

ESCAPE: A Channel Evacuation Protocol for Spectrum-Agile Networks

Xin Liu and Zhi Ding
University of California, Davis

Abstract—Preliminary studies and general observations indicate a significant amount of “white space” in radio spectrum, varying in frequency, time, and geographic locations. Enabled by regulatory initiatives and radio technologies advances, opportunistic usage of the white space can potentially mitigate the spectrum scarcity. In this paper, we consider a scenario in which secondary users can opportunistically access unused spectrum vacated by idle primaries. In such a spectrum-agile system, a critical and unique issue is the fast and reliable evacuation of secondary users upon the return of primary users. To address this need, we propose an in-band signaling scheme, named ESCAPE. We present the design of the ESCAPE protocol in physical, MAC, and routing layers in the paper. The proposed scheme is distributed, reliable, and has minimum requirements on the system.

I. INTRODUCTION

In this paper, we consider a spectrum-agile network with two types of users. Primary users are the rightful owners and have strict priority on spectrum access. Secondary users are cognitive devices that can sense the environment and adapt to appropriate frequency, power, and transmission schemes. They can opportunistically access unused spectrum vacated by idle primaries. The objective is to design secondary networks that are *non-intrusive* and pose minimum disruption to primary-network users. One critical issue in such communication scenarios is the timely evacuation of secondary users upon the return of primary users, which is the focus of this paper.

Consider an **illustrative scenario**: a set of spectrum-agile communication devices (secondary users) communicate over a spectrum unused by primary users. These secondary users are deployed in an ad-hoc manner with no centralized controllers. Upon the return of primary users, the secondaries need to exit the channel quickly in order to be the least intrusive. Due to differences in device capability, environment, and communication load, not all users can sense the return of primary users. Therefore, the secondary users that detect the return of primaries need to propagate such evacuation information to other secondaries quickly and reliably. We call this information dissemination process *the evacuation process* and the evacuation information the *warning message*. Such information exchange encounters a few challenges:

- The transmission of evacuation information is subject to interference from both primary users and other secondary transmissions.
- A secondary user in transmission cannot receive simultaneously.

- A secondary user may not have global information regarding the topology of the network; e.g., node locations and density.
- The evacuation process must meet the delay and interference constraints imposed by primary users or regulatory bodies.

Because of these challenges, naive strategies may not effectively evacuate the channel. Failure to evacuate secondary users may cause undesirable interference to primary users. In addition, a secondary user failed to evacuate (leave the primary channel) may lose connectivity. In this paper, we propose a simple yet reliable protocol, named Embedded SpeCtrally Agile radio Protocol for Evacuation (or ESCAPE), to enable fast and reliable evacuation information dissemination among secondary users. The design of ESCAPE involves the physical, link, and network layers, as briefly explained next.

In the physical layer, a secondary user that detects the presence of primary user(s) sends a predefined warning message that declares “primary-active”. The message is sent as CDMA signal to its neighbors using a predefined spreading code. Other neighboring secondary users hearing the message will abort their own transmissions and repeat a verbatim copy of the warning message “primary-active” for a few times. A node can transmit the warning message as soon as it receives it. The routing protocol is a simple flooding protocol. To satisfy the performance constraints set by primary users (such as evacuation delay and peak interference during evacuation), we need to determine the design variables in the ESCAPE protocol including transmission power, code type and length, number of repetitions, and detection threshold.

The intuition of ESCAPE design is as follows: 1) Spreading provides good interference tolerance properties because such evacuation information is subject to interference of both primary and other secondary transmissions; 2) A well constructed spreading code can have very good autocorrelation property. Thus, multiple users can announce “primary-active” with little coordination among themselves. This significantly simplifies the MAC and alleviates the potential instability and congestion due to a high burst of traffic demand; and 3) Routing is based on flooding and thus requires little prior knowledge on network topology. It also provides redundancy in the warning message dissemination and thus improves reliability.

In summary, our objective is to disseminate the evacuation information among all secondary users and thus evacuate the primary channel reliably. The ESCAPE protocol can tolerate interference from both primary and other secondary transmissions. It requires little prior information on the network topol-

ogy and density. It is distributed in nature. It works regardless the number of secondary users that detect the primary. It does not require synchronization among users. Thus, the ESCAPE protocol is suitable for evacuation purposes.

The paper is organized as follows: system model is introduced in Section II. We also discuss the limitations of some common protocols if used for evacuation. The ESCAPE protocol is presented in Section III, followed by its parameter design in Section IV. The performance analysis is included in Section V and simulation results in Section VI. Related work is discussed in Section VII, followed by conclusion and future work in Section VIII.

II. SYSTEM MODEL

We consider two types of users. Primary users have strict priority on spectrum access, who are often conventional legacy users whose hardware and protocols should not be required to retrofit secondary user access needs. Secondary users are cognitive devices that opportunistically access unused spectrum vacated by idle primaries. Secondary users are deployed in an ad-hoc manner with no central controller. For instance, multiple WLAN devices in a building may use an unoccupied TV band [5].

Our focus is on channel evacuation. Therefore, we assume that the detection of primary users is achieved by at least one cognitive user using methods proposed in the many studies on cognitive radio, e.g., [2]. We note that the primary-user detection is a critical and challenging research problem by itself. Proposals include power-sensing, feature detection, centralized database, and other approaches, which is beyond the scope of this paper.

We do not assume that all secondary users that should evacuate can detect the return of (a) primary user(s). The reason is multi-fold. First, secondary devices can have different detection capabilities. For instance, one device may have a very sensitive detector and feature detection capabilities that others do not have. It is also possible that a fraction of secondary users (e.g., access points) has Internet access to check a certain database frequently while others do not. Another possibility is that due to channel fading and/or shadowing, some nodes may fail to sense the primary transmissions while other peers do. Furthermore, certain detection of primary users requires that secondary users listen to a channel for a sizable duration. Thus, secondary users with light communication load and more power resource are more suitable for such detection tasks. In summary, joint detection has the advantage of shared work load and detection reliability. It is thus important to share information on primaries among secondary users. ESCAPE is specifically designed for such a purpose.

We focus on one primary band. When a primary user activates, it will occupy the channel and no secondary user is allowed to co-exist in the vicinity. We assume that the primary power level is known to the secondaries. This information is needed to determine the power/length of the spreading code used by secondaries.

We do not assume that cognitive radios can transmit and receive simultaneously, which is the case for widely used TDD

(Time Division Duplexing) devices (such as WLAN devices) due to antenna constraints. We further assume that each node has a peak transmission power constraint as most radio devices do. We consider in-band signaling and show that it suffices for the purpose of channel evacuation.

