links.

Each network node n computes the ETT $ETT_{nk}(e, f)$ given by

$$ETT_{nk}(e, f) = \frac{L_k}{T_n(e, f) \times (1 - p_n(e, f))} \quad (5)$$

¹In a wireless environment, the transmission of *neighbor links* can interfere with each other and significantly impact their effective transmission time. Hence, the action of a node can impact and be impacted by the action of the other relay nodes. To coordinate these neighboring nodes, we construct the IM with binary "1" and "0."

with $e \in \mathbf{E}_n$, $f \in \mathbf{F}_n$ for transmitting delay-sensitive applications in priority class C_k . $T_n(e, f)$ and $p_n(e, f)$ represent the transmission rate and the packet error rate of network node n using frequency channel f over link e , respectively. $T_n(e, f)$ and $p_n(e, f)$ can be estimated by the MAC/PHY layer link adaptation [18]. Specifically, we assume that the channel condition of each link-frequency channel pair can be modeled using a continuous-time Markov chain [16] 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 ν_i (rate of transition at state i in transitions per second). We assume that the maximum transition rate² of the network is ν and that the variation of the channel conditions in a time interval $\tau \leq 1/\nu$ is regarded negligible.

Define the action vector $\mathbf{A}_i = [A_n | n \in \sigma_i]$ as the vector of the actions of all the NR nodes for transmitting V_i . Assume that the i th delay-sensitive application V_i is 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 $\sigma_i = \{\sigma_{ij} | j = 1, \dots, q_i\}$, where σ_{ij} is the route of the j th packet in V_i . A route σ_{ij} is a set of link-frequency pairs that the packets flow through, i.e.,

$$\sigma_{ij} = \{(e, f) | \text{the } j\text{th packet of } V_i \text{ flows}$$

through link e using frequency channel $f\}$. (6)

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 NR node receiving a packet can decide to where to relay the packet and, using which frequency channel, 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 needs to obtain the delay information from other nodes according to the actions taken by the relay nodes, i.e.,

$$d_{ij}(\sigma_{ij}(\mathbf{A}_i)) = \sum_{n \in \sigma_{ij}} ETT_{nk}(\mathbf{A}_i), \quad \text{for } j \in C_k. \quad (7)$$

IV. RESOURCE MANAGEMENT PROBLEM FORMULATION OVER MULTIHOP 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.

1) *Centralized Problem Formulation With Global Information Available at the Sources*: If we assume that global information³ \mathcal{G}_i is available to source node n_i^s for the delay-sensitive application V_i , route $\sigma_{ij}(\mathbf{A}_i, \mathcal{G}_i)$ can be determined for each packet j of V_i . The centralized optimization can be performed at

²In the case in which some of the channel conditions severely change in the network, a threshold ν_{th} can be set by protocols to avoid these fast-changing nodes, and ν is, hence, selected as the maximum transition rate below this threshold value.

³The word "global information" means the information gathered from every node throughout the network. We discuss the required information in Section V.

every source node to maximize utility u_i . Hence, for application V_i , we have

$$\begin{aligned} \mathbf{A}_i^{\text{opt}} &= \arg \max u_i(\mathbf{A}_i, \mathcal{G}_i) \\ \text{s.t. } & A \in \hat{\mathbf{A}}_n \text{ for all } A \in \mathbf{A}_i, \quad \text{where} \\ u_i(\mathbf{A}_i, \mathcal{G}_i) &= \sum_{j=1}^{q_i} \lambda_{ij} \cdot \text{Prob}\{d_{ij}(\sigma_{ij}(\mathbf{A}_i, \mathcal{G}_i)) \leq D_{ij}\} \\ D_{ij} &= D_k \quad \text{and} \quad \lambda_{ij} = \lambda_k, \quad \text{if } j \in C_k. \end{aligned} \quad (8)$$

However, due to the limited wireless network resource, the end-to-end delay constraint $d_{ij}(\sigma_{ij}(\mathbf{A}_i, \mathcal{G}_i)) \leq D_k$ can make the optimization solution infeasible. Hence, suboptimal greedy algorithms that sequentially perform optimizations from the highest priority class to the lowest priority class are commonly adopted [13], [23]. Specifically, for class C_k , the following optimization is considered:

$$\begin{aligned} \mathbf{A}_{ik}^{\text{opt}} &= \arg \min \sum_{j \in C_k} d_{ij}(\sigma_{ij}(\mathbf{A}_{ik}, \mathcal{G}_i)) \\ \text{s.t. } & d_{ij}(\sigma_{ij}(\mathbf{A}_{ik}, \mathcal{G}_i)) \leq D_k \\ & A \in \hat{\mathbf{A}}_n(k) \text{ for all } A \in \mathbf{A}_{ik} \end{aligned} \quad (9)$$

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

Due to the informationally decentralized nature of the multi-hop wireless networks, the centralized solution is not practical for the multiuser delay-sensitive applications, as the tolerable delay does not allow propagating 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 timely be relayed to a source node. For instance, 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 exponentially grows with the number of classes and nodes in the network. 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. Hence, a ‘‘decomposition’’ of the optimization problem into distributed strategic adaptation based on the available local information is necessary.

2) *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 multihop wireless infrastructure. Note that, at each node n , the end-to-end delay for sending a packet $j \in C_k$ in (9) can be decomposed as

$$d_{ij}(\sigma_{ij}) = d_n^P(\sigma_{ij}) + E[\tilde{d}_n(k, \sigma_{ij})] \quad (10)$$

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 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 impact factor λ_k . The information about λ_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 (9)]. Since, at node n , the value of $d_n^P(\sigma_{ij})$ is fixed, the optimization problem at node n becomes

$$\begin{aligned} A_n^{\text{opt}} &= \arg \min E[\tilde{d}_n(k, \sigma_{ij}(A_n, \mathcal{L}_n))] \\ \text{s.t. } & E[\tilde{d}_n(k, \sigma_{ij}(A_n, \mathcal{L}_n))] \leq D_k - d_n^P(\sigma_{ij}) - \rho \\ & j \in C_k \quad A_n \in \hat{\mathbf{A}}_n \end{aligned} \quad (11)$$

where $E[\tilde{d}_n(k, \sigma_{ij}(A_n, \mathcal{L}_n))]$ represents the expected delay from relay node n to the destination of the packets in class C_k . ρ represents a guard interval such that 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 [20]). To estimate the expected delay $E[\tilde{d}_n(k, \sigma_{ij}(A_n, \mathcal{L}_n))]$ in (11), 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 [16]. We assume that each node n maintains and updates a delay vector $\mathbf{d}_n = [E[\tilde{d}_n(2)], \dots, E[\tilde{d}_n(K)]]$ (note that the first priority class is reserved for the PUs) with elements for each priority class. We will discuss the minimum-delay routing/channel-selecting algorithm in Section VI. Compared with the centralized approach in (8), the distributed resource management in (11) can adapt better to the dynamic wireless environment by periodically gathering local information. Next, we discuss the distributed resource management with information constraints in more detail.

