Spectrum Sharing Applications

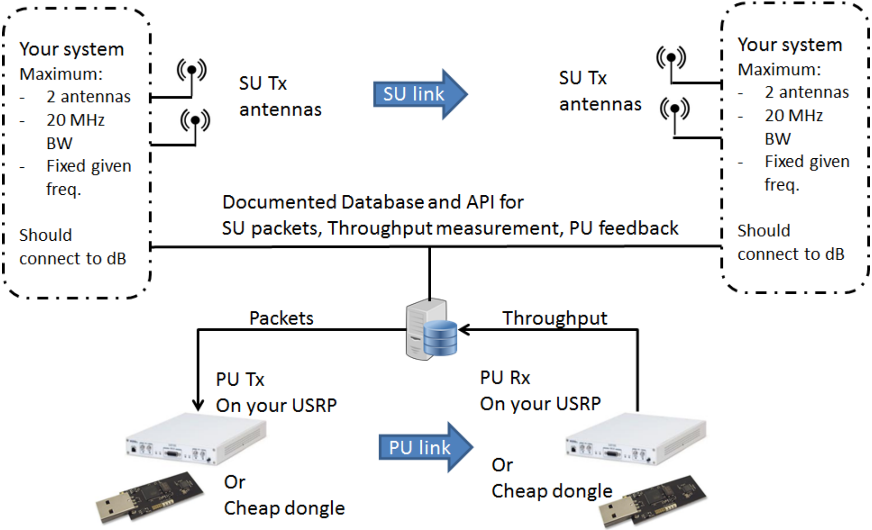
Sreeraj Rajendran

rsreeraj@gmail.com

FOSDEM'15, Brussels
February 1, 2015

Contents

Challenge Setup



Scoring criteria

- Final Score

$$Score = T_{SU} \times S_{PU}$$

$$S_{PU} = max\left(0, \frac{10}{9}T_{PU} - \widehat{T}_{PU}\right)$$

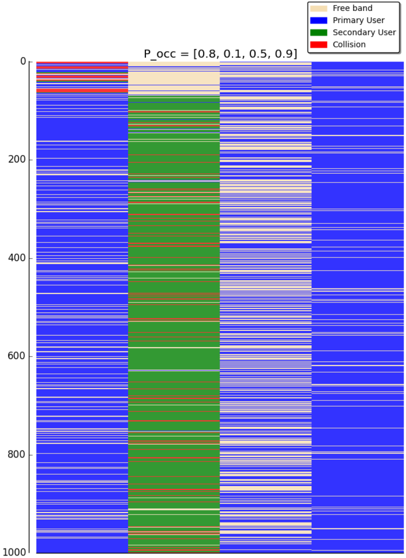
  - T_{SU} - Delivered secondary user throughput
  - S_{PU} - Primary user satisfaction
  - T_{PU} - Delivered primary user throughput
  - $\widehat{T}_{PU}$ - Offered primary throughput

- Objective winner
  - Based on the highest score

- Subjective winner
  - Based on the quality of the paper

- More details: `ieee-dyspan.org`

Simple Scenario (Single Channel selection)



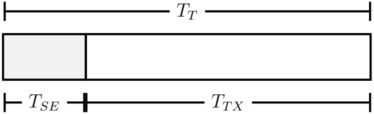
- Assumptions
  - Entire time duration divided into slots
  - Secondary collisions are the only cause for primary throughput reduction
  - Channel occupancy distribution is known
- Objective
  - Maximize SU throughput
- Channel selection
  - Select the channel with minimum occupancy

Problems

- Issues
  - SU lacks the information about the channel
  - SU has to explore the channel to estimate its occupancy
  - Exploration-Exploitation trade-off
- Models
  - Popular multi-armed bandit problems

Upper confidence bound (UCB) based strategies

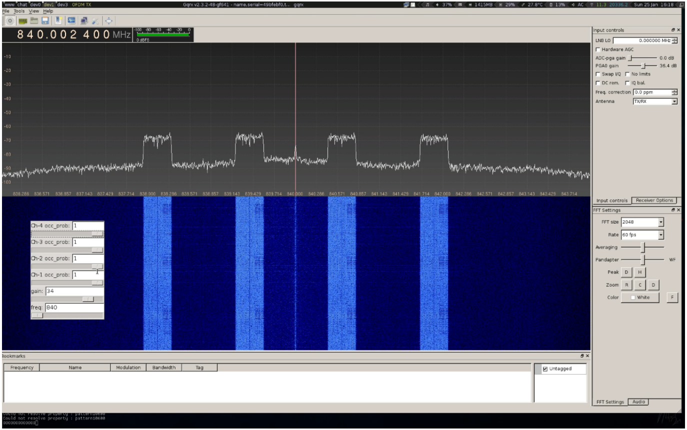
- Sensing and Transmission slot



- Assign positive reward if the channel is sensed free
- Average the reward and calculate an upper confidence bound for the sample mean
- Select the channel based on this UCB