The performance metrics of primary users include

- Evacuation time: From the moment a secondary user detects the primary to the instant all interfering secondary users evacuate the channel,
- Peak aggregated interference during the evacuation process,
- Average aggregated interference during the evacuation process, and
- Evacuation failure probability: the probability that some secondary users fail to evacuate the channel due to the failure of the warning message dissemination protocol.

The performance metrics for secondary users include

- Evacuation time: the longer the evacuation time, the larger the discrepancy of services among secondary users, and
- False alarm rate: it is possible that a secondary user falsely detects a warning message due to noise and interference while there is no such message in the air. A secondary user will broadcast such a (false) warning message and cause additional secondary users to vacate the channel when not needed. False alarm causes service interruption and resource waste and thus should be kept extremely low. (Note that this is a false detection of warning message instead of primary users. We assume no latter case in this study.)

ESCAPE is used for the sole purpose of spectrum-agile radio evacuation and does not handle PHY, MAC, or routing issues for regular secondary communications. The protocol performs well under the scenario when multiple or even all secondary users detect the return of primary users. Such multiple detections in general can accelerate the evacuation process.

We note that because of the specific challenges in the channel evacuation scenario, naive strategies may not work effectively. We next consider the following possible strategies and their limitations for channel evacuation. First, a secondary user can broadcast the warning message using extremely high power. This scheme may not work because 1) users have peak transmission power constraints; and 2) devices cannot transmit and receive simultaneously. In addition, transmission power, within its limit, is a control variable in the ESCAPE protocol. Second, a primary user can jam the channel and thus evacuate secondary users with high power. This scheme may not be feasible due to peak transmission power constraints on primary devices. In addition, primary users are legacy users, who are not required to retrofit the need secondary users. Third, an existing random access scheme, such as CSMA/CA or ALOHA, can be used to disseminate the warning message. ALOHA-family schemes suffer from instability. Carrier-sensing-based schemes may incur collision and backoff among warning messages and regular secondary transmissions. The need to flood the warning message may result in a high burst

of traffic and thus cause large delay and delay variation (due to the uncertainties in backoff schemes). In addition, without protection, the warning message may not be correctly received in the presence of primary interference.

Last, an out-of-band control channel can be used for warning message. We note that a control channel carries other control information, such as traffic load and link condition. Because of such traffic in the control channel, the control channel incurs the same issues, such as delay and interference, as a regular data channel, although less severe in general. Therefore, the ESCAPE protocol can also be used in the control channel to handle interference and reduce delay and delay variations. Certain parameters may need to be adjusted based on the traffic load, packet length, and other factors in the control channel. We note that both in-band and out-of-band control channels are being considered for spectrum-agile communication. Out-of-band schemes distinguish data communication channel and control channel physically. It requires either multiple radio interfaces or periodic (synchronized) switching. In comparison, in-band signaling requires a single radio interface, no frequent switching/hopping, and has a lower cost. On the other hand, because our signaling is based on a predefined CDMA warning message, we can also consider it as a logical control channel, as discussed in CORVUS [2].

III. ESCAPE PROTOCOL

The objective of ESCAPE is for multiple cognitive radios to evacuate the channel quickly and reliably. Consider the initiate state when the primary channel is unused. A group of secondary users detect the band and start to occupy the channel opportunistically. Later, primary users return, which is detected by one or more secondary users. The evacuation step begins. Secondary users who detect the primaries will transmit a pre-defined warning message declaring "primary-active". The warning message is modulated using a predefined CDMA spreading code. Other secondary users hearing the announcement will repeat the same warning message "primary-active". In this section, we discuss the details of the ESCAPE protocol.

A. Initialization Phase

A group of secondary users operating on a primary band need to agree on a few parameters of the warning message, including the pattern of the warning message, CDMA spreading code to be used, transmission power, and a few other parameters. An evacuation group is a group of connected secondary nodes sharing the same spread warning message. One user in the group detecting a primary user(s) will initiate the warning message that evacuates the whole group. The size and membership of the group are determined by the geographic area that need to be evacuated for active primaries. For instance, all WLAN devices in a building may belong to the same group while devices in different building do not. Nodes in a group may belong to different networks (or authority) and do not have regular communications (except for the warning message). A node may also belong to different (partially overlapping) evacuation groups. The purpose of the

initialization is to establish and broadcast the spread message. Any communication protocol may be used for such a purpose. A secondary user switched to the primary band after the initialization stage would obtain information on the warning message of the area from its neighboring nodes or a central authority.

B. Protocol Description

After the initialization, secondaries would sense and utilize the idle primary spectrum. During its normal operational phase, a secondary user performs the following procedure individually.

- Step 1: If a secondary user has a packet to transmit, it transmits its packet according to its regular access protocol. Then go to Step 2. If a user has no packet to transmit, go directly to Step 3.
- Step 2: Listen to the channel for L_s time unit.
 - If the user notices that the primary is back or it detects the warning signal, go to Step 4.
 - If the user has a packet to transmit or retransmit (e.g., due to a received or missing ACK/NACK or a newly generated packet), go to Step 1.
 - Otherwise, go to Step 3.
- Step 3: Listen to the channel.
 - If the user detects a primary or a warning message, go to Step 4.
 - If it has a packet to transmit, go to Step 1.
 - Otherwise, stay in Step 3.
- Step 4: The secondary user sends/relays the warning signal at the predetermined power level for N times as shown in Figure 1. Go to Step 5.
- Step 5: The user leaves the current band and moves back to the default band.

We notice that Step 2 is a required listening phase after each transmission. In current access protocols, a node needs to listen for its ACK/NACK packet after a transmission. On the other hand, in this procedure, L_s can be set larger to enhance the chance of receiving a warning message, as discussed later. Step 3 is the "idle" listening stage where users listen to the channel when they have no packet to transmit.

In Step 4, a secondary user broadcasts the warning message as node "B" in Figure 1. In the figure, N is the number of warning message transmissions each user should send, L_t is the maximum transmission time of a regular secondary packet, L_s is the amount of time a secondary user listens to the channel, L_p is the prefix transmission time of the warning message, L_w is the transmission time of the warning message, and L_i is the idle interval between two consecutive warning messages from the same secondary user. If a secondary user is listening to the channel during the transmission of the prefix of the warning message and detects the prefix successfully (known as acquisition), the user will listen to the channel for the following L_w period of time expecting a warning message. If it receives the warning message successfully, it will broadcast the warning message. Otherwise, the secondary user returns to its regular state.