V. DISTRIBUTED RESOURCE MANAGEMENT WITH INFORMATION CONSTRAINTS

In this paper, we assume that the required local information \mathcal{L}_n is exchanged using a designated coordination control channel similar to [12]. The transmission is time slotted, and the time slot structure of a node is provided in Fig. 1. We denote the time slot duration as t_I . Action A_n is selected at each node, during each time slot, after the coordination interval (which includes the channel sensing for SOM and the information exchange for IM). In addition to the SOM and IM, the information required in the coordination interval should also include delay vectors \mathbf{d}_n and the control messages for request-to-send (RTS)/clear-to-send (CTS) coordination [7], [11]. The goal of the coordination interval at each time slot is to provide the feasible action set $\hat{\mathbf{A}}_n$ for the channel access and the relay selection of the packet transmission. We denote the coordination interval at network node n as $d_I(\mathcal{L}_n)$.

A. Benefit of Acquiring Information and Information Cell Constraints

For network node n , the local information \mathcal{L}_n gathered from different network nodes has different impacts on decreasing the objective function $E[\tilde{d}_n(k, \sigma_{ij}(A_n, \mathcal{L}_n))]$ in (11). Let $\mathcal{I}_n(x) = \{\mathbf{I}_k(n_x, A_{n_x}), A_{n_x}, \mathbf{d}_{n_x} | n_x \in \mathbf{N}_x^n\}$ denote the set of local information gathered from the neighbor nodes, which

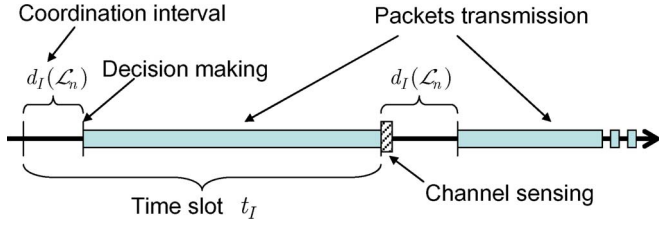


Fig. 1. Transmission time line at node n .

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 $\mathcal{L}_n(x) = \{\mathcal{I}_n(l) | l = 1, \dots, x\}$ as the local information gathered from all of these neighbor nodes. Given local information $\mathcal{L}_n(x)$, we define the optimal expected delay as $K_n(k, x) = E[\tilde{d}_n(k, \sigma_{ij}(A_n^{\text{opt}}, \mathcal{L}_n(x)))]$. The larger x will have a smaller expected delay $K_n(k, x)$. The benefit (reward) of information $\mathcal{I}_n(x)$ for class C_k traffic is denoted as $J_n(k, \mathcal{I}_n(x))$. In a static network case, $J_n(k, \mathcal{I}_n(x))$ is defined as

$$J_n(k, \mathcal{I}_n(x)) \triangleq K_n(k, x-1) - K_n(k, x), \quad \text{if } x > 1. \quad (12)$$

We define $J_n(k, \mathcal{I}_n(1)) = K_n(k, 1)$ since $\mathcal{L}_n(1) = \mathcal{I}_n(1)$. The reward of information $J_n(k, \mathcal{I}_n(x))$ can be regarded as the benefit (decrease in the expected delay) in terms of expected delay $E[\tilde{d}_n(k, \sigma_{ij})]$ if information $\mathcal{I}_n(x)$ is received by node n . Note that the optimal expected delay $K_n(k, x)$, given information $\mathcal{L}_n(x)$, is given by

$$K_n(k, x) = K_n(k, 1) - \sum_{l=2}^x J_n(k, \mathcal{I}_n(l)). \quad (13)$$

Equation (13) states that the optimal expected delay is a decreasing function of x , meaning that smaller expected delays can be achieved as more information is gathered. The improvement is quantified by the reward of the information $J_n(k, \mathcal{I}_n(l))$. Here, we ignore the cost of exchanging such information, which will be defined in the next section. Fig. 2 shows a simple illustrative example of reward of information at node n , which is five hops away from the destination node of class C_k traffic. The more information $\mathcal{I}_n(x)$ that is available from nodes that are x hops away, the smaller the optimal expected delay $K_n(k, x)$ that can be obtained.

Let $\mathbf{J}_n(k) = [J_n(k, \mathcal{I}_n(x))]$, for $1 \leq x \leq H_n$ denote the reward vector from one-hop information to H_n -hop information, where $H_n = \max\{H_n^I, H_n^d\}$. H_n^d represents the shortest hop counts from node n to the destination node of class C_k traffic, and H_n^I represents the interference range in terms of hop counts for node n . We assume that reward vector $\mathbf{J}_n(k)$ is obtained when the network is first deployed and only infrequently updated, when SUs join or leave the network. Note that all the elements in $\mathbf{J}_n(k)$ are nonnegative, i.e., $J_n(k, \mathcal{I}_n(x)) \geq 0$, for $1 \leq x \leq H_n$, due to the fact that knowing additional information cannot increase the expected delay $E[\tilde{d}_n(k, \sigma_{ij})]$ in a static network. However, if we consider the propagation delay of such information exchange across the network in the dynamic network, the dynamic reward of information $J_n^d(k, \mathcal{I}_n(x))$ decreases as hop count x increases. When the information of the further nodes reaches

decision node n , the information is more likely to be out of date (i.e., the information cannot reflect the exact network situation in a dynamic setting since the network conditions and traffic characteristics are time varying). Once the information is out of date, $J_n^d(k, \mathcal{I}_n(x)) = 0$, i.e., there is no benefit from gathering information that is out of date. Note that, in a dynamic network, once $J_n^d(k, \mathcal{I}_n(x)) = 0$, $J_n^d(k, \mathcal{I}_n(x')) = 0$ for $x \leq x' \leq H_n$.