Reference: W. Jouini, D. Ernst, C. Moy, and J. Palicot, "Upper confidence bound based decision making strategies and dynamic spectrum access," in *Proceedings of the IEEE International Conference on Communications (ICC '10)*

Single channel selection Demo (UCB)

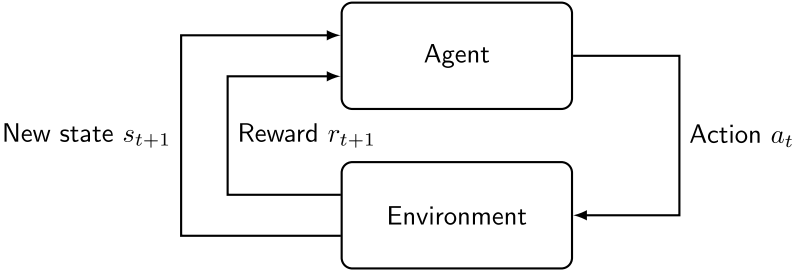


We could do better

- There will be sensing errors
  - Obvious since PU and SU are not synchronized
- Exploit the feedback information (T_{PU}, T_{SU})
- Sensing and transmission slots are fixed
  - Stop sensing if channel is always free

Reinforcement Learning

- A more general framework
  - A discrete set of states, $\mathbb{S}$
  - A discrete set of actions, $\mathbb{A}$
  - A policy π that maximizes the expected reward



Q-Learning

- Most popular model-free algorithm for reinforcement learning
- Learns from delayed reinforcement
- Model
  - Action set: $\{$`sense, transmit, channel_switch` $\}$
  - States, S: $\{0, .., n\}$ where n is the number of available channels
- QL update

$$Q_{t+1}(s, a_t) = Q_t(s, a_t) + \alpha \left(r(s, a_t) + \gamma \max_a Q_t(s, a) - Q_t(s, a_t) \right)$$

α is the learning rate and γ is the discount factor

More details: How to select a channel 's'?

- $Q(s, a_{se})$: Sensing reward

$$r(s, a_{se}) = \begin{cases} 0 & \text{if channel is occupied} \\ 1 & \text{if channel is free} \end{cases}$$

- $Q(s, a_{tx})$: Transmission reward

$$r(s, a_{tx}) = T_{SU} - T_{CO}$$

- $Q(s, a_{sw})$

$$V(s) = Q_t(s, a_{se}) + Q_t(s, a_{tx})$$

$$\widehat{s} = \arg\max_{h \in S} V(h)$$

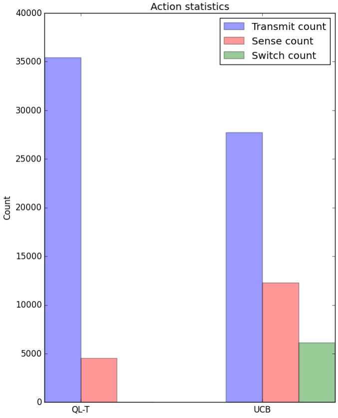
$$Q_{t+1}(s, a_{cs}) = V(\widehat{s}) - V(s)$$

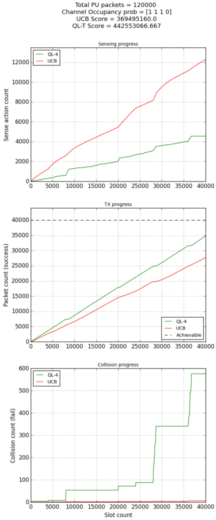
- Soft-max selection policy, $\pi_t(s, a)$

Reference: Marco Di Felice, Kaushik Roy Chowdhury, Andreas Kassler and Luciano Bononi, "Adaptive Sensing Scheduling and Spectrum Selection" in *Proceedings of the 20th International Conference on Computer Communications and Networks (ICCCN '11)*

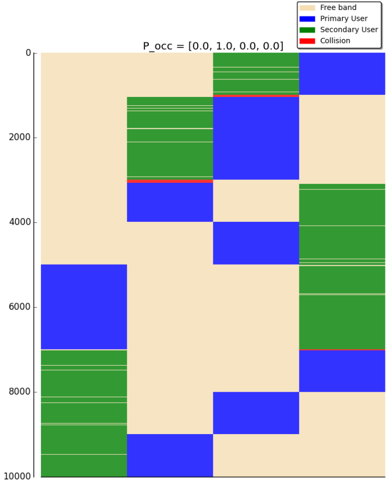
Simulation Results

- After every $4000 \times T_{slot}$ a random channel is made free





We could do better



- Room for improvement
  - Multi-Channel solutions
- Consider PU throughput
  - PU will back-off due to the presence of carrier sense (802.15.4)
  - No intelligence in PU to maximize throughput

Prototyping tools

- GNURadio examples
  - `gnuradio.org`
- OOT modules from Bastian Bloessl
  - `github.com/bastibl/gr-ieee802-15-4`
  - `github.com/bastibl/gr-foo`
- RFNoC
  - `github.com/EttusResearch/uhd/wiki/RFNoC`
- Labview
  - `dyspanchallenge@esat.kuleuven.be`

THANK YOU