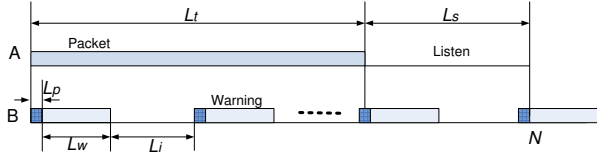


Fig. 1. Transmission of the warning message. Node A transmits a regular secondary packet and Node B is sending the warning message.

A user may miss a warning message for two reasons: First, it is transmitting (a regular secondary packet) and thus cannot receive the warning message. Second, the warning message signal is received but cannot be correctly decoded due to signal propagation loss and interference. Such possibilities need to be accommodated in the protocol design. We address the first issue by repetition (N) plus an enforced listening window (L_s) and the second by appropriate selection of spreading code and transmission power.

Consider that “B” is transmitting a warning message and “A” is a one-hop neighbor of “B”. We have two options in choosing the values of L_s and N , depending on whether or not a warning message can reach a two-hop neighbor with a high success rate. If so, then the first issue can be solved as follows: Suppose “A” is transmitting to a one-hop neighbor, “R”. Because “R” is within two hops of “B”, “R” can receive the warning message. Instead of sending an ACK to “A”, “R” will start repeating the warning message. Therefore, “A” will not receive its ACK but the warning message. In this case, L_s can be set as small as possible. This is the preferred mode because it maximizes the possibility of regular secondary transmissions. In other words, as soon as a node receives its ACK, it can return to the transmission state.

On the other hand, if, due to physical constraints, a warning message may not reach a two-hop neighbor with a high success rate, then we have to set L_s and N to be large enough to receive the warning message. To elaborate, L_s has to be longer than $(L_p + L_w - \epsilon) + L_i + L_p$, where $\epsilon > 0$. The first term is the delay if “A” just missed the prefix of the warning message, the second one, L_i , is the time between two warning messages, and L_p is the prefix length of the warning message. Thus, we have

$$L_s = 2L_p + L_w + L_i \quad (1)$$

In addition, to guarantee that a transmitting secondary has a chance to listen to a warning message, we have

$$L_t + L_s = L_p + (N - 1)(L_p + L_w + L_i) + L_p.$$

To elaborate, “A” will not start its transmission if it receives the prefix of the warning message sent by “B”. After transmitting a packet, “A” will listen to the channel for ACK and for the warning message. Its listening period is L_s . Thus, we have

$$L_t = (N - 2)(L_p + L_w + L_i) + L_p \quad (2)$$

In the above equation, we have two degrees of freedom, N and L_i , while other parameters are fixed. In the extreme case,

we can set $L_i = 0$. In this case, a secondary user will send N warning messages back to back. The delay is minimized in such a case. However, there can be (minor) negative impact: 1) More warning message transmissions imply more cumulative interference to primary users; 2) More simultaneous transmissions of the warning message can gradually decrease the reception success probability of the warning message.

A similar argument enables the system to support sleep-wake scheduling of users. In particular, a user can sleep up to L_t time units and it has to listen to the channel for L_s time units. In this case, a sleep node can be informed of the evacuation with a high probability.

The above design guarantees that a secondary node has a chance to listen for the warning message, and we can further increase the value of N to improve reliability. In Section IV-C, we explain the design of the warning message such that the message can be received with a high success rate.

C. MAC and Routing

The objective of the ESCAPE protocol is to evacuate the channel as fast and reliable as possible. To elaborate, we need to disseminate the warning message “primary-return” to all secondary users as soon as possible. It does not matter where and who start to broadcast the message. It does not matter from whom a user gets the message.

Because of the unique objective and the physical layer property, the MAC of the warning message is simple: a user transmits the warning message as it wishes, which is similar to the ALOHA protocol (but without backoff and retransmission). The overlapping transmissions, which is considered as collision in ALOHA, are handled by spreading code. As discussed in Section IV-C, numerical results show that an m-127 code enables a receiver to successfully detect the warning message in the presence of tens of simultaneous copies of the warning message.

The MAC here is similar to that of the spreading ALOHA. In spreading ALOHA, different users exploit the same spreading code for multiple access. Different users can be distinguished because of the asynchronization of the transmission. When the number of users increase, it is more likely that the transmissions of two or more users are synchronized in a chip-level and thus cannot be distinguished, which results in collision. In comparison, in ESCAPE, different users transmit the verbatim copies of the warning message. Asynchronized transmissions are processed by spreading code. Chip-level synchronized transmissions, which happens by coincidence (since we do not synchronize users on purpose), actually benefit the reception because the signal strength of the received warning message is the summation of multiple copies. This simple MAC does not claim any capacity gain but for the purpose of the warning message propagation.

The routing of the warning message is also simple. Basic flooding is assumed. This simplicity again benefits from the auto-correlation property of the spreading code. Flooding also provides reliability enhancement because of its redundancy. If a secondary user misses a warning message, it is highly likely that it will receive the message from other neighboring nodes

who echo and flood the same warning message. Flooding protocols been studied in various scenarios. One focus is to reduce flooding overhead by eliminating (unnecessary) transmissions. Such schemes are orthogonal to ESCAPE and can be combined in the ESCAPE to reduce the number of transmissions. On the other hand, we stay with the most simple form of flooding for simplicity and for reliability caused by redundancy in this paper.

It is possible that more than one user detect the return of primaries and start the warning messages. This is similar to flood the area from several locations with the same message. It will speed up the evacuation process; i.e., the propagation of the warning message. A user will not resend the warning message if it has already done so.

IV. PROTOCOL DESIGN

ESCAPE is designed for the purpose of fast and reliable evacuation. This is different from typical protocol designs where capacity, fairness, and coexistence are critical. From the perspective of primary users, the performance metrics of ESCAPE include evacuation time, peak interference, average interference, and evacuation failure probability. The metrics for secondary users include evacuation time and false alarm rate. Thus, design parameters of ESCAPE are spreading code length, transmission power of the warning message, message repetition time (N), and warning message detection threshold given performance constraints of primary and secondary users. We report our design of ESCAPE parameters in the following.

A. Evacuation Time Constraint

Consider the case where evacuation time is the constraint set by primary users. Assume that the channel is expected to be vacated in T_E time units. Assume that the primary transmitter is far away from secondary users and thus the average receiving power of the primary at all secondary users are the same, denoted as P_{ps} . Let the chip length be L_c and the number of symbols of the warning message be M . The average receiving power of other one-hop secondary users is P_{ss} . It is desirable that the transmission power of the warning message to be at least at the same level as other secondary transmissions. This selection is both physically feasible and desirable. If other secondary transmissions are narrow-band, then the transmission of the warning message is likely to cause collision, which is desirable to stop the regular secondary transmissions. On the other hand, if other secondary transmission are spread spectrum, we would like the warning message to use a spreading code at least as long as and the transmission power as high as other secondary transmissions. Given such information, the spreading gain can be determined. If the chosen spreading gain does not satisfy the T_E constraint, we choose the longest code that satisfies that T_E constraint and select power level such that the primary interference can be efficiently suppressed.