Therefore, in the dynamic network, we define the information horizon $h(k, \nu)$ such that

$$h_n(k, \nu) \triangleq \arg \max_x \quad \text{s.t. } J_n^d(k, \mathcal{I}_n(x)) > \phi(k, \nu), \quad 1 \leq x \leq H_n \quad (14)$$

where $\phi(k, \nu) \geq 0$ represents a minimum delay variation specified by the application, which determines the minimum benefit of receiving local information for class C_k traffic. In fact, $h_n(k, \nu)$ depends on the variation speed ν of the wireless network condition (i.e., see Section III-D). In a dynamic network with higher variation speeds ν (e.g., with high mobility), a higher threshold $\phi(k, \nu)$ is needed to guarantee that information $\mathcal{I}_n(x)$ is still valuable, and it should be exchanged. This results in a smaller information horizon $h_n(k, \nu)$. We illustrate this mobility issue in Section VII. Note that the information horizon $h_n(k, \nu)$ also varies for different classes of traffic. Since higher priority class traffic has more network resources than lower priority class traffic, the threshold value $\phi(k, \nu) \leq \phi(k', \nu)$, if $k < k'$; therefore, $h_n(k, \nu) \geq h_n(k', \nu)$, if $k < k'$. In other words, the information horizon $h_n(k, \nu)$ of a higher priority class C_k is larger than the information horizon $h_n(k', \nu)$ of a lower priority class $C_{k'}$.

For simplicity, we assume in this paper that the information horizon is only a function of network variation speed ν , i.e., $h_n(k, \nu) = h(\nu)$. Information horizon $h(\nu)$ is determined for the most important class among the SUs in the network. This definition of information horizon $h(\nu)$ is aligned with [13], in which $h(\nu)$ is defined as the maximum number of hops that the information can be conveyed in τ , such that the network is considered unchanged. (Recall that any network changes within interval $\tau(\nu) \leq 1/\nu$ can be regarded negligible.)

Based on this information horizon $h(\nu)$, we assume that the network nodes within the $h(\nu)$ hops form an *information cell*. Only the local information $\mathcal{L}_n(h)$ within the information cell is useful to node n since the reward of information is zero, i.e., $J_n(k, \mathcal{I}_n(x)) = 0 \forall x > h(\nu)$. Recall that the neighbor nodes of node n are defined as the nodes that can interfere or can be interfered by node n (within H_n^I hops), which may not align with the range of the information cell (within $h(\nu)$ hops). If all neighbor nodes are within the h -hop information cell, all necessary information are timely conveyed to node n . Otherwise, the neighbor nodes that are too far away cannot convey the interference information to node n in time. We refer to this problem as the “*information exchange mismatch*” problem.

B. Cost of Information Exchange

In the previous section, 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 in the expected delay) due to this information

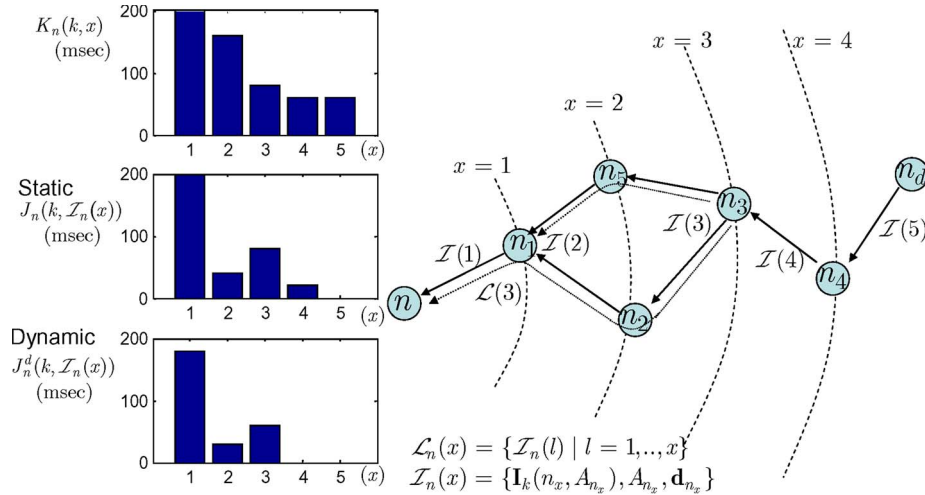


Fig. 2. Example of the static reward of information $J_n(k, \mathcal{I}_n(x))$, the dynamic reward of information $J_n^d(k, \mathcal{I}_n(x))$, and optimal expected delay $K_n(k, x)$ (where information horizon $h_n(k, \nu) = 3$, average packet length $L_k = 1000$ B, and average transmission rate $T = 6$ Mb/s over the multihop network).

exchange. The duration of time slot $t_I(\nu)$ is also the interval between the repeated information exchanges in the network. We define c time slots in τ seconds, i.e.,

$$t_I(\nu) = \frac{\tau(\nu)}{c} \quad (15)$$

where c defines the frequency of the decision making and 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 with τ . Hence, the network changes in t_I are also negligible.

Note that, even though the information exchange is implemented in a designated coordination channel [12], a network node with a single antenna cannot transmit both the data and the control signals at the same time. This information exchange time overhead decreases the effective transmission rate at node n using line e and frequency channel f , i.e.,

$$T'_n(e, f) = \frac{t_I(\nu) - d_I(\mathcal{L}_n(h))}{t_I(\nu)} \times T_n(e, f). \quad (16)$$

Hence, the ETT at a node n using link e and frequency channel f to transmit a packet in class C_k becomes

$$ETT'_{nk}(e, f) = \frac{t_I(\nu)}{t_I(\nu) - d_I(\mathcal{L}_n(h))} \times ETT_{nk}(e, f). \quad (17)$$

In conclusion, the increase in the ETT degrades the performance of the delay-sensitive applications. The degradation depends on the content of the local information exchange $\mathcal{L}_n(h)$ and network variation speed ν . Hence, the benefit $J_n^d(k, \mathcal{I}_n(x))$ in (14) will decrease due to this cost of the information. Hence, we denote the value of information with this cost consideration as $J_n^c(k, \mathcal{I}_n(x))$, i.e.,

$$\begin{aligned} J_n^c(k, \mathcal{I}_n(x)) &= K'_n(k, x-1) - K'_n(k, x) \\ &= K_n(k, x-1) \times \frac{t_I(\nu)}{t_I(\nu) - d_I(\mathcal{L}_n(x-1))} \\ &\quad - K_n(k, x) \times \frac{t_I(\nu)}{t_I(\nu) - d_I(\mathcal{L}_n(x))}. \end{aligned} \quad (18)$$

In addition, the optimal information horizon $h_n(k, \nu)$ in (14) also decreases due to the cost. Next, we discuss the proposed distributed resource management algorithm based on the information exchanges and learning capabilities to tackle the optimization problem in (11).

