Cooperation and Communication in Cognitive Radio Networks based on TV Spectrum Experiments

Kaushik Chowdhury, Rahman Doost-Mohammady, Waleed Meleis
Dept. of Electrical and Computer Engineering
Northeastern University
Boston, MA, USA
{krc,doost,meleis}@ece.neu.edu

Marco Di Felice, Luciano Bononi
Dept. of Computer Science
University of Bologna
Bologna, Italy
{difelice, bononi}@cs.unibo.it

***Abstract*—Cognitive radio (CR) ad hoc networks are composed of wireless nodes that may opportunistically transmit in licensed frequency bands without affecting the primary users of that band. In such distributed networks, gathering the spectrum information is challenging as the nodes have a partial view of the spectrum environment based on the local sensing range. Moreover, individual measurements are also affected by channel uncertainties and location-specific fluctuations in signal strength. To facilitate the distributed operation, this paper makes the following contributions: (i) First, an experimental study is undertaken to measure the signal characteristics for indoor and outdoor locations for the TV channels $21-51$, and these results are used to identify the conditions under which nodes may share information. (ii) Second, a Cooperative reinforcement LearnIng scheme for Cognitive radio networKs (CLICK) is designed for combining the spectrum usage information observed by a node and its neighbors. (iii) Finally, CLICK is integrated within a MAC protocol for testing the benefits and overhead of our approach on a higher layer protocol performance. The proposed learning framework and the protocol design are extensively evaluated through a thorough simulation study in ns-2 using experimental traces of channel measurements.**

***Keywords*-cognitive radio; cooperation; learning; TV spectrum;**

I. Introduction

Cognitive radio (CR) is an enabling technology that allows opportunistic transmission in under-utilized or vacant frequencies, thereby resulting in high spectrum efficiency. In CR ad hoc networks, owing to the distributed operation of the nodes, ensuring that the licensed users or primary users (PUs) of the spectrum are not affected is a challenging task. CR nodes are neither aware of the global network topology, including the locations of the PUs, and nor do they have *a priori* access to PU transmission schedules [1]. Thus, the CR nodes must rely on learning the local channel utilization through individual measurements, and share these results with neighboring nodes. The cooperation between nodes facilitates quick dissemination of the knowledge of the spectrum environment, and also reduces missed detection errors by merging together different data sets from close-by locations. In this paper we explore how to select the best channels for cooperation, how to merge together spectrum knowledge gained by the nodes, and use these learnt channel utilization models during link-layer transmission.

Prior studies on spectrum utilization have revealed significant spectrum availability in the UHF band [1]. Consequently, the FCC has recently opened up the vacant spectrum in the TV channels 21 ($512\,\mathrm{MHz}$) to 51 ($698\,\mathrm{MHz}$), with the exception of channel 37, for use by unlicensed devices [8]. However, FCC specifically points out the need to ensure that the reception of both high and low power TV signals are not adversely impacted. Thus, any practical CR deployment must be preceded by a comprehensive study of the characteristics of the expected received power in these channels, the utilization in these frequencies, in both indoor and outdoor environments. In the first part of our work, we undertake the spectrum study over the complete range of the TV channels $21-51$ and note the variation of the signal strength with distance, location, and frequency. From this study, we identify which channels exhibit reliable and time-invariant behavior at a given location. Moreover, we point out how a CR node can independently identify these reliable channels without external administrator help, an important consideration in an ad hoc network, where nodes rely on each other's spectrum measurements.

CR nodes periodically undertake spectrum sensing to ensure that the channel knowledge stays current, and the choice of the channel does not affect the licensed users. While most works assume on-off Markovian models based on the birth-death process, recent experiments have pointed out that the actual PU activity is captured better by long-tailed exponential distributions [4]. Thus, during actual usage, it is beneficial if the CR user samples the channels and builds its own estimation of the PU activity. We propose a Cooperative reinforcement LearnIng scheme for Cognitive radio networKs (CLICK) to achieve this, that differently from previous works, also takes into account (i) the reliability of the channel information, and (ii) the level of trust that can be assigned to the readings of the collaborating node. From our experimental findings we observe that even CR nodes a few meters apart from each other may exhibit a widely varying channel measurement based on location-specific wave reflections. Thus, both distance between nodes, and the peculiar characteristics of the wireless channel at the measurement location are key factors during the collaboration process.

While CLICK allows each CR user to learn about the channel availability over time, choosing a specific channel for communication requires the participation of both the sender-receiver pair. In order to demonstrate the benefit and the overhead of integrating CLICK in a higher layer protocol, we propose an extension of an existing multi-channel medium access control (MAC) protocol for CR ad hoc networks. We demonstrate how the collaboration among nodes results in better choice of channels, and the need for shorter individual sensing times. Both of these factors result in higher link layer throughput and better PU detection.

The rest of this paper is organized as follows. The experimental study that motivates our approach, and the related work are given in Section II. In Section III, we describe our proposed CLICK scheme in detail. The implementation of CLICK in a MAC protocol is given in Section IV. We undertake a thorough performance evaluation in Section V, and finally, Section VI concludes our work.

II. Related Work and Motivation

A. Related Work

Classical cooperation techniques mainly propose combining binary decisions (hard decision) made by the deployed nodes, or their spectrum sensing measurements (soft decision) to a centralized location or fusion center. As an example of the hard decision, the Bayesian hypothesis rule is applied in [6]. The authors point out that the main research problem is the need to appropriately select the decision thresholds that signal the detection event, both at the individual node and during combination at the fusion center. A similar approach using a voting rule is described in [20], where at least half of the nodes must agree on a decision. A more general formulation of collaboration based on K nodes concurring out of N, also called as the K by N rule, is given in [13], and later extended in [14] to also incorporate the MAC layer throughput optimization (note that $\frac{K}{N} = \frac{1}{2}$ in [20]). A more realistic case is presented in [3] where a distinct channel behavior between a given PU transmitter and a CR receiver is considered. Our proposed approach, CLICK, does not combine binary decisions. Instead, it merges together the values of the *states* and *actions* that define the extent of the learning within the node. These values determine the probabilistic availability of the channel. Thus, our collaborative approach predicts long-term spectrum availability information, rather than its immediate binary availability.