B. Peak Interference Constraint

Consider the case where peak interference is the constraint set by primary users. The worst case peak interference happens where all secondary users transmit the warning message

simultaneously. This occurs in two scenarios. The first is when all secondary users detect the primary and start to transmit the warning message. The second is that different secondary users may repeat the warning message at a certain time. In this case, the maximum transmission power of the warning message is determined. Based on this information and the power of the primary and other secondary, we select the code that has sufficient interference tolerance capability.

We note that information such as P_{ps} , P_{ss} can be estimated in general. For instance, consider the scenario when unlicensed WLAN devices use unoccupied TV broadcast channels. In this case, information on the transmission power, the location of the TV tower, the service contour, the transmission power of a WLAN device, the rough density of a WLAN are all available. Such information can be used to estimate P_{ps} and P_{ss} and determine the code and transmission power level. When the information is less accurate, the design needs to be more conservative. Note that code selection should be done infrequently since such information is needed BEFORE users operate in opportunistic bands.

C. Warning Message Detection

In ESCAPE, the warning message is sent as CDMA using a predefined spreading code. The transmission of the warning message is subject to 1) interference from both primary and other regular secondary transmissions; and 2) interference from other (unsynchronized) copies of the warning message due to the design of MAC and routing. Therefore, the spreading code needs to provide a good spreading gain and superior auto-suppression capability. The technique of spreading has been studied extensively in the field of communication. We explore the following properties in the ESCAPE protocol.

We require the spreading code to have a good auto-correlation property so that a user can detect the “warning” signal even when the transmissions of the warning message at multiple secondary users are overlapping. A secondary user may lock onto any one of the transmissions by treating others as interference. A good auto-correlation ensures a good reception probability under multiple transmissions. The properties of various spreading codes have been studied extensively in the field of CDMA communications. We choose m-sequence as the spreading code for ESCAPE because of its superior auto-correlation property as shown in the Appendix. Numerical results show that an m-127 code can sustain tens of simultaneous copies of the warning message with little performance degradation.

CDMA spreading also provides interference tolerance against transmission of primary and other secondary users. For example, the m-127 spreading code provides a spreading gain of more than 20dB. The detection and false alarm rate are discussed in the Appendix. We also note that the false alarm rate can be kept very low by appropriately tuning parameters.

If a Rake receiver is available, the interference and noise suppression capability can be improved significantly. In this case, multiple overlapping transmissions of the warning message can be exploited to significantly improve the performance. (All numerical results in the paper assume no Rake receiver to be conservative.)

Note that the transmission power of the spread warning message is not necessarily low compared to the transmission power of a regular packet. In addition, the warning message can use the same band as a regular packet instead of a wide band as in a CDMA system. In other words, a chip length in the warning message can be as large as a bit length in a regular packet. In this case, the effect of spreading code is to make the warning message longer instead of over a wider band. We spread the warning message for the purpose of interference tolerance and auto-suppression.

We wish to clarify that we do not propose any new spreading or signal processing techniques. Our purpose is to exploit the physical capabilities of a spread warning message to design a protocol for fast and reliable channel evacuation in spectrum-agile communication networks.

V. DELAY AND FAILURE PROBABILITY ANALYSIS

When a secondary user starts to transmit the warning message, its neighbor may be in transmission and thus miss the warning message. (We do not assume two-hop reception with high success rate.) This introduces randomness in the delay and potential failure of the evacuation process. In this section, we analyze the average delay and failure probability for a one-hop transmission of the warning message. We consider two cases where the packet length of a regular secondary transmission is fixed and is exponentially distributed.

With some abuse of the terminology, we call the time between the end of the enforced listening period and the beginning of the next regular secondary transmission *vacation time*. We assume that the vacation time is exponentially distributed with mean $L_v = 1/\lambda$. Because the prefix length is relatively short in comparison to the warning packet length, we ignore it here in the analysis for simplicity. We set the enforced listening window $L_s = L_w + L_i$. Let L_t be the average packet length of a regular secondary transmission. Before it receives the warning message, a secondary user can be in one of the three states: transmission (t), listening (s), and vacation (v). Let p_x be the probability that a secondary user is in state x . We have

$$p_x = \frac{L_x}{L_t + L_s + L_v}, \quad x \in \{t, s, v\}.$$

Let N be the number of repetitions. We consider delay D which is defined as the time before the user begins to receive the warning message. We analyze the first and second moments of D for fixed and exponentially-distributed packet length.

There are two reasons for delay: 1) the secondary user may be in transmission and thus cannot receive until its transmission ends; or 2) the secondary user fails to detect the warning message due to interference. As discussed earlier, the warning message can be received with a high success rate through appropriate code and power selections. Therefore, we ignore the receiving failure probability in the delay analysis. Let P_i be the probability that the receiver misses the first i warning messages due to its own transmission. When the

regular secondary packet has a fixed length L_t , we have

$$P_i = \begin{cases} p_v + p_s & i = 0 \\ p_t \frac{L_w + L_i}{L_t} & 1 \leq i \leq n' \\ p_t \frac{\Delta}{L_t} & i = n' + 1, \end{cases}$$

where $n' = \lfloor L_t / (L_w + L_i) \rfloor$ and $\Delta = L_t - n'(L_w + L_i)$.

When the packet is exponentially distributed with mean $L_t = 1/\mu$, we have $P_0 = p_v + p_s$ and

$$P_i = \exp(-(i-1)\mu(L_w + L_i)) - \exp(-i\mu(L_w + L_i)),$$

for $i = 1, \dots, N-1$. In addition, the user will miss all copies of the warning message with probability P_∞ where $P_\infty = \exp(-(N-1)\mu(L_w + L_i))$.

Therefore, the first and second moment of one hop delay in the case of fixed packet length is

$$\begin{aligned} E(D_f) &= L_w + \sum_{i=0}^{N-1} i(L_w + L_i)P_i \\ &= \frac{(L_w + L_i)p_t(n' + 1)}{L_t} \left(\frac{n'(L_w + L_i)}{2} + \Delta \right) \\ E(D_f^2) &= (L_w + L_i)^2 p_t \frac{n' + 1}{L_t} \cdot \\ &\quad \left(\frac{n'(2n' + 1)}{6} (L_w + L_i) + (n' + 1)\Delta \right) \end{aligned} \quad (3)$$

In the case of exponentially distributed packet length, conditioning on the user receives the warning message, we have

$$\begin{aligned} E(D_e) &= \frac{(L_w + L_i)p_t}{1 - P_\infty} \left(1 - Na^{N-1} + \frac{a(1 - a^{N-1})}{1 - a} \right) \\ E(D_e^2) &= \frac{1}{1 - P_\infty} (L_w + L_i)^2 p_t \cdot \\ &\quad (-1 - N^2 a^{N-1} + 2 \frac{1 + Na^{N+1} - a^N - Na^N}{(1 - a)^2} \\ &\quad - \frac{a(1 - a^{N-1})}{1 - a}), \end{aligned} \quad (4)$$

where $a = \exp(-\mu(L_w + L_i))$.