VI. DISTRIBUTED RESOURCE MANAGEMENT ALGORITHMS

Fig. 3 shows a system diagram of the proposed distributed resource management. First, a packet $j \in C_k$ is selected from the application scheduler at node n , based on the impact factor λ_k of the packet, and an action A_n is taken for that packet. The application layer information including C_k , L_k , and D_k is conveyed to the network layer for this action decision. Network conditions $T_n(e, f)$ and $p_n(e, f)$ are then conveyed from the MAC/PHY layer for computing the ETT values using (5).

In addition to $T_n(e, f)$ and $p_n(e, f)$, 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 $\mathcal{L}_n(h) = \{\mathcal{I}_n(l) | l = 1, \dots, h\}$. We use the notation $-n(h)$ to represent the set of neighbor nodes of 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, node n knows the estimated delay $\mathbf{d}_{-n(h)}$ from its neighbor nodes to the destinations, such as the actions $A_{-n(h)}$ of its neighbor nodes and their IM $\mathbf{I}_k(-n(h), A_{-n(h)})$. Based on the delay information from the neighbor nodes $\mathbf{d}_{-n(h)}$, a network node can update its own estimated delay to the various destinations and determine the minimum-delay action based on the Bellman–Ford algorithm [16].

We separate the distributed resource management into two blocks at node n , as in Fig. 3: 1) the information exchange interface block that regularly collects required local information and 2) the route/channel selection block for determining the optimal action. We now discuss the role of the exchanged information and the two algorithms implemented in these blocks, respectively.

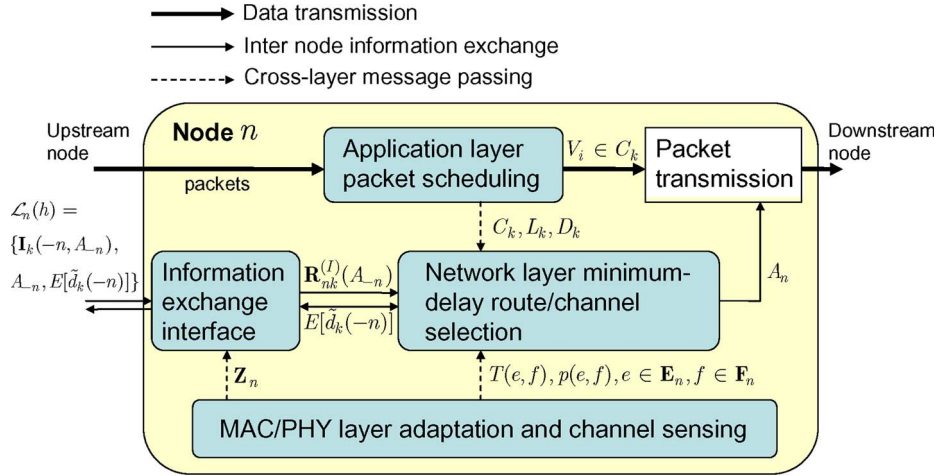


Fig. 3. System diagram of the proposed distributed resource management.

A. Distributed Resource Management Algorithms

The next algorithm is performed at network node n at the information exchange interface in Fig. 3.

Algorithm 1. Periodic Information Exchange Algorithm:

- Step 1) Collect the required information. Node n first collects the required information SOM \mathbf{Z} from channel sensing, and $\mathcal{L}_n(h) = \{\mathbf{I}_k(-n(h), A_{-n(h)}), A_{-n(h)}, \mathbf{d}_{-n(h)}\}$ 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 $\mathbf{s}(n') = [s_A(n') | A = (e \in \mathbf{E}_{n'}, f \in \mathbf{F}_{n'})]$, $n' \in -n(h)$, where $s_A(n')$ represents the probability (strategy) of selecting an action A by node n' , which will be discussed in the next section.
- Step 3) Estimate the resource matrix. From the SOM and the IM $\mathbf{I}_k(n', A_{n'})$ gathered from neighbor node n' , the resource matrix can be obtained for each class of traffic by $\mathbf{R}_{nk}^{(I)} = \mathbf{R}_n \otimes \bar{\mathbf{I}}_{k-1} \otimes \cdots \otimes \bar{\mathbf{Z}}_n$, which will be explained in Section VI-A in more detail. Then, the available resource $\mathbf{R}_{nk}^{(I)}(A_{-n})$ is provided to the network layer route/channel selection block stated in Algorithm 2.
- 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 IM $\mathbf{I}_k(n, A_n)$. Two types of interference model are considered in this paper when constructing the IM $\mathbf{I}_k(n, A_n)$ from (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 $\mathbf{I}_k(n, A_n)$ that is associated with the frequency channel are set to 1. Then, the IM is exchanged to the nodes within the predetermined information horizon h .

- 2) A network node cannot transmit and receive packets at the same time. In this case, if frequency channel $f \in \mathbf{F}_n$ is used, all the elements in the column of the IM $\mathbf{I}_k(n, A_n)$ 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 IM $\mathbf{I}_k(n, A_n)$ that is associated with node n are also set to 1, no matter what frequency channel it uses. Then, the IM is exchanged to the nodes within the predetermined information horizon h .
- Step 5) Broadcast the information $\{\mathbf{I}_k(n, A_n), A_n, \mathbf{d}_n\}$, and periodically repeat the algorithm in every $t_I(\nu)$ seconds.

The next algorithm is performed at network node n at the network layer minimum-delay route/channel selection block in Fig. 3.

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 node n is scheduled to be transmitted. Assume that packet $j \in C_k$ and the information of C_k, L_k , and $D_k - d_n^P$ are extracted or computed from the application layer.
- Step 2) Construct the feasible action set. Construct the feasible action set $\hat{\mathbf{A}}_n(k)$ from the resource matrix $\mathbf{R}_{nk}^{(I)}$ given by the information exchange interface for priority class C_k at node n [see (4)].
- Step 3) Estimate the channel condition. The transmission rate $T_n(e, f)$ and packet error rate $p_n(e, f)$ 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 [18].
- 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

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

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 neighbor node $n'(A_n)$. $ETT_{nk}(A_n)$ can be calculated based on L_k , $T_n(e, f)$, and $p_n(e, f)$ using (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. Determine the optimal action A_n^{opt} from the feasible action set $\hat{\mathbf{A}}_n(k)$, which is given by

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

Note that the feasible action set $\hat{\mathbf{A}}_n(k)$ in (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 the next section.