A distributed soft-decision based cooperation scheme for single channel ad hoc networks is proposed in [10] that is limited to forming pairs from a larger set of nodes. However, as we show in our experiments in Section II-B, such a collaboration between *any* two nodes on *any* channel is not guaranteed to return accurate results in a practical environment. As mentioned in [2] and independently verified in our experiments, the assumptions of equal SNR for all nodes with respect to a particular PU transmitter (assumed in [11] [20] [14]) or perfect channel behavior (assumed in [19] [12]), are incorrect. Moreover, the locations of the nodes, say indoors/outdoors, or in an elevated roof, also have bearing on the accuracy of the sensed data. CLICK appropriately weights the contributing information for each <channel,node> pair to account for these variations. Other node-only weighting schemes have been previously proposed in [15] [18], under the limited assumptions of perfect and non-fading channel behavior. Thus the above works fail to consider in a comprehensive manner both the effects of the wireless channel, and the need for location-specific weighting of the nodes in a practical setting. Moreover, as in the case of hard decision, the above schemes are useful for the current sensing instant only, and do not reflect on the long term suitability, in terms of CR network throughput and PU protection, when the channel is used.

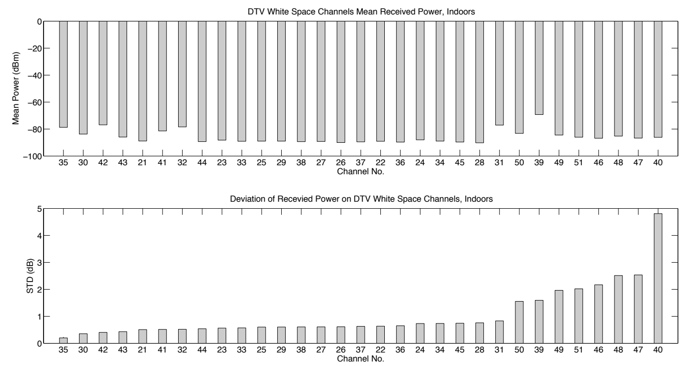


Figure 1. The variation in the mean and the standard deviation in the received power and pilot signal are shown for indoor location for TV channels $21-51$

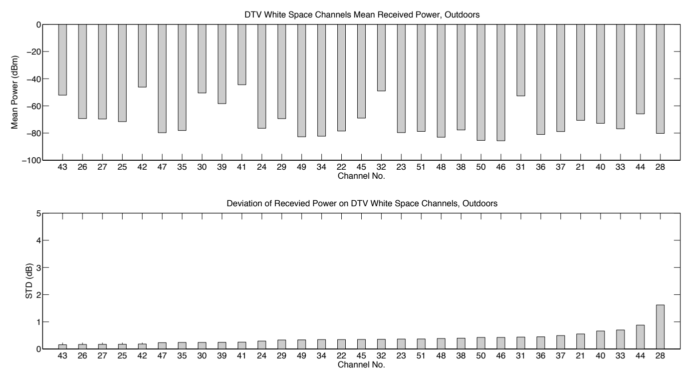


Figure 2. The variation in the mean and the standard deviation in the received power are shown for roof-top location for TV channels $21-51$

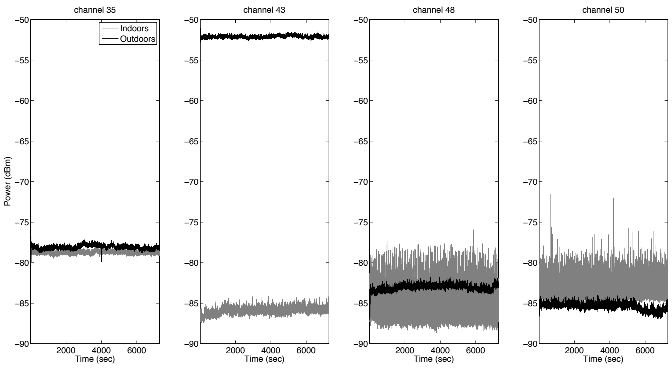


Figure 3. The time variation in the received signals at indoor (broken lines) and outdoor locations (bold lines)

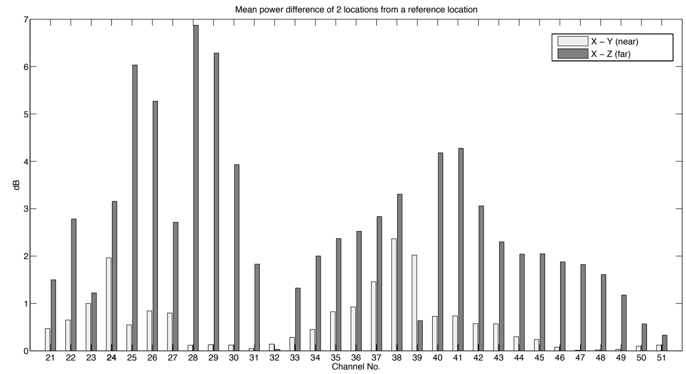


Figure 4. The pairwise difference of the mean power for near (X-Y) and far (X-Z) separation distances