We next analyze the probability that the secondary user fails to receive the warning message of its one-hop neighbor. Let q be the probability that the warning message is not received due to interference. Let Z_i be the probability that the i th warning message is not received. We have

$$\begin{aligned} P_{fail} &= P(\cap_{i=1}^N Z_i) \\ &= P(Z_1) \prod_{i=2}^N P(Z_i | \cap_{j=1}^{i-1} Z_j) \\ &\approx P(Z_1)P(Z_2|Z_1)P(Z_3|Z_1Z_2)P(Z_i|S_{i-1} = t), \end{aligned}$$

where S_{i-1} is the state of the secondary user at the beginning of the $(i-1)$ th warning message and $P(Z_i|S_{i-1} = t)$ is the probability that the i th warning message is not received given state $S_{i-1} = t$. In the above equation, we note that events Z_i s are not independent. Therefore, for a large value of N , we use the approximation presented in the last step. For a large value of N , in the case of fixed packet length, the failure probability is extremely low when enforced listening window is $L_s = L_w + L_i$ and q small. Therefore, we focus the

analysis on the case where the packet length is exponentially distributed. Due to page limit, we omit the steps and present the result here. We have

$$\begin{aligned}
P(Z_1) &= p_t + (p_v + p_s)q \\
P(Z_2, Z_1) &= p_t(p_{tt} + p_{vt}q + p_{st}q) \\
&\quad + \sum_{y=\{v,s\}} p_y q(p_{ty} + p_{vy}q + p_{sy}q) \\
P(Z_3, Z_2, Z_1) &\approx \sum_{y=\{v,s,t\}} (p_{ty} + p_{vy}q + p_{sy}q) \\
&\quad P(S_2 = y, Z_2, Z_1) \\
P(Z_i | S_{i-1} = t) &= p_{tt} + p_{vt}q + p_{st}q,
\end{aligned}$$

where $L_c = L_w + L_i$, and

$$\begin{aligned}
p_{tt} &= \exp(-\mu L_c) \\
p_{st} &= 1 - \exp(-\mu L_c) \\
p_{vt} &= 0 \\
p_{ts} &= \frac{1}{\mu L_c} + \frac{\mu \exp(-\lambda L_c) - \lambda \exp(-\mu L_c)}{\mu L_c (\lambda - \mu)} \\
p_{vs} &= \frac{1 - \exp(-\lambda L_c)}{\lambda L_c} \\
p_{ss} &= 1 - p_{ts} - p_{vs} \\
p_{tv} &= \lambda \frac{\exp(-\mu L_c) - \exp(-\lambda L_c)}{\lambda - \mu} \\
p_{vv} &= \exp(-\lambda L_c) \\
p_{sv} &= 1 - p_{tv} - p_{vv}.
\end{aligned}$$

VI. SIMULATION

In this section, we present simulation results of the ESCAPE protocol. We consider two types of topologies: a 5×5 regular grid and a 25-node random network. We consider interference from primary users, other secondary transmission, multiple copies of the warning message, and background noise. The transmission power of the warning message is the same as the transmission power of a regular secondary message (unless otherwise specified). We assume, at secondary users, the received power from the primary transmission are 3dB higher than the power from the nearest secondary transmission. The power pathloss exponent is 4 and the detection threshold is 21 dB above the background noise level. We consider a short spreading code (for fast simulation). We consider an m-15 spreading code with a symbol length of 4. Therefore, each warning message is 60-bit long. The prefix is set to be 6 bits. If a secondary user detects a prefix successfully, it suspends its transmission (if any) and listens to the channel for the duration of the warning message. Idle interval between two consecutive transmissions of the warning message is set to be 10 bits. A small uniformly distributed random delay with mean 5-bit is introduced before a secondary user starts to forward the warning message. This delay is introduced to model the random processing time of secondary users. In the simulation, it avoids perfect synchronization of the warning message among secondary users (which will be too optimistic).

We first run simulations on a 5×5 grid for 1000 iterations. We simulate two cases: fixed packet length and exponentially

distributed packet length. In both cases, we set $N = \{4, 9\}$, average packet length 200 bits, and a secondary user is transmitting 49.6% of time (including the time for enforced listening). The warning message starts to propagate from the node in the left-upper corner. This represents the worst case performance in terms of delay and failure probability. In the case of fixed packet length, in all iterations, all 25 nodes leave the system; i.e., evacuation failure does not occur.

In the case of exponentially distributed packet length, we notice 20% and 1.2% of evacuation failure for $N = 4$ and $N = 9$, respectively. An evacuation fails if one or more secondary users does not receive the warning message. This is mainly due to the existence of large transmission packets. We note that a packet length exceeds the length of $(N - 1)$ warning messages with probability 35% and 6% for $N = 4$ and $N = 9$, respectively. Because we expect a maximum packet length in practical wireless systems, evacuation failure will occur much less frequently. To improve reliability, we could increase the value of N .

Figure 2 is the evacuation delay histogram for the exponential case with $N = 9$. In the figure, the x-axis is the evacuation time normalized over the average packet length (200 bits). The actual time unit depends on the physical layer. For instance, if the data rate is 11Mbps, each chip-length is $0.9\mu s$. The evacuation time is around 1-2ms. The y-axis is the number of times that the delay occurs in 1000 iterations. We observe a few clusters. The delay histogram of other cases are similar. As a benchmark, we note that the normalized delay is 3.75 in the most optimistic case, where the first warning message is heard by all other nodes.

In all cases, we notice that with a large chance, the last node receives the warning message before the first node finishes its N transmissions in the 5×5 grid. Therefore, the peak interference usually occurs when all nodes are transmitting warning messages. In Figure 3, we plot the transmission power of the 5×5 grid in one iteration with exponentially distributed packet length and $N = 9$. We show two curves. The solid curve is the total transmission power including both warning message and other secondary transmissions. The y-axis is the transmission power, its exact value depends on the physical layer. For instance, if the transmission power of each user is 100mW, the maximum value in the y-axis (250) represents 2.5W. The dashed curve is the total emission power by the transmissions of warning messages. We note that in the beginning, other transmission power dominates and after 800 time units (the unit is chip-length), only warning messages are transmitted. In this case, the peak interference (250) is due to a combination of warning message transmission and regular secondary transmission (e.g, at 400 time units).