- Step 7) Send an RTS request. After determining the next relay and frequency channel, send an RTS request indicating the determined action information A_n^{opt} to the next relay.
- Step 8) Wait for the CTS response, and transmit the packets.
- Step 9) Update the delay and the current action information. After selecting the optimal action, update the estimated delay $E[\tilde{d}_n(k)]$ using an exponential moving average with a smoothing factor α , i.e.,

$$E[\tilde{d}_n(k)] = \alpha E[\tilde{d}_n(k)]^{\text{old}} + (1 - \alpha) E[\tilde{d}_n(k, A_n^{\text{opt}})] \quad (21)$$

and provide the updated delay vector $\mathbf{d}_n = [E[\tilde{d}_n(2)], \dots, E[\tilde{d}_n(K)]]$ to Algorithm 1 at the information exchange interface. In Fig. 4, we provide a block diagram of the proposed distributed resource management. For the blocks that are beyond the scope of this paper, we refer to [4] and [5] for channel sensing, [7] and [11] for RTS/CTS coordination, and [16] for the delay vectors.

B. Adapting Information Horizon Using AFP

We now provide a learning approach for the SUs to learn the feasible action set $\hat{\mathbf{A}}_n(k)$ in (20) for our distributed resource-management algorithms. 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. This motivates us to apply a well-known learning approach, i.e., FP [14], which is applied when the SUs are willing⁴

⁴If the action information is not provided by the other SUs, a node can learn its own strategy from its action payoffs, i.e., the estimated delay $E[d_n(k)]$. The learning approach refers to reinforcement learning (a model-free learning or a payoff-based learning).

to reveal their current action information; thereby, they are able to model the behaviors (strategies) of other SUs (a model-based learning). However, due to the information constraint discussed in the previous section, only the information from the neighbor nodes in the information cell is useful. Hence, we adapt the FP learning approach to our considered network setting.

Note that 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 $\mathbf{s}(n', t) = [s_A(n', t)|A = (e \in \mathbf{E}_{n'}, f \in \mathbf{F}_{n'})]$ for each of its neighbor nodes $n' \in -n(h)$ in the information cell. $s_A(n', t)$ represents the frequency selection strategy of node n' making action A at time t , which is obtained using

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

where $r_A(n', t)$ is the *propensity* [15] of node n' for taking action A at time t , which can be computed by

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

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

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

Fig. 5 shows how network variation speed ν affects the size of the information cell and, ultimately, the video performance. We will consider the mobility of the NRs to show this network variation impact in the next section.

$s_A(n', t)$ represents the probability that network node n' will choose an action A . Hence, the probability $s_A(n', t)$ for modeling node n' making an action A at time t will increase with the actual times that the action A is selected. Based on strategy $s_A(n', t)$, the AFP provides the estimated IM \mathbf{I}_k , and then, the feasible action set $\hat{\mathbf{A}}_n(k)$ can be computed.

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

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

Then, node n can estimate the IM \mathbf{I}_k for the traffic in class C_k , i.e.,

$$\mathbf{I}_k = \left[I_{ij} | I_{ij} = \begin{cases} 1, & \text{if } I_{ij}^e \geq \mu \\ 0, & \text{if } I_{ij}^e < \mu \end{cases} \right] \quad (26)$$

where μ represents a threshold value that determines whether 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 resource matrix $\mathbf{R}_{nk}^{(I)} = \mathbf{R}_n \otimes \bar{\mathbf{I}}_{k-1} \otimes \dots \otimes \bar{\mathbf{Z}}_n$ using (4). By learning the feasible action set $\hat{\mathbf{A}}_n(k)$, the best response actions are computed using (20).

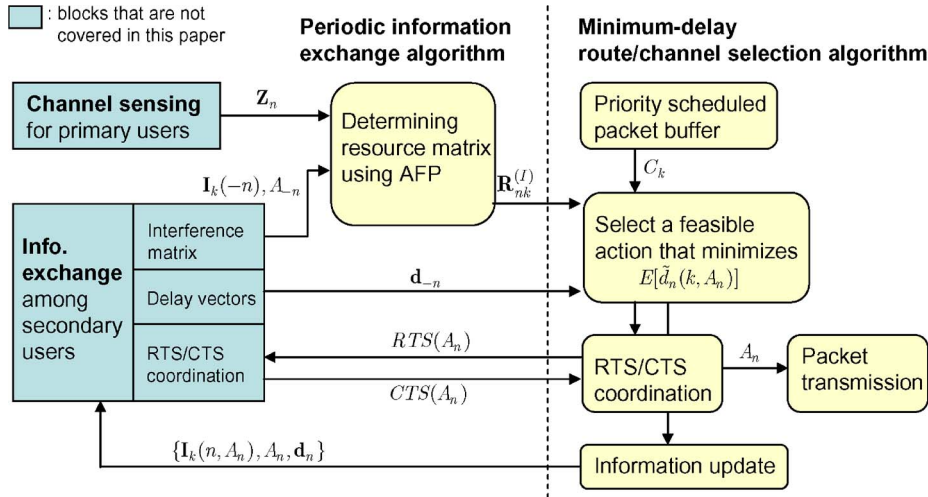


Fig. 4. Block diagram of the proposed distributed resource management at network node n .

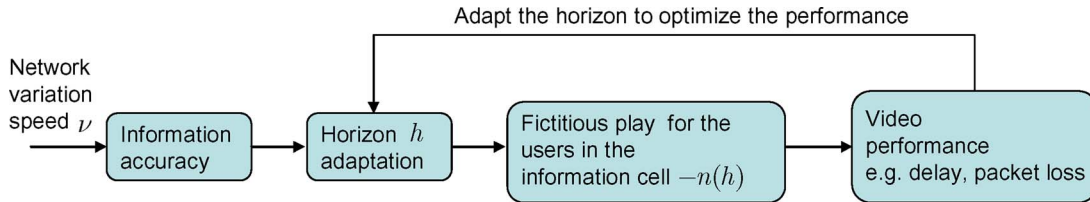


Fig. 5. Impact of the network variation on the FP and the video performance.

VII. SIMULATION RESULT

We simulate two video streaming applications that are transmitting videos V_1 “Coastguard” and V_2 “Mobile” (with 16 frames per group of pictures and 30-Hz frame rate in Common Intermediate Format) over the same multihop cognitive radio network. Each video sequence is divided into four priority classes ($K_i = 4, K = 9$) with average packet length $L_k = 1000$ B and delay deadline $D_k = 500$ ms. Although the first priority class C_1 is reserved for the PUs, let us first consider the case when there are no PUs, 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 Fig. 6 in a 100×100 meters region with $N = 15$ nodes and $L = 22$ links similar to the network settings in [19]. A link is established as long as the channel condition (described in this paper by the link signal-to-noise-plus-interference ratio) is acceptable within the transmission distance (approximately 36 m). Note that this transmission distance is not aligned with interference range H_n^I . Neighbor nodes that are beyond the transmission distance can still interfere with each other.