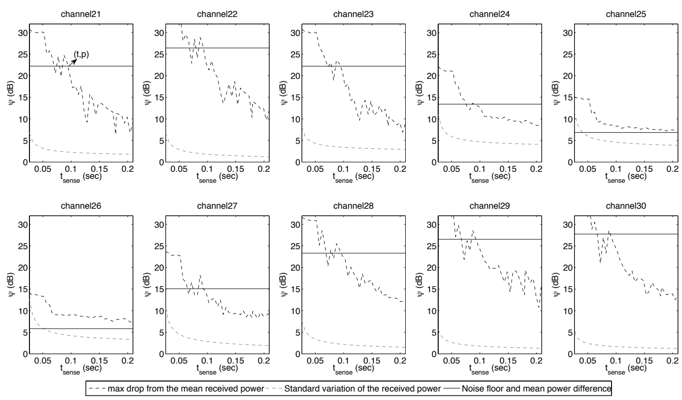


Figure 5. The maximum drop in the mean of the received signals, and its breaching the noise floor, for sensing time t^k_{sense} for the channels $k = 21 - 30$

B. Motivation

We believe that for CR networks, network design must be guided by real world experiments, with parameters carefully chosen so that the PU operation is not affected. We note that signal strength, TV transmitter locations, CR user placements, constructive and destructive effects on the signal caused by reflections from structures and the terrain, frequency of the channel, among others, affect the spectrum sensing performance. Moreover, while cooperating CR users may be spread over both indoor and outdoor locations, the PU receiver antennas for the case of the TV transmissions is typically outdoors, at an elevated surface. Thus, the TV signal measured by the indoor CR users may be markedly different, and local decisions by the node based on signal strength may not capture the received power levels experienced by the PU receiver antennas. We study these effects further on the digital TV frequencies corresponding to UHF channels $21 - 51$ to motivate the design of CLICK.

1) Experimental Setup and Results: The Northeastern University campus was chosen as the site for our study with experiments conducted, both indoors and outdoors, in two adjacent research buildings. In particular, the sole outdoor measurement was carried out on a high platform on the roof, which provided a line of sight (LOS) reception with a select set of TV transmission towers that were at approximate distances from $6.8 - 7.3\,\text{miles}$ [17]. The non-LOS transmitters ranged from 32 to well over 100 miles. To measure the signal strength, we used the Universal Software Radio Platform 2 (USRP2) equipped with WBX daughterboard

2) Setting the PU detection threshold: In Figure 1, we see both the mean received power in the channels $21 - 51$ (upper plot) arranged in the increasing order of the standard deviation σ of the power values (lower plot) for the reference indoor location X. We verified from [17] that at the time of experimentation, all the sensed channels were occupied. We observed that for all the 30 indoor measurement sites, including X, each channel exhibited a channel power of at least $-87\,\text{dBm}$, which we set as the detection threshold, i.e., greater than this value signaled the presence of the PU. This importance of this selection in a learning scheme was pointed out in [6], but no solution was proposed. From the mean calculations in the outdoor roof setting in Figure 2 (upper plot), the decision threshold is higher at $-81\,\text{dBm}$, owing to the reduced reflections, and better propagation paths with some of them being along the LOS.

C. Standard deviation or mean as a channel quality metric?

Though the transmission towers are spread out in large geographical areas, the mean received power does not show significant difference over a large extent of the spectrum. We argue that while the mean power is important in setting the decision threshold, and is indeed used by energy-detection schemes at the receiver, sudden and indeterministic fluctuations of the signal about the mean may result in detection errors. From Figure 3 we observe that for the outdoor location (bold lines), the signal does not show significant variation. Yet, for the indoor location X, and especially for channels 48 and 50, the variance with time is significant compared to the channels 35 and 43. During spectrum sensing, the node relies on a small window of collected samples to measure the signal energy and compare against a threshold. Hence, large variations in the signal may not be smoothened out in the short sample window, leading to detection errors. As an example, from the indoor standard deviation plot in Figure 1, though the mean powers are comparable in channels 43 and 48, the standard deviation σ in the latter is almost 200% higher. Thus, channel 43 exhibits more consistent behavior with time, and is more suited for exchanging channel information during collaboration. CLICK weights these channels with reliable readings (i.e. with lower σ) higher using a novel concept of the *expert* or E-value for a channel. We believe that simply weighting specific nodes is not sufficient in a CR environment, but the channel weights must also be considered.

D. Range of collaboration

How large should the physical separation between two collaborative nodes be? To answer this question, we fixed the reference location, say X, and considered the following two sets of nodes. The first set is represented by Y and composed of nodes $y \in Y$ such that their distance from X, is within $100\,\mathrm{m}$, i.e., $0 \leq d^{X,y} \leq 100\,\mathrm{m}$ shown by the dotted circle. Similarly consider the second set of nodes given by Z such that $100\,\mathrm{m} \leq d^{X,z} \leq 200\,\mathrm{m}$, shown by the space between the two circular regions. Let the mean received PU signal powers at the locations for the nodes in Y and Z be $\bar{m_y}$ and $\bar{m_z}$, respectively. The plots of the pairwise difference of these means from the mean signal power at location X, i.e., $\bar{m_x}$ is given by $|m_x - m_y|$ (white bars) and $|m_x - m_z|$ (shaded bars) in Figure 4, respectively. Results reveal that distance and orientations of the nodes in an indoor setting affects spectrum measurements considerably. Though the TV transmitters are located at distances several order of magnitude greater than, say, $d^{X,z}$, the difference in the means is caused by the varying reflective environment and paths experienced by the signals. We have repeated this three way measurements in 30 trials, each time replacing the reference node X. We find that for our indoor locations, the mean difference at $100\,\mathrm{m}$ is contained within $3\,\mathrm{dBm}$. Thus, we limit the collaboration radius of nodes to $100\,\mathrm{m}$ so that nodes assimilating the spectrum information from their neighbors are placed in a similar spectrum environment.

