**Cooperative Spectrum Sensing Using Q-Learning with Experimental Validation**

Zhe Chen, and Robert C. Qiu, Tennessee Technological University

**Introduction**

Spectrum sensing detects the availability of the radio frequency spectrum and it is essential to cognitive radio. Cooperative spectrum sensing conducted among multiple secondary users (SUs) has the potentials for solving the hidden terminal problem and achieving a better performance.

In this paper, an approach for cooperative spectrum sensing based on Q-learning is proposed. Real-world Wi-Fi signals measured from four different locations simultaneously are employed to evaluate the performance of the proposed approach. Experimental result shows the performance of the proposed approach is better than that of the popular M-out-of-N rule.

**Cooperative Spectrum Sensing**

N secondary users:

\[ SU_1, SU_2, \ldots, SU_N \]

**Time slot**

- Time slot \( t \)
- Frequency band \( f \)
- Decision \( d \)
- Channel state:
  - “busy” \( d = 1 \)
  - “idle” \( d = 0 \)

**M-out-of-N rule**: ChannelState = \( \begin{cases} \text{busy} & \text{if } \sum_{i=1}^{N} d_{i,t} \geq M \\ \text{idle} & \text{otherwise} \end{cases} \)

**Q-Learning**

A Q-table \( Q(s,a) \) is a matrix indexed by state \( s \) and action \( a \), which is the expected discounted reinforcement of taking action \( a \) in state \( s \). At each time, an agent is assumed to be in a certain state \( s \), and it chooses an action \( a \) according to the Q-table and other algorithms to interact with the environment. Then the agent receives a reward \( r \) from performing action \( a \) and observes a new state \( s' \). After that, the Q-table is updated by the following rule:

\[
Q(s,a) = (1 - \alpha)Q(s,a) + \alpha[r + \gamma \max_{a'}(s',a')] \quad (2)
\]

where \( \alpha \) is the learning rate, and \( \gamma \) is the discount factor.

**Proposed Cooperative Spectrum Sensing Using Q-Learning**

Assume \( N \) SUs sense the spectrum cooperatively. Each SU senses the spectrum and outputs one-bit information \( d_{i,t,f} \) that represents the channel state of frequency band \( f \) during time slot \( t \), where \( i = 1, 2, \ldots, N \) and \( d_{i,t,f} \in \{0,1\} \). Let “0” represent the “idle” channel state and “1” represent the “busy” channel state. Then \( N \) one-bits can form an integer \( s_{t,f} \) whose value ranges from 0 to \( 2^N - 1 \):

\[
s_{t,f} = \sum_{i=1}^{N} d_{i,t,f} \cdot 2^{i-1} \quad (3)
\]

In cooperative spectrum sensing using Q-learning, let the set of states \( S \) be \( \{0,1,\ldots,2^N - 1\} \), and the set of actions \( A \) be \( \{0,1\} \), where \( s \in S \) and \( a \in A \). An action with index “0” means the channel is in “idle” state and available for SUs, and an action with index “1” means the channel is in “busy” state and unavailable for SUs.

Assigning values to reward \( r \) for Q-learning can be tricky. In the proposed algorithm, \( r \) is assigned as follows:

\[
r = \begin{cases} R_p & a_{t,f} = c_{t,f} \\ R_n & a_{t,f} \neq c_{t,f} \end{cases} \quad (4)
\]

where \( a_{t,f} \) is the chosen action for frequency band \( f \) during time slot \( t \), \( c_{t,f} \) is the actual channel state of frequency band \( f \) during time slot \( t \), \( R_p \) and \( R_n \) are constants.

The proposed Q-learning based approach for cooperative spectrum sensing is shown in Fig. 2, where \( a \) is the output.

1. Let current state be \( s \), where \( s = s_{t,f} \), and \( s_{t,f} \) is calculated by (3) using \( N \) outputs \( d_{i,t,f} \) from \( N \) SUs.
2. Choose an action \( a \) by:

\[
a = \begin{cases} 0 & Q(s,0) \geq Q(s,1) \\ 1 & \text{otherwise} \end{cases} \quad (5)
\]

3. Receive a reward \( r \) from choosing an action \( a \), where the value of \( r \) is given by (4). Denote the resulting state as \( s' \), where \( s' = s_{t+1,f} \) and \( s_{t+1,f} \) is calculated by:

\[
s_{t+1,f} = \sum_{i=1}^{N} d_{i,t+1,f} \cdot 2^{i-1} \quad (6)
\]

4. Update \( Q(s,a) \) using (2).
5. \( t = t + 1 \). Goto step 1.

**Experimental Results**

- Parameters of the proposed approach for performance evaluation are: \( N = 3, f = 2.418 \text{ GHz}, t = 1, 2, \ldots 2000 \), \( R_p = 100 \), \( R_n = -100 \), \( \alpha = 0.1 \), and \( \gamma = 0.1 \).

**Conclusion**

- An approach for cooperative spectrum sensing based on Q-learning has been proposed.
- Performance of the proposed approach has been evaluated using measured real-world Wi-Fi signals.
- Experimental result shows the proposed approach is indeed effective.

**Acknowledgments to National Science Foundation and Office of Naval Research**