A. Reward and Cost of the Information Exchange

First, we simulate the impact of the information including the reward J_n^d [see (12)] and cost J_n^c [see (18)] from the expected delay $E[d_n]$ using the AFP in Section VII with different information horizons. Fig. 7 shows the resulting reward and the cost of information at different locations for streaming video V_1 (at nodes $n = 1, 7,$ and 13 on one of the routes of video V_1). The results show that a one-hop information cell is enough when

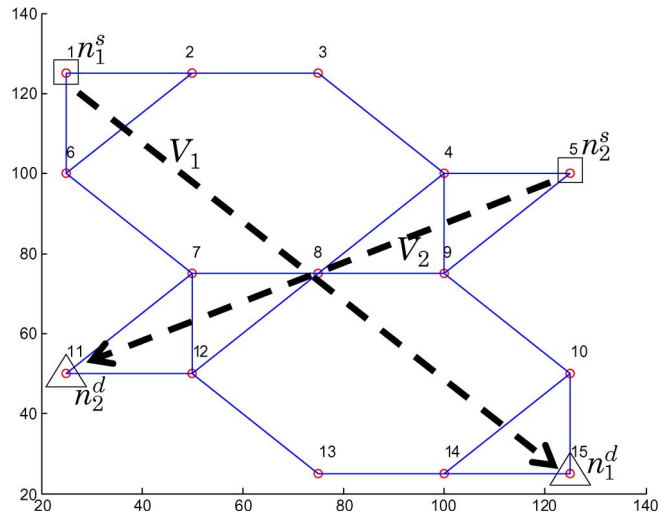


Fig. 6. Wireless network settings for the simulation of two video streams.

the interference range is 40 m since only the nodes that are one hop away can interfere with each other. If the interference range is 80 m, the information exchange mismatch problem (see Section V) occurs, and the appropriate information horizon for information exchange is then increased to 2.

B. Application Layer Performance With Different Information Horizons and Interference Ranges

We next compare the proposed dynamic resource management algorithm using AFP with two other resource management methods: 1) AODV [21] with load balancing over the two

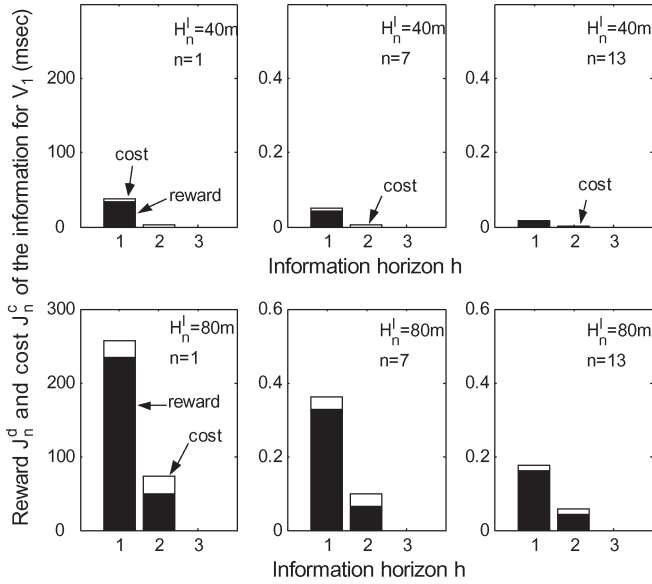


Fig. 7. Reward J_n^d and cost J_n^c of different information horizons at different nodes for video V_1 .

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

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 ($H_n^I = 80$ m)

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

available frequency channels (AODV/LB) and 2) the dynamic least interference channel selection (DCS) [22] extended to a network setting. Tables 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. In particular, when the interference range is large ($H_n^I = 80$ m), the proposed AFP approach significantly improves the video quality. (X represents a PSNR of below 26 dB, which is unacceptable for a viewer.)

For delay-sensitive applications, we measure the packet loss rate (i.e., the probability that the end-to-end delay exceeds the delay deadline) for different approaches in Fig. 8(a). The results of both applications are shown. The AODV represents the on-demand routing solution with only one frequency channel.

The AODV/LB approach randomly distributes packets over the two available frequency channels. The DCS approach with cognitive ability selects a better frequency channel based on the link measurements and, hence, improves the performance, as opposed to the AODV/LB. The AFP further improves the performance of both applications by learning the behaviors of the neighbor nodes. Interestingly, the benefit brought by the learning capability decreases as the network bandwidth increases. In other words, it is not worthy to be too intelligent in an environment with plenty of resources. Moreover, as shown in Fig. 8(b), the improvement of the two-hop information cell is limited when the interference range is 40 m. This is because the nodes that are two hops away have no impact on the current node, and their information is not valuable (i.e., it does not impact the utility).

C. Impact of the PUs

The simulation implies that the reward of information is also impacted by the existence of the PUs. Next, we consider the impact of the PUs, which always have higher priority to access the preassigned frequency channels than the network nodes in Fig. 6. Assume that frequency channel F_1 is occupied by the PUs with time fraction $\rho = 0\%$, 20%, 40%, 60%, and 80% around a certain congestion region (network nodes $n = 7, 11, 12$) in Fig. 6. Fig. 9 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 Mb/s, $b_n/c = 1$, and the interference range is 80 m.