E. Intervals between collaboration events

The difference between the maximum standard deviation and the mean (a metric represented by ψ) calculated from different sample lengths is shown in Figure 5 for a limited set of channels from $21-30$ owing to space considerations. The ψ plot is shown by the bold broken lines, which intersects the horizontal line representing the difference between the noise floor and mean power level. Thus for shorter sample lengths, owing to shorter sensing times, the received signal is yet to converge to the true mean, and hence, ψ is higher than the mean to noise floor difference. For channel 21, this intersection is depicted by the point $\{t, p\}$, where t is the sensing time $t^{21}_{sense} = 0.09$. This can be interpreted as follows: for channel 21, the t^{21}_{sense} should be at least $0.09\,\mathrm{s}$. The minimum time between collaboration messages t_C is a critical design component, as waiting too long may also result in longer convergence time. At minimum, t_C must be at least $\sum_{k=1}^{|S|} t^k_{sense}$, i.e. the cumulative time for sensing all the channels.

We describe later in Section IV how the standard deviation in a channel is used to calculate a metric called as *expert value* for that channel, during collaboration.

III. CLICK: A Cooperative Reinforcement Learning Scheme

A. Overview

CLICK is a cooperative reinforcement learning scheme, and adapted from the Q-learning technique. It is composed of three stages:

Stage 1 -Intra-node Measurement: Each node undertakes spectrum sensing on the channels $21 - 51$ of the TV spectrum, and records the mean and standard deviation of the signal power measured for all of these channels. During this stage, the node detects the presence of the PU by comparing against the receive threshold (Section II-B2), and returns a reward $r = 1$ for those channels where the PU is present.

Stage 2. -Intra-node Learning: Nodes calculate the Q and expert (E) values defined in Sections III-B1 and III-B2 based on the channel measurements. Basically, the PU activity on the channels is captured by the Q-learning tables, where Q is a function of the total rewards for a given state, each PU detection event earning a reward of 1. Thus, the nodes learn about the channel availability for long-term use. This learning process is gradual, and as the PU may have intermittent transmissions, or have wireless channel induced sensing errors, a considerable time is needed before a highly accurate channel usage information is gathered locally. The E-values qualify how reliable the Q-values are, based on the signal deviations of the channel in which the measurements were undertaken.

Stage 3. -Inter-node Cooperation: The nodes periodically share their learnt information with the neighbors, while distinguishing with whom the collaboration is undertaken, and on which channels. In this stage, the different Q-table entries of the nodes are merged together after being appropriately weighted. Thus, collaboration accelerates the overall learning at a much higher rate than that possible by local measurements alone.

Stage 1 has been explained in detail earlier in Section II-B, where the channel measurements were collected. This stage is periodically repeated to maintain the spectrum information current. The choice of sensing time t_{sense} for a given channel was also discussed (further optimization strategies to reduce the total sensing duration are presented in Section IV). The rest of this section describes the stages 2 and 3 in the operation of CLICK, while mainly answering the following set of questions: How are the experimental channel measurements, especially the location based uncertainties, incorporated in the collaboration scheme? Which subset of nodes from a larger neighborhood pool are suitable for collaboration, and which channels do they share their information on? What is the interval over which the collaboration is undertaken, and what is the overhead?

B. Stage 2 - Intra-node Learning

The learning model is defined by a Markov Decision Process (MDP) which is composed of states represented by S. Each state $s \in S$ maps to a specific TV channel numbered between $21-51$. The set of actions A maps the transition between states, which is the formal definition of the function of channel switching. The probability with which the node switches the channel (i.e. the state), say from s to s', through the action a, is calculated from the distribution function $\pi(s,a) \in [0,1]$. The goal of the learning process is to determine the optimal policy that returns the channel with the least Q, meaning that it is free from PU activity for longer durations of time. During Q-learning, the node observes the current channel s, selects a possible action a, and receives a reward r from the environment for that specific action. Different from the classical Q-learning technique that only maintains the Q (or reward) value table for each state-action pair, we also propose a new metric called as the *expert* (E) value. Here, for a given state action pair (s,a) and a node i, the value $E^i(s)$ defines how accurate the corresponding Q-value $Q^i(s,a)$ is. The formal expression of the Q and E values, and their update equations are as follows:

1) Q-Value:

$$Q^i(s,a) = (1-\alpha)Q^i(s,a) + \alpha r, \tag{1}$$

where the reward $r=1$ if PU is detected, and 0 otherwise. α is a tuning parameter that decides the speed of learning. The probability $\pi^i(s,a)$ of arriving in the state s from a prior state s' for node i by choosing action a is given by the Boltzmann distribution,

$$\pi^i(s,a) = \frac{e^{Q^i(s,a)/T}}{\sum_{j\in A} e^{Q^i(s,j)/T}}, \tag{2}$$

where T is the temperature parameter and adjusts the tradeoff between exploration and exploitation actions. The values of α and T are taken as 0.2 and 10 [16], though T itself being a progressively decreasing metric have also been explored in the learning literature.

2) Expert or E-Value:

$$E^i(s) = 1 - \min[1, \frac{\sigma}{M}], \tag{3}$$

The E-value $0 \le E^i(s) \le 1$ is a function of standard deviation σ of the signal received on a channel, represented by the state s, and the mean M. Each node maintains a history of the last H signal samples for a sensed channel s. We use $H = 30$ to ensure adequate buffer space in the nodes Thus, the higher is the deviation in the signal strengths, lesser is the reliability of the channel. Hence, the PU occupancy predicted by $Q(s,a)$ results in lower accuracy.