We next show the impact of N and P_w on average evacuation delay and failure probability. Each simulation is run 100 times. In Figures 4 and 5, we show the average delay (normalized over the average packet length of 200 bits) and failure probability as a function of N . As N increases, average delay increases and the failure probability decreases. In Figures 6 and 7, we show the average delay (normalized over the average packet length of 200 bits) and failure probability as a function of P_w . The x-axis is P_w normalized over P_s . As P_w

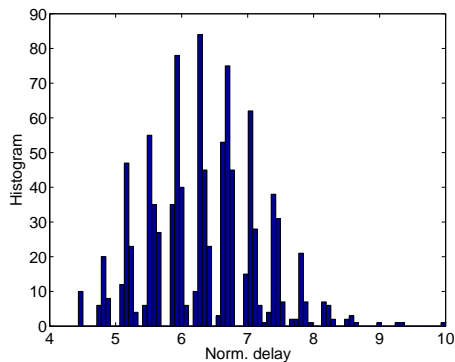


Fig. 2. Delay histogram of a 5×5 grid with exponential packet length.

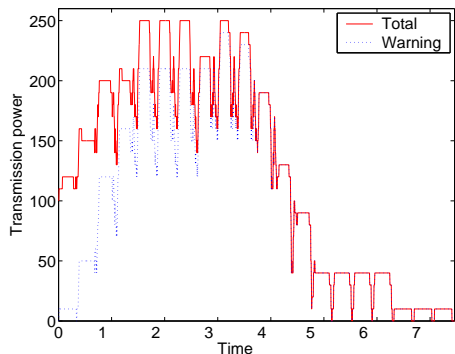


Fig. 3. Interference of a 5×5 grid with exponential packet length. The x-axis is time normalized over the average packet length of 200 bits.

increases, both the failure probability and delay decreases. The gain is significant when P_w is relatively small and diminishes as P_w increases. We note that increasing P_w along is not enough to totally eliminate evacuation failure, partially due to the existence of extreme long secondary packets. The results for a random network are similar and thus omitted here.

In summary, simulation results demonstrate that the proposed ESCAPE protocol can reliably evacuate secondary users in the presence of interference from primary and secondary transmissions. In addition, our simple MAC and routing schemes work well due to the superior auto-suppression prop-

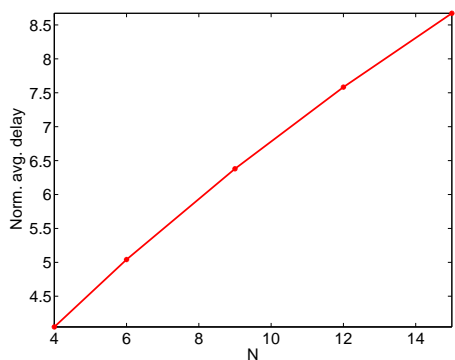


Fig. 4. Normalized average delay in a 5×5 grid with exponential packet length as a function of N .

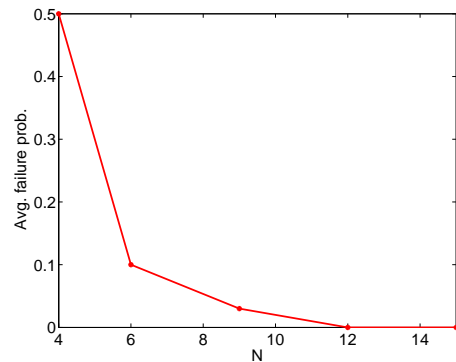


Fig. 5. Failure probability in a 5×5 grid with exponential packet length as a function of N .

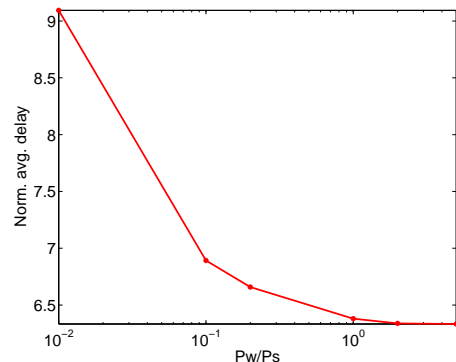


Fig. 6. Normalized average delay in a 5×5 grid with exponential packet length as a function of P_w .

erty of m-sequence codes.

VII. RELATED WORK

We note that the access scheme used in our protocol is different from spreading ALOHA. In spreading ALOHA, different users use the same spreading code to convey their *individual* messages [1], [17]. Users can be distinguished due to the asynchronous transmissions of different users. The performance of spreading ALOHA is limited by chip-level collision. In other words, if one receives two or more packets with delay difference within a chip-interval, the packets collapse and cannot be recovered. As the number of users increases, so does the collision probability. In comparison, because different users are sending the SAME warning message in our scheme, if chips synchronize, it will indeed benefit the reception by providing a stronger signal, to which a receiver is more likely to detect. In other words, such “synchronization” benefits the reception of the warning message. Numerical results indicate that the access scheme can tolerate a large number of simultaneous transmissions of the warning messages. Our scheme is also different from traditional CDMA systems where different communications use different spreading code.

A few MAC protocols have been discussed in the context of cognitive radio. In [2], [10], dedicated control channels are proposed for secondary users. Our protocol applies in-band signaling and does not require a dedicated control channel. Of

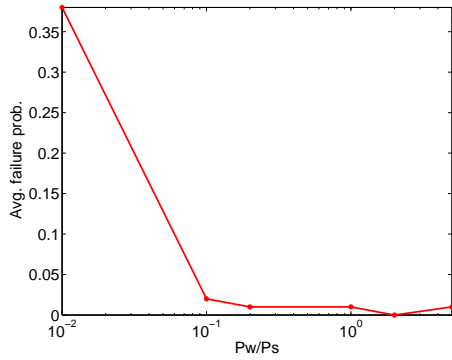


Fig. 7. Failure probability in a 5×5 grid with exponential packet length as a function of P_w .

course, the control channel can benefit many other purposes, such as resource sharing among secondary users, which is not among the functionality of ESCAPE. In this sense, the proposed scheme can be combined with the proposed protocols.

Many MAC protocols have been proposed in the literature that focus on resource sharing among users. These protocols can be potentially used by secondary users. One particular challenge is to distinguish among primary transmission and secondary transmissions. If one can detect primary transmission, our protocol can help solving the problem. In other words, only the “warning message” indicates the return of primaries. We note that the ESCAPE is complimentary these MAC protocols.