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

D. Impact of Mobility

In this section, we consider the impact of mobility on the video performance. We adopt a well-known mobility model, the “random walk” [24], in which the relay nodes (SUs) shown in Fig. 6 randomly select a direction at each time slot and move at a fixed speed v . We simulate the speed v ranging from 0 to 1 m/s. We assume that there is no PU, i.e., $\rho = 0$. The average transmission rate is set to 8 Mb/s, $b_n/c = 1$, and the interference range is 80 m. Fig. 10 shows the packet loss rate as the mobility changes for different information horizons. The results show that the mobility degrades the performance

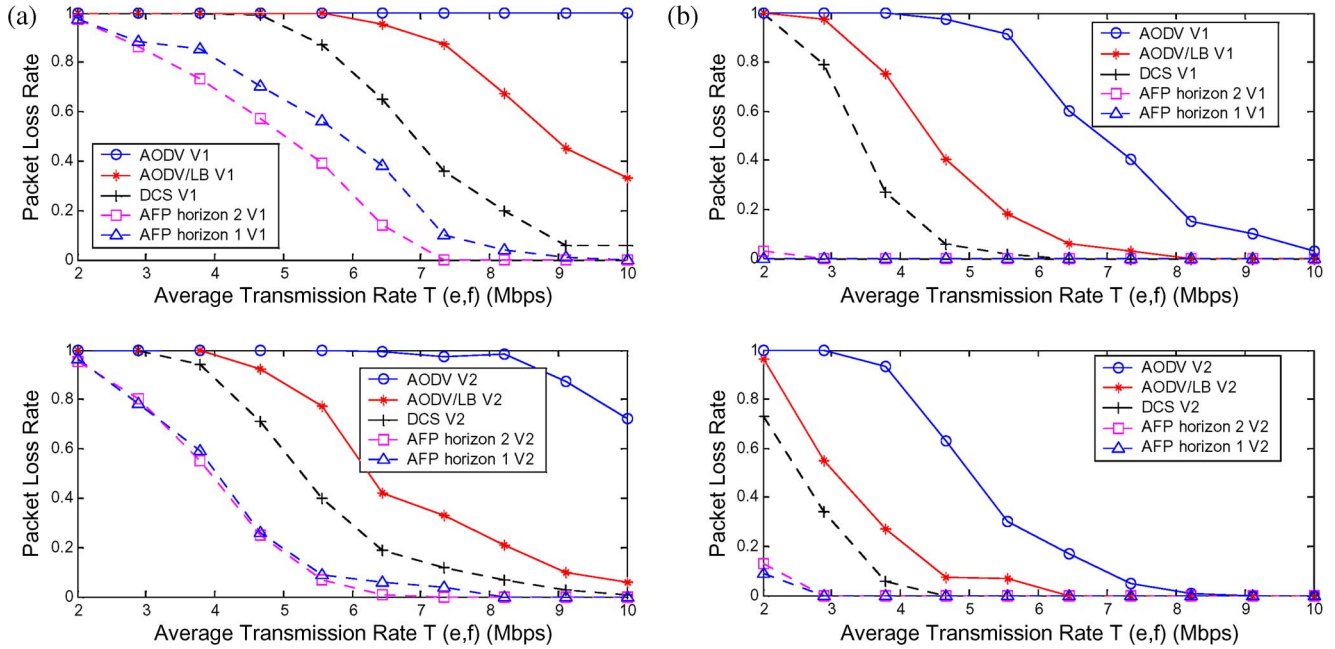


Fig. 8. (a) Packet loss rate versus the average transmission bandwidth using different approaches ($H_n^I = 80$ m). (b) Packet loss rate versus the average transmission bandwidth using different approaches ($H_n^I = 40$ m).

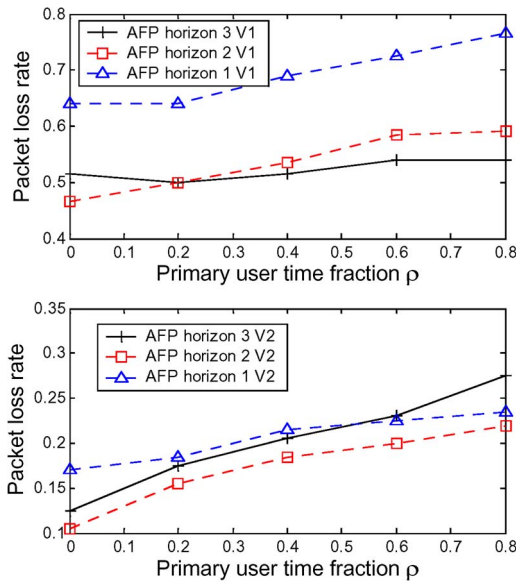


Fig. 9. Packet loss rate versus time fraction ρ of the PUs occupying frequency channel F_1 around network nodes $n = 7, 11,$ and 12 (average $T = 5.5$ Mb/s, $b_n/c = 1$, and $H_n^I = 80$ m).

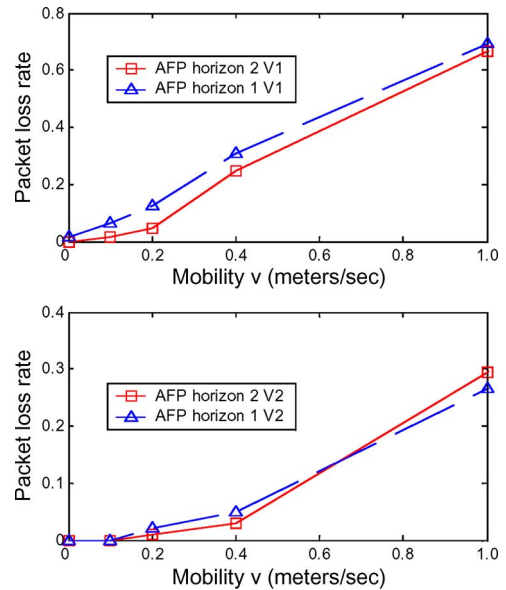


Fig. 10. Packet loss rate versus mobility v of the SUs (NRs) (average $T = 8$ Mb/s, $\rho = 0$, $b_n/c = 1$, and $H_n^I = 80$ m).

of both applications. When the mobility v is small, the AFP with information horizon $h = 2$ performs better than that with information horizon $h = 1$, as in the previous simulations with $H_n^I = 80$ m. However, for video V_2 , when the mobility exceeds 0.6 m/s, the best information horizon changes from $h = 2$ to $h = 1$. This is because the increased mobility will decrease the information accuracy; hence, the required information horizon also decreases. Note that, for video V_1 , the AFP with information horizon $h = 2$ still performs better than that with information horizon $h = 1$. This is because video V_1 has a longer route; thus, modeling more interfering neighbor nodes, using a larger information horizon, is still beneficial.

VIII. CONCLUSION

In this paper, we have shown that the distributed resource-management solution using AFP significantly improves the performance of delay-sensitive applications transmitted over a multihop cognitive radio network. We assume that the autonomous SUs are able to learn the spectrum opportunities based on the information exchange. The proposed approach can also be used to support QoS for general multiradio wireless networks, when there is no PU. This situation is also brought up in [4], when the SUs are competing in the unlicensed band (i.e., ISM band), where there is no PU. Importantly, based on the value of the obtained information (i.e., the impact

on decreasing the expected end-to-end delay), we define the information horizon in our AFP. 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.