C. Stage 3 - Inter-node Cooperation

The cooperative learning scheme allows the agents to share their $Q-$ and $E-$ values, in order to increase the convergence speed of the learning. The cooperation is achieved in the following three steps: (i) The *sharing* function (Section III-C1) allows each node to decide the Q values of which channels should be shared with the other nodes. (ii) The *combine* function (Section III-C2) tells the node how to combine its own stored Q and E-values, with the respective values received from other collaborating nodes.

1) Sharing Function: A given node, say i, creates a list L of only those channels for which the E-value is above a pre-decided threshold ϕ_E, meaning that the channel measurements are reliable, and consistent over time. For these channels $s \in L$, the respective tuples $< Q^i(s,a), E^i(s)$ are sent to the collaborating nodes. Thus, the list L that is eventually shared can be formally defined as:

$$L = \{< Q^i(s,a), E^i(s) > | s \in S, a \in A | E^i(s) > \phi_E\}, \tag{4}$$

where ϕ_E is the decision threshold set at $1 - \frac{\sigma=1}{|M|=80} = 0.9875$. The rationale for selecting this value is that our experimental results show in Figure 1, the average indoor mean power was found to be at $-87\,\mathrm{dBm}$ and the deviation σ for the reliable channels was contained within $1\,\mathrm{dB}$. Thus, we select ϕ_E of a comparative value, while maintaining a small safety factor. This sharing function is repeated every t_C time steps, with the minimum value of t_C derived in Section II-E. While lower values of t_C increase network overhead, higher values of t_C lower the learning rate, and may lead to PU interference. We investigate the effect of different intervals of t_C on the PU detection and network overhead in Figure 7(a) during the performance evaluation.

2) Combine Function: When node i receives the sharing list L^j from the collaborating node j, with $L^j = \{< Q^j(s,a >, E^j(s,a) >\}$, it will update its Q-values by calculating the weighted average as follows:

$$Q^i(s,a) = (1-W^{i,j}(s,a)) \cdot Q^i(s,a) + W^{i,j}(s,a) \cdot Q^j(s,a), \tag{5}$$

where $W^{i,j}$ is a weight function which decides how much the information from node j will contribute to adjust the learning process of node i. In our formulation, $W^{i,j}$ is a combination of two weighting factors, ϵ and η:

- *Expert Weight* ($\epsilon^{i,j}(s,a)$): When a node i receives Q-values from another node j, it may only integrate these new values in its own Q-table if the expert value or E-value of j is greater or equal to its own. This prevents dilution of the accuracy of its own measurements on the channels during collaboration. Formally, this is expressed as the ratio of the difference in the expert weights for a given collaborating pair to the cumulative difference of the expert values for all the K collaborating nodes combined, i.e.,

$$\epsilon^{i,j}(s,a) = \begin{cases} \frac{E^j(s)-E^i(s)}{\sum_{n=1}^{K} E^n(s)-E^i(s)} & \text{if } E^j(s) > E^i(s) \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

Similarly, if $E^j(s) - E^i(s) > 0$, then the node i will increase its expert value $E^i(s)$ as more accurate channel measurements have been incorporated. This increase is given by the weighted sum:

$$E^i(s) = E^i(s) \cdot (1 - \eta^{i,j}) + \eta^{i,j} \cdot E^j(s) \quad (7)$$

- *Trust weight* ($\eta^{i,j}$): It measures how much a given node i can trust the $Q^j(s,a)$ values coming from node j. The difference between η and ϵ is that the former depends on the distance between the collaborating node pair (i.e. the $< i,j >$ pair) and not on the specific channel behavior represented by the state-action values (i.e. the $< s,a >$ pair). From Figure 4, neighbors closer in distance exhibit channel measurements that have a higher correlation with a node's own observations, and can therefore be trusted. Thus,

$$\eta^{i,j} = 1 - \frac{d^{i,j}}{R}, \quad (8)$$

where R is the transmission range, and $d^{i,j}$ is the distance between the nodes i and j.

The cumulative weight function $W^{i,j}$ in eq(5) can now be written as:

$$W^{i,j}(s,a) = \epsilon^{i,j}(s,a) \cdot \eta^{i,j} \quad (9)$$

IV. Leveraging Cooperation at the Link Layer

In this section we develop a simple MAC protocol to demonstrate how the cooperation between the nodes can be implemented and analyzed in a practical setting. We mainly study the following effects at the link layer:

- Most existing MAC protocols rely on the transceiver pair to select a currently vacant channel. Instead, CLICK selects the channel with the least Q-value with the aim of maximizing the long term availability, and enhanced PU protection. The benefit and tradeoff of such a selection strategy, as opposed to the classical approach of *any* vacant channel selection, must be evaluated.
- The collaboration imposes a load on the network, resulting in transmission delays and packet losses. The effect of transmitting this additional state-action information on link layer performance must be quantitatively measured.

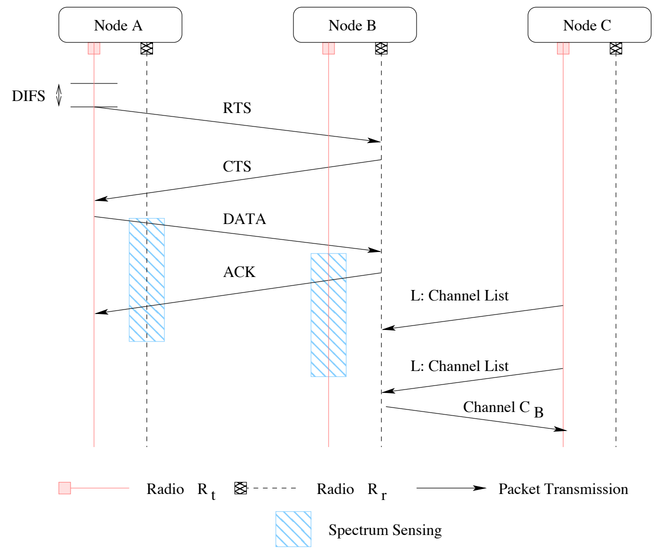


Figure 6. The MAC protocol design that implements CLICK

- The spectrum sensing time can be reduced though collaboration, wherein a node rapidly gains a high expert or E-value by assimilating other's state-action tables. Thus, expending a lot of time for sensing to collect its own measurements may no longer be needed, and the radio may instead be used for data transmissions to improve throughput.