Many flooding schemes have been proposed in the literature, many with the focus on overhead reduction. We first note that “overhead” can cause significant delay in tradition MAC schemes, e.g., CSMA/CA, especially when the network is dense. In other words, when a large number of nodes try to access the channel, congestion occurs which results in delay and packet loss. On the other hand, multiple transmissions will not cause serious concern in our protocol. An m-sequence of 127 chips with simultaneous transmissions of tens of users can cause little degradation in the performance without the presence of other transmissions. Furthermore, we should note that overhead reduction schemes proposed in the literature can be applied in combination with ESCAPE to reduce interference to primaries and to enhance the transmission success probability. In this paper, we ignore overhead reduction to emphasize the characteristics of the proposed ESCAPE and also for simplicity.

Cognitive radio has the ability to sense and learn from the environment and adapt to appropriate frequency, power, and transmission schemes. It has attracted a lot of research interests (e.g., [9], [8], [2], [6], [3]). Research efforts include spectrum pooling (e.g., [16], [11]), game-theoretic analysis (e.g., [12], [15], [7]), channel sensing and detection (e.g., [14], [13]), dynamic spectrum sharing (e.g., [10], [4]), etc. To the best of our knowledge, this paper is the first attempt to design an in-band signaling protocol for channel evacuation.

VIII. CONCLUSION AND FUTURE WORK

Channel evacuation is an important yet unique issue in frequency-agile communication networks. In this paper, we present the ESCAPE protocol for channel evacuation of secondary users. The objective of the protocol is to disseminate evacuation information among secondary users fast and reliably with minimum requirement on topology information, network synchronization, and routing maintenance. The ESCAPE protocol is based on joint considerations of physical, MAC, and routing layers. In the physical layer, a secondary user that detects the presence of primary user(s) sends a predefined warning message that declares “primary-active”. The message is spread using a predefined spreading code. Other secondary users hearing the message will abort their own transmissions and send a verbatim copy of the warning message “primary-active” as soon as the message is received. We choose an m-sequence code for spreading because of its superior auto-correlation characteristics so that MAC and routing are significantly simplified. In addition, the spreading code provides sufficient processing gain for interference tolerance. Numerical results indicate that the protocol performs well.

There are a few issues that need to be further studied. The current work includes an average delay and failure probability analysis for a single hop. We are extending it to the multi-hop delay and failure probability. In particular, the tail distribution of end-to-end delay is closely related to evacuation time and needs to be carefully examined. In addition, we plan to investigate information authentication and verification issues. First, the current scheme is prone to malicious user abuse, virus attacks, and false alarms. Indeed, a malicious user can broadcast an evacuation message without detecting the return of primary user(s). Second, a secondary user falsely detects the primary user or warning message can cause unnecessary evacuation. We plan to address the issues by including authentication schemes that verify the identity of the initial warning message. The authentication and identification process can be used to reduce false evacuation. For instance, the system is less prone to false evacuation if it evacuates when two or secondary users corroborate the detection of primaries. Other issues to be considered include theoretical and numerical comparison of the ESCAPE protocol with other broadcasting schemes.

IX. APPENDIX: WARNING DETECTION ANALYSIS

As discussed earlier, the spreading code needs to provide a good spreading gain and superior auto-suppression capability to tolerate interference from other transmissions and to simplify the design of MAC and routing schemes. We choose m-sequence codes for its superior auto-correlation properties over other spreading codes (such as Gold codes).

Consider an m-sequence of length 127 for illustration purpose. The m-127 spreading code provides a spreading gain of more than 20dB and superior auto-suppression property. In Figure 8, the left subplot shows the auto-correlation of an m-sequence code of length 127. In this figure, the x-axis is the shift of chips and the y-axis is the autocorrelation value. The right subplot is the auto-correlation of a random 127-bit code. Note that the m-sequence code has a much

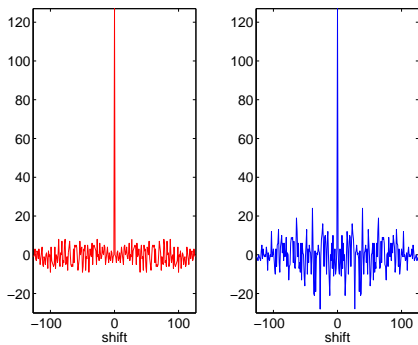


Fig. 8. Compare the autocorrelation functions of an m-sequence code and a random sequence with length 127.

better auto-correlation property than a random 127-bit sequence code. This property is exploited to handle simultaneous transmissions of the same warning message and to simplify MAC and routing issues. Numerical results show that a 127 m-sequence can sustain tens of simultaneous copies of the warning message with little performance degradation.

Next, we discuss the detection and false alarm issue of a spread warning message. Consider a warning message with symbols $W = [w(1), w(2), \dots, w(M)]$, where M is the number of symbols in the warning message and is predefined. Let $C = [c(1), c(2), \dots, c(l_c)]$ be the spreading code used. WLOG (Without Loss of Generality), assume that a receiver detects the prefix and thus is synchronized to the 0th copy of the warning message. The received signal is then

$$\begin{aligned} y(m, k) &= \sqrt{P_w(0)}w(m)c(k) \\ &+ \sum_{i=1}^{N_w} \sqrt{P_w(i)}w(m')c(m + \delta_i) \\ &+ \sqrt{P_p}s_p(m, k) + \sum_{i=1}^{N_s} \sqrt{P_s(i)}s_s(m, k) + n(m, k), \end{aligned}$$

where $P_w(i)$, P_p , $P_s(i)$ are received power of the i th copy of the warning message, the primary user, and the i th secondary transmission, respectively, δ_i is the shift of the i th copy of the warning message. In the above equation, the first term is the signal of the warning message to be detected. The second term represents multiple copies of the warning message with a random shift with respect to the 0th copy. The random shift is due to multipath and other unsynchronized transmissions of the warning message. The third term is the interference caused by primary users, and the fourth is the interference caused by other secondary transmissions. The last one represents zero-mean white Gaussian noise with variance σ_0^2 .