The information horizon is assumed to be fixed in this paper for different priority classes over the whole wireless networks. However, our simulation results show that the benefit from various information horizons can be different for distinct applications with various delays and quality impacts, particularly when PUs are present in the network at different locations. Exploring what are optimal information horizons if the applications and network conditions are changing forms an interesting future research topic in the multihop cognitive radio networks.

REFERENCES

- [1] Fed. Commun. Comm. (FCC), *Spectrum Policy Task Force*, Nov. 15, 2002. ET Docket no. 02-135.
- [2] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
- [3] J. Mitola *et al.*, "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," *Comput. Netw.: Int. J. Comput. Telecommun. Netw.*, vol. 50, no. 13, pp. 2127–2159, Sep. 2006.
- [5] T. X. Brown, "An analysis of unlicensed device operation in licensed broadcast service bands," in *Proc. IEEE DySPAN*, Nov. 2005, pp. 11–29.
- [6] C. Cordeiro, K. Challapali, D. Birru, and S. N. Shankar, "IEEE 802.22: An introduction to the first wireless standard based on cognitive radios," *J. Commun.*, vol. 1, no. 1, pp. 38–47, Apr. 2006.
- [7] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE J. Sel. Areas Commun.—Special Issue Adaptive, Spectrum Agile Cognitive Wireless Networks*, vol. 25, no. 3, pp. 589–600, Apr. 2007.
- [8] D. Niyato and E. Hossain, "A game-theoretic approach to competitive spectrum sharing in cognitive radio networks," in *Proc. IEEE WCNC*, Hong Kong, Mar. 11–15, 2007, pp. 16–20.
- [9] Z. Han, C. Pandana, and K. J. R. Liu, "Distributive opportunistic spectrum access for cognitive radio using correlated equilibrium and no-regret learning," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2007, pp. 11–15.
- [10] K.-D. Lee and V. C. M. Leung, "Fair allocation of subcarrier and power in an OFDMA wireless mesh network," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 11, pp. 2051–2060, Nov. 2006.
- [11] H. Wu, F. Yang, K. Tan, J. Chen, Q. Zhang, and Z. Zhang, "Distributed channel assignment and routing in multiradio multichannel multihop wireless networks," *IEEE J. Sel. Areas Commun.—Special Issue Multi-Hop Wireless Mesh Networks*, vol. 24, no. 11, pp. 1972–1983, Nov. 2006.
- [12] V. Brik, E. Rozner, S. Banarjee, and P. Bahl, "DSAP: A protocol for coordinated spectrum access," in *Proc. IEEE DySPAN*, Nov. 2005, pp. 611–614.
- [13] H.-P. Shiang and M. van der Schaar, "Informationally decentralized video streaming over multihop wireless networks," *IEEE Trans. Multimedia*, vol. 9, no. 6, pp. 1299–1313, Sep. 2007.
- [14] D. Fudenberg and D. K. Levine, *The Theory of Learning in Games*. Cambridge, MA: MIT Press, 1998.
- [15] H. P. Young, *Strategic Learning and its Limits*. New York: Oxford Univ. Press, 2004.
- [16] D. Bertsekas and R. Gallager, *Data Networks*. Upper Saddle River, NJ: Prentice-Hall, 1987.
- [17] R. Draves, J. Padhye, and B. Zill, "Routing in multi-radio, multi-hop wireless mesh networks," in *Proc. ACM Int. Conf. MOBICOM*, 2004, pp. 114–128.
- [18] D. Krishnaswamy, "Network-assisted link adaptation with power control and channel reassignment in wireless networks," in *Proc. 3G Wireless Conf.*, 2002, pp. 165–170.
- [19] D. Krishnaswamy *et al.*, "A cross-layer cross-overlay architecture for proactive adaptive processing in mesh networks," in *Proc. 2nd IEEE WiMesh*, 2006, pp. 74–82.
- [20] D. Jurca and P. Frossard, "Media streaming with conservative delay on variable rate channels," in *Proc. IEEE ICME*, 2006, pp. 1841–1844.
- [21] C. E. Perkins and E. M. Royer, "Ad hoc on-demand distance vector routing," in *Proc. 2nd IEEE Workshop Mobile Comput. Syst. Appl.*, Feb. 1999, pp. 90–100.
- [22] G. D. Kondylis and G. J. Pottie, "Dynamic channel allocation strategies for wireless packet access," in *Proc. IEEE VTC*, Amsterdam, The Netherlands, Sep. 1999, pp. 2819–2824.
- [23] J. Chakareski and P. Frossard, "Rate-distortion optimized distributed packet scheduling of multiple video streams over shared communication resources," *IEEE Trans. Multimedia*, vol. 8, no. 2, pp. 207–218, Apr. 2006.
- [24] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Commun. Mobile Comput.*, vol. 2, no. 5, pp. 483–502, Aug. 2002.
- [25] H.-P. Shiang and M. van der Schaar, "Delay-sensitive resource management in multihop cognitive radio networks," in *Proc. IEEE DySPAN*, Oct. 2008.



Hsien-Po Shiang is currently working toward the Ph.D. degree with the Department of Electrical Engineering, University of California, Los Angeles.

During his graduate study, he was with Intel Corporation, Folsom, CA, researching the overlay network infrastructure over wireless mesh networks, in 2006. He has recently been selected as one of the eight Ph.D. students for the 2007 Watson Emerging Leaders in Multimedia Workshop organized by IBM Research. He has authored several journal and conference proceeding papers. His research interests

include cross-layer optimizations/adaptations for multimedia transmission over wireless mesh networks and dynamic resource allocation based on collaborative information exchange for delay-sensitive applications.



Mihaela van der Schaar (SM'08) received the Ph.D. degree from Eindhoven University of Technology, Eindhoven, The Netherlands, in 2001.

She is currently an Associate Professor with the Department of Electrical Engineering, University of California, Los Angeles. Since 1999, she has been an active participant to the ISO MPEG standard, to which she has made more than 50 contributions. She is an Editor (with P. Chou) of the book *Multimedia over IP and Wireless Networks: Compression, Networking, and Systems* (Academic, 2007). She is also

the holder of 30 U.S. patents.

Ms. van der Schaar was the recipient of the National Science Foundation CAREER Award in 2004, the IBM Faculty Award in 2005 and 2007, the Okawa Foundation Award in 2006, the Best IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY Paper Award in 2005, the Most Cited Paper Award from the *EURASIP Journal Signal Processing: Image Communication* from 2004 to 2006, and three ISO Recognition Awards.