A. MAC Protocol Design

We assume that the CR users are equipped with two radio transceivers, which we call as the receive radio (R_r) and transmit radio (R_t) as shown in Figure 6. Each node also has a default receive channel which is chosen as the channel with the least Q-value, i.e. least PU activity at its location. The receive radio R_t is always tuned to this default channel, unless requested otherwise, and the choice of this channel is broadcast through *hello* messages periodically by the receiver. This packet also contains the Q and E-values used by the neighboring nodes for collaboration. The transmit radio R_t tunes to the receive channel of the intended recipient to begin the link layer packet transfer. Unlike the work in [5], we do not assume any network administrator assigned channel, or specialized low power radios.

The two cases for MAC layer coordination when a node is attempting a packet transfer to node B are as follows:

1) Channel Common to Sender and Receiver: Consider the interaction between A and B, with say, channel x as the default receive channel for B. If the Q-value of this default channel x at node A is lower than the pre-decided threshold (assumed at 0.4 to have at least greater than 50% safety), then it immediately tunes its transmit radio R_t to x. It listens to the channel for DIFS time, sensing both the PU and the transmission of its neighboring CR nodes, before sending out the request to send (RTS) packet, as defined in the 802.11

standard. The RTS is received by the node B on channel x, and the remainder of the interaction follows the classical clear to send (CTS)-DATA- acknowledgement (ACK) cycle. If the RTS-CTS handshake is successfully concluded, just as the data transmission begins, the free radios (R_r for A and R_t for B) begin the spectrum sensing of the channels, till all the channels from $21-51$ are sensed.

2) Different Channels between Sender and Receiver: After successful completion of the transfer between A and B, consider another data transmission, this time initiated by node C for recipient B. Unlike the previous case, assume that channel x has a Q-value greater than the permissible threshold for C, and it is likely that transmission in this channel will get interrupted due to PU activity in C's neighborhood. Hence, C now proposes a channel list L to B with permissible Q-values, and allows B to pick the best channel. B chooses the channel from the entries of the list L, which has the least Q-value at its own location, say channel $C_B = y$, and informs the sender C. The receiver radio R_r for B is now tuned to y for the duration of the transfer, and the RTS-CTS-DATA-ACK cycle is repeated. If no permissible channel is found, the channel reply is denied to node C.

B. Cooperation Benefits in the MAC Protocol

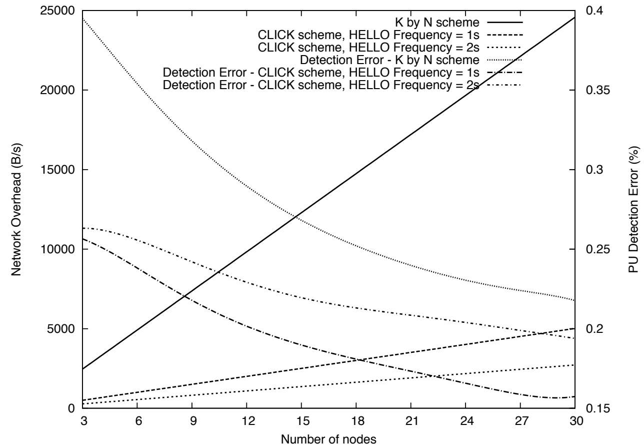
1) Sensing Time Optimization: The time used for collecting the spectrum information results in lowering the link layer throughput as the transceiver is busy. As seen in Figure 6, node B is engaged in spectrum sensing, and is unable to accept new requests from node C, leading to retransmissions of the channel list L. Through collaboration, the individual sensing time can be shortened as the state table is enhanced by integrating the information contained by the neighboring nodes. This actual sensing time, T_t^i for node i considering all the channels in S is,

$$T_t^i = \sum_{k=1}^{|S|} t_{sense}^k \left[1 - \frac{E^i(k) - E'^i(k)}{E'^i(k)}\right] \tag{10}$$

where $E^i(k)$ and $E'^i(k)$ are the E-values after collaboration, and from the individual measurements, respectively. This expression implies that the minimum sensing time t_s^kense for a given channel $k \in S$ (obtained from measurements described in Section II-E by each node) is scaled by the fractional increase in the expert or E-value at a node. This results in a total sensing time savings of $\triangle_t^i = \sum_{k=1}^{|S|} t_{sense}^k - T_t^i$.

C. PU protection and Switching Minimization

The Q-value, with time, captures not only the spectrum availability at the node location, but is also a measure of its availability in the immediate neighborhood as well. This is because, the the Q-value table is periodically exchanged between 1-hop neighbors, and reliable channel readings integrated. Thus, choosing the channel based on its long



(a)

Figure 7. Overhead of CLICK and *K by N* scheme.