We consider a special case where the second term is zero and both the primary and other secondary transmissions are independent of the transmission of the warning message and the spreading gain over these transmissions is one. In other words, the warning message use the same spectrum as primary and other secondary transmissions. In this case, we can approximately calculate the detection and false alarm rate of the warning message using the central limit theorem. We assume that the bit streams of primary users are independent;

i.e., $s_p(k, m)$ s are independent. By correlating the received signal with $w(k)c(m)$, we have

$$\begin{aligned} y_0 &= \sum_{m=1}^M \sum_{k=1}^{l_c} y(m, k)w(m)c(k) \\ &= \sqrt{P_w}l_cM + \sqrt{P_p} \sum_{m=1}^M \sum_{k=1}^{l_c} s_p(m, k)w(m)c(k) \\ &+ \sum_{i=1}^{N_s} \sqrt{P_s(i)} \sum_{m=1}^M \sum_{k=1}^{l_c} s_s(m, k)w(m)c(k) \\ &+ \sum_{m=1}^M \sum_{k=1}^{l_c} n(m, k)w(m)c(k) \\ &\approx \sqrt{P_w}l_cM + N_a \end{aligned}$$

where N_a is the aggregated interference. It is approximated by a zero-mean Gaussian variable with variance $(P_p l_c M + \sum_{i=1}^{N_s} P_s(i) l_c M + \sigma_0^2 l_c M)$. Let D_{th} be the detection threshold. Set $D_{th} = p_{th} \sqrt{P_w} l_c M$. We have

$$\begin{aligned} P_{det} &= P(y_0 \geq D_{th}) \\ &= P(\sqrt{P_w} l_c M + N_a \geq D_{th}) \\ &= P(N_a \leq \sqrt{P_w} l_c M - D_{th}) \\ &= P\left(N_0 \leq \frac{\sqrt{P_w} l_c M - D_{th}}{\sqrt{l_c M} \sqrt{P_p + \sum_{i=1}^{N_s} P_s(i) + \sigma^2}}\right) \\ &= 1 - Q\left(\frac{\sqrt{P_w} l_c M (1 - p_{th})}{\sqrt{P_p + \sum_{i=1}^{N_s} P_s(i) + \sigma^2}}\right), \end{aligned} \quad (5)$$

where N_0 is a normalized Gaussian random variable. The corresponding false alarm rate for the threshold is calculated as follows:

$$\begin{aligned} P_{FA} &= P(N_a \geq D_{th}) \\ &= P\left(\bar{N}_a \geq \frac{p_{th} \sqrt{P_w} l_c M}{\sqrt{P_p + \sum_{i=1}^{N_s} P_s(i) + \sigma^2}}\right) \\ &= Q\left(\frac{p_{th} \sqrt{P_w} l_c M}{\sqrt{P_p + \sum_{i=1}^{N_s} P_s(i) + \sigma^2}}\right), \end{aligned} \quad (6)$$

We note that the energy per message is $P_w l_c M$. Therefore, for fixed energy per message, the detection and false alarm rate for the same value of p_{th} is fixed regardless of spreading code length. For fixed energy per message, the longer the code length, the lower the transmission power, but the longer the length of the warning message. Using Eqs. (5) and (6), we can determine the appropriate threshold, transmission power, and code/message length.

We should note that the false-alarm rate must be kept very low. There are many possible measures. First, by choosing an appropriate transmission power and code length. Second, by choosing a good detection threshold. For instance, to keep the false-alarm rate at 10^{-8} , we have $p_{th} = 0.6099$ and $P_{det} = 0.9998$ when $P_p = 2$, $P_w = P_s = 1$, $N_s = 4$, $\sigma^2 = 0.01$, $l_c = 127$, and $M = 4$. Note that we can also

choose different thresholds to lower the false alarm rate. We choose one (higher) threshold for a user to rebroadcast the warning message. Another (lower) threshold for a user to suspend its regular transmission. It can either listen to the channel or send probing message asking whether a warning message was sent. It will only broadcast if the received signal is above the first (higher) threshold.

REFERENCES

- [1] N. Abramson. Fundamentals of packet multiple access for satellite networks. *IEEE Journal on Selected Areas in Communications*, 10(2):309 – 316, 1992.
- [2] R. Brodersen, A. Wolisz, D. Cabric, S. Mishra, and D. Willkomm. CORVUS: a cognitive radio approach for usage of virtual unlicensed spectrum. White Paper, 2004.
- [3] D. Cabric, S. M. Mishra, and R. W. Brodersen. Implementation issues in spectrum sensing for cognitive radios. In *the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers*, volume 1, 2004.
- [4] R. Etkin, A. Parekh, and D. Tse. Spectrum sharing for unlicensed bands. In *the 1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DYSPAN)*, Baltimore, MD, Nov. 2005.
- [5] FCC. Unlicensed operation in the TV broadcast bands, ET Docket No. 04-186; additional spectrum for unlicensed devices below 900 mhz and in the 3 ghz band, ET Docket No. 02-380, FCC 04-113.
- [6] S. Haykin. cognitive radio: brain-empowered wireless communications. *IEEE Journal on Selected Areas in Communications*, 23(2):201 – 220, 2005.
- [7] J. Huang, R. Berry, and M. Honig. Spectrum sharing with distributed interference compensation. In *the 1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DYSPAN)*, 2005.
- [8] J. Mitola III. Cognitive radio: making software radios more personal. *IEEE Personal Communications*, 6(4):13 – 18, 1999.
- [9] Joseph Mitola III. Cognitive radio for flexible mobile multimedia communications. In *Sixth International Workshop on Mobile Multimedia Communications*, 1999.
- [10] Xiangpeng Jing and D. Raychaudhuri. A spectrum etiquette protocol for efficient coordination of radio devices in unlicensed bands. In *Proceedings of PIMRC 2003*, Beijing, China, 2003.
- [11] T. Weiss; J. Hillenbrand A. Krohn; F. Jondral. Mutual interference in ofdm-based spectrum pooling systems. In *IEEE 59th Vehicular Technology Conference*, volume 4, 2004.
- [12] J. Neel, R. M. Buehrer, B. H. Reed, and R. P. Gilles. Game theoretic analysis of a network of cognitive radios. In *The 2002 45th Midwest Symposium on Circuits and Systems*, volume 3, 2002.
- [13] S. Shankar, C. Cordeiro, and K. Challapali. Spectrum agile radios: Utilization and sensing architectures. In *the 1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DYSPAN)*, Baltimore, MD, Nov. 2005.
- [14] R. Tandra and A. Sahai. Fundamental limits on detection in low snr under noise uncertainty. In *WirelessCom 05 Symposium on Emerging Networks, Technologies and Standards*, Maui, Hawaii, June 2005. IEEE.
- [15] R. W. Thomas, L.A. DaSilva, and A. B. Machenzie. Cognitive networks. In *the 1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DYSPAN)*, 2005.
- [16] Timo Weiss and Friedrich Jondral. Spectrum pooling: An innovative strategy for the enhancement of spectrum efficiency. *IEEE Communications Magazine*, 42(4), 2004.
- [17] A. Yener and R. Yates. Multiuser access capacity of packet switched CDMA systems. In *Proceedings of IEEE Vehicular Technology Conference VTC*, 1999.