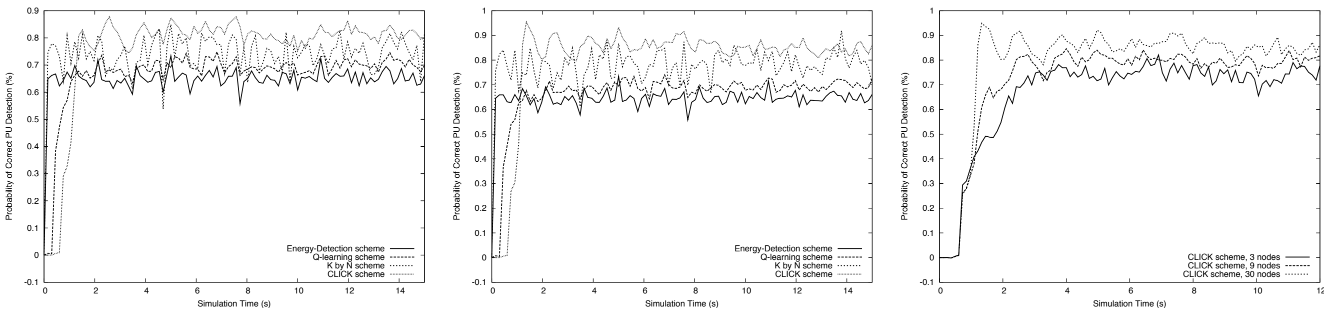
term availability saves on repeated spectrum switching, and ensures continued communication without possible service degradation to the PUs. Moreover, as a control channel is absent to make the link layer more spectrally efficient, it is vital that the default receiver channel R_r be available for a longer extent of time.

V. Performance Evaluation

In this section, we evaluate the performance of the proposed cooperation scheme CLICK and its performance at the link layer using the ns-2 simulator extended for CR networks [9]. We use real measurement traces to derive the PDF of successful PU detection for different durations of sensing time. This follows from the discussion in Section II-E, where we measure the number of samples for which the received signal strength of an active TV transmitter falls below the noise floor leading to missed-detection. We consider a grid topology of 9 cells at the Northeastern University campus, in a $300 \times 300\,\mathrm{m}^2$ area , and the missed detection PDFs for TV channels $21-51$ are obtained by carrying out measurements in each cell by USRP2 devices. The latter are deployed randomly, and several sets of time series measurements are saved to later simulate the real world environment. Thus, practically observed error probability and signal strength traces are used for the packet level simulation in ns-2 We restrict the analysis to a subset of TV channels $25-29$ and $38-40$ which we exhibit higher variance of the PU signal for different grid locations.

A. PU Protection Analysis

In this analysis, we evaluate the benefits and costs of cooperation and learning on the PU detection process. The sensing duration and inter-sensing intervals are fixed at $0.026\,\mathrm{s}$ and $0.2\,\mathrm{s}$, respectively to ensure correct PU detection. For fairness, the sensing time optimization (Section IV-B1) is disabled in this experiment. We compare the performance



(a) (b) (c)

Figure 8. The probability of correct PU detection for 15 nodes (a) and 30 nodes (b), while the convergence behavior for CLICK is in (c).

of CLICK with (i) *Individual sensing*, i.e., classical energy detection using threshold from Section II with no cooperation, (ii) Q-learning at the node level, again without cooperation, and (iii) the *K by N* collaboration scheme proposed in [20].

In Figure 7(a), we show the overhead of the *K by N* scheme and of two configurations of CLICK with different collaboration intervals (i.e. every $1.0\,\mathrm{s}$ and $2.0\,\mathrm{s}$, respectively). The *K by N* scheme incurs the highest overhead due to the fact that each CR node must collect measurements from other nodes after each sensing interval (i.e. $0.2\,\mathrm{s}$ in our experiments). Figure 7(a) shows that the frequency of sharing actions has also a critical impact on CLICK, i.e. higher accuracy can be guaranteed with more frequent updates. However, the sharing action can be performed asynchronously and at lower rate than the sensing, and this explains the overhead reduction of our scheme. In Figures 8(a) and 8(b) we depict the probability of PU detection over time in the configurations with 15 and 30 nodes, respectively. In both figures, we notice that the cooperation accelerates the learning at a higher rate, when compared to the non-cooperative Q-learning scheme, giving about $12-15\%$ improvement. Figure 8(c) shows the improvement on the increasing number of nodes on the convergence time of CLICK.

B. Analysis of MAC Protocol using CLICK

In this section, we study the benefits of CLICK integrated into our proposed MAC protocol, which we call as CLICK MAC. We consider a topology with 12 CR nodes with 6 active connections, and we vary the system load produced by each connection. The traffic type is Constant Bit Rate (CBR) with the UDP protocol at transport layer. Figure 9(a) shows the throughput of the proposed MAC scheme for different sensing time intervals. From Figure 9(a), that the communication performance at MAC layer is improved by an accurate setting of the sensing interval. When the sensing time is set to the minimum value, PU interference may lead to packet losses during the transmission period of CR users. However, long sensing time reduces the transmission opportunities for CR nodes. The CLICK MAC scheme with Adaptive Sensing is able to balance the tradeoff between PU protection and CR transmission opportunities, as shown in Figure 9(a). Figure 9(b) shows the throughput of the proposed MAC scheme with different underlying sensing schemes for deciding the default channel for the receiver interface R_r at each node. In the case of *Individual*, Q-learning and *K by N* scheme, channel selection is performed by receiver node only. By increasing the long-term knowledge on each channel, CLICK reduces the overhead of channel switching, and the impact of PU interference. As a result, the integrated CLICK MAC scheme guarantees the highest throughput also under high traffic loads. The same benefit can be seen in terms of end-to-end delay, which is shown in Figure 9(c).

VI. Conclusion

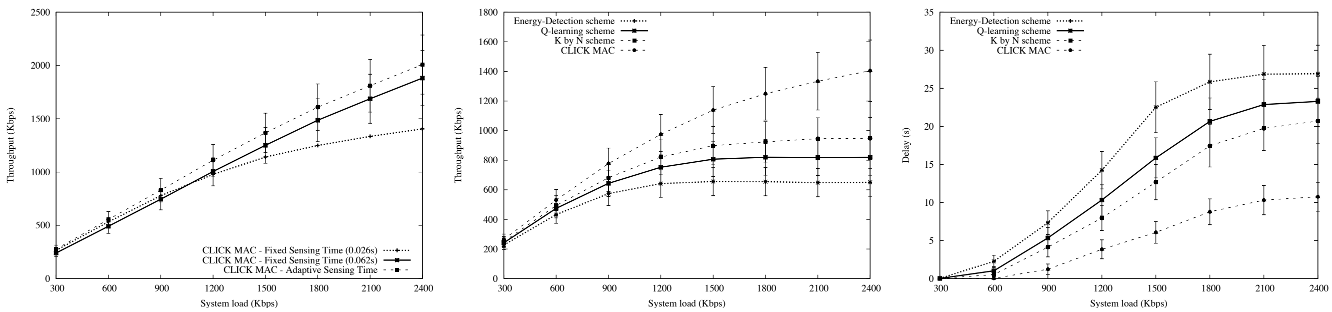
We have undertaken actual channel measurements in the TV spectrum band covering channels $21-51$. Using these measurements, we have drawn inferences on how cooperation could be achieved in a distributed environment. Our reinforcement learning approach CLICK takes into account channel characteristics and node location to decide *which* channels and nodes are suitable for collaboration. Finally, we have demonstrated the benefits of cooperation by extending a MAC protocol for CR operation. Our future research directions will involve leveraging cooperation in the operation of the higher layer network protocols.

References

[1] I. F. Akyildiz, W. Y. Lee, and K. Chowdhury, CRAHNs: cognitive radio ad hoc networks. *Elsevier Ad Hoc Networks Journal*, vol. 7, no. 5, pp. 810-836, July 2009.

[2] K. Arshad, M. A. Imran, and K. Moessner, Collaborative spectrum sensing optimisation algorithms for cognitive radio networks. *Intl. Journal of Digital Multimedia Broadcasting*, vol. 2010, Article ID 424036, 20 pages, 2010.

[3] T. Aysal, S. Kandeepan, and R. Piesiewicz. Cooperative spectrum sensing over imperfect channels. *in Proc. of Globecom Workshops*, pp. 15, December 2008.



(a) (b) (c)

Figure 9. The throughput of CLICK MAC for different values of sensing time interval (a), and sensing schemes (b). The link delay is given in (c).

[4] D. Chen, S. Ying, Q. Zhang, M. Liu, and S. Li. Mining spectrum usage data: A large-scale spectrum measurement study. *Proc. ACM International Conference on Mobile Computing and Networking (MobiCom)*, September 2009, Beijing, China.

[5] K. R. Chowdhury, P. Chanda, D. P. Agrawal, and Q. A. Zeng. Channel allocation and medium access control for wireless sensor networks. *Ad Hoc Networks Journal, (Elsevier)* , vol.7, no.2, pp. 307-321, March 2009.

[6] C. R. C. M. da Silva, B. Choi, and K. Kim, Distributed spectrum sensing for cognitive radio systems. *in Proc. Information Theory and Applications Workshop*, 2007.

[7] Ettus Research LLC. Website [Online]:. http://www.ettus.com/

[8] FCC Press Release. FCC adopts rules for unlicensed use of television white spaces. Second Report and Order (FCC 08-260), November 2008.

[9] M. D. Felice, K. R. Chowdhury, and L. Bononi, Modeling and performance evaluation of transmission control protocol over cognitive radio ad hoc networks. *ACM MSWIM Conf.*, pp. 4-12, Spain, October 2009.

[10] G. Ganesan and Y. G. Li. Cooperative spectrum sensing in cognitive radio: Part I: two user networks. In *IEEE Transactions on Wireless Communications*, vol. 6, pp. 2204-2213, June 2007.

[11] A. Ghasemi and E. S. Sousa. Opportunistic spectrum access in fading channels through collaborative sensing. *IEEE Journal of Communications*, vol. 2, no. 2, pp. 7181, 2007.

[12] S. M. Mishra, A. Sahai, and R. W. Brodersen. Cooperative sensing among cognitive radios. In *Proc. of IEEE ICC*, pp. 16581663, August 2006.

[13] E. Peh and Y. C. Liang Optimization for cooperative sensing in cognitive radio networks. In *Proc. of IEEE WCNC*, pp. 2732, April 2007.

[14] E. Peh, Y. C. Liang, and Y. L. Guan. Optimization of cooperative sensing in cognitive radio networks: A sensing-throughput tradeoff view. In *Proc. of IEEE ICC*, June 2009.

[15] Z. Quan, S. Cui, and A. Sayed. Optimal linear cooperation for spectrum sensing in cognitive radio networks. *IEEE Journal on Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 2840, 2008.

[16] R. S. Sutton and A. G. Barto Reinforcement learning: An introduction. *MIT Press*, Cambridge, MA, 1998.

[17] TV Fool Coverage Maps. Website [Online]:. http://www.tvfool.com/

[18] F. E. Visser, G. J. M. Janssen, and P. Paweczak. Multinode spectrum sensing based on energy detection for dynamic spectrum access. In *Proc. of IEEE Vehicular Technology Conference*, pp. 13941398, June 2008.

[19] E. Visotsky, S. Kuffher, and R. Peterson. On collaborative detection of TV transmissions in support of dynamic spectrum sharing. In *Proc. of IEEE International Symposium on Dynamic Spectrum Access Networks*, pp. 338345, December 2005.

[20] W. Zhang, R. K. Mallik, K. Ben Letaief, Cooperative spectrum sensing optimization in cognitive radio networks. In *Proc. of IEEE ICC*, 19-23 May 2008.