Cognitive Anti-jamming Satellite-to-Ground Communications on NASA’s SCaN Testbed

Sudharman K. Jayaweera, Shuang Feng, Abriel Holland, and Christos Christodoulou.

BLUECOM Systems and Consulting LLC, Albuquerque, NM.
ECE Department, University of New Mexico, Albuquerque, NM.

Work performed under NASA STTR contract NNX17CC01C

Dale Mortensen, Marie Piasecki, and Mike Evans

NASA Glenn Research Center, Cleveland, OH.

Presenter: Dale Mortensen
Wideband Autonomous Cognitive Radio (WARC) Architecture

Spectrum Knowledge Acquisition

Cognitive Engine

Cognitive processing

Cognitive multi-mode Communication protocols

Wideband Antennas

Control logic + Down Conversion

LPF

ADC

Reconfigurable Antennas

Control logic + Up Conversion

LPF

DAC
Satellite-to-Ground
Cognitive Anti-Jamming (CAJ)
Communications: Concept of Operations

Frequency shift to avoid interference
Implemented WACR System

WARC operation with two separate SDR modules instead of a single SDR module.
Radiobot Cognitive Engine: CAJ Policy Options

1. Load a pre-learned policy from a file and keep updating the policy during the communications phase.

2. Learn a policy during a training period and keep updating the policy during the communications phase.

3. Learn a policy during a training period and keep it fixed during the communications phase.
CAJ Policy with Reinforcement Learning: Watkin’s Q-Learning Algorithm

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

NOTE: Learning rate (\( \alpha \)) and Forgetting factor (\( \gamma \)) both held constant for this experiment.
Exploration vs Exploitation

\[ \pi_t(s) = \arg\max_a Q_t(s, a) \]

\[ a_t(s) = \begin{cases} 
\pi_t(s) & \text{with probability } 1 - \varepsilon \\
U(A \setminus \{s\}) & \text{with probability } \varepsilon 
\end{cases} \]

- Learned policies can be used with an exploration rate (\(\varepsilon\)) during the communications phase.
  - Allows discovery of possible new optimal (state, action) pairs.
  - Must be balanced with exploitation of the already learned policy.
- Complete exploitation of previously learned policy is obtained setting \(\varepsilon\) to zero.
CAJ with a Random Policy

• Set the exploration rate to unity during communications phase to achieve a random channel selection policy.

• Random channel selection policy does not mean it is a traditional radio.
  – Even when the policy is to select channels randomly, the radio is still a WACR.

• Random policy is used to evaluate the effectiveness of the learning process, not the effectiveness of cognitive communications.
  – To perform anti-jamming communications, even with a random policy, the radio still needs the spectrum knowledge of the cognitive radio.
  – Hence, it is still autonomously mitigating the jammer.
Flight Testing System Configuration

Flight systems
- DVB-S2 Tx
- Tx Control Information
- BPSK Rx Demod
- JPL RADIO
- NASA SCAN TESTBED (on ISS)

Ground systems
- S-BAND DVB-S2
- RF Hardware
- Doppler Compensation
- ML-605 SDR BPSK Tx
- USRP (SDR Platform)
- Spectrum Knowledge Acquisition Module
- Cognitive Anti-jamming Decisions
- Control Signal Synthesis
- Policy Learning

Bluecom Systems RADIOBOT Cognitive Engine
Bluecom Systems RADIOBOT Wideband Autonomous Cognitive Radio (WACR) System
NASA GRC GROUND STATION
Flight Testing
Ground Station Antenna Setup

Over-the-air jammer antenna setup on same rooftop as main ground station.
Flight Testing: Relative Powers of Satellite and Jammer Signals
# Flight Testing Event Data

<table>
<thead>
<tr>
<th>Test #</th>
<th>Jammer Type</th>
<th>Policy Type</th>
<th>Exploration Rate (ε)</th>
<th>Total Number of Sensing Periods During the Complete Event-pass</th>
<th>Total Number of Sensing Periods with Sufficient Signal Quality Between Channel Transitions</th>
<th>Number of Channel Transitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>sweep</td>
<td>random</td>
<td>1.0</td>
<td>214710</td>
<td>29380</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>sweep</td>
<td>random</td>
<td>1.0</td>
<td>218545</td>
<td>96337</td>
<td>77</td>
</tr>
<tr>
<td>3</td>
<td>sweep</td>
<td>pre-learned</td>
<td>0.3</td>
<td>235192</td>
<td>132380</td>
<td>81</td>
</tr>
<tr>
<td>4</td>
<td>sweep</td>
<td>pre-learned</td>
<td>0.3</td>
<td>120370</td>
<td>298</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>Markov</td>
<td>random</td>
<td>1.0</td>
<td>192751</td>
<td>51412</td>
<td>67</td>
</tr>
<tr>
<td>6</td>
<td>Markov</td>
<td>pre-learned</td>
<td>0.0</td>
<td>229520</td>
<td>72908</td>
<td>79</td>
</tr>
<tr>
<td>7</td>
<td>Markov</td>
<td>pre-learned</td>
<td>0.3</td>
<td>266661</td>
<td>112660</td>
<td>115</td>
</tr>
</tbody>
</table>

Learning rate (α) set to 0.3, and Forgetting factor (γ) set to 0.8 for all tests.
# Flight Testing: Performance Evaluation of CAJ Communications

<table>
<thead>
<tr>
<th>Test #</th>
<th>Jammer type</th>
<th>Policy Type</th>
<th>Exploration Rate ($\epsilon$)</th>
<th>Average time in a Channel Without Being Jammed</th>
<th>Average Fraction of time in a Channel Without Being Jammed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sweep</td>
<td>Random</td>
<td>1.0</td>
<td>1399</td>
<td>0.14</td>
</tr>
<tr>
<td>2</td>
<td>Sweep</td>
<td>Random</td>
<td>1.0</td>
<td>1251</td>
<td>0.44</td>
</tr>
<tr>
<td>1 &amp; 2</td>
<td>Sweep</td>
<td>Random</td>
<td>1.0</td>
<td>1325</td>
<td>0.20</td>
</tr>
<tr>
<td>3</td>
<td>Sweep</td>
<td>Pre-learned, continuously updated through exploration</td>
<td>0.3</td>
<td>1634</td>
<td>0.56</td>
</tr>
<tr>
<td>5</td>
<td>Markov</td>
<td>Random</td>
<td>1.0</td>
<td>767</td>
<td>0.27</td>
</tr>
<tr>
<td>6</td>
<td>Markov</td>
<td>Pre-learned and fixed.</td>
<td>0.0</td>
<td>922</td>
<td>0.32</td>
</tr>
<tr>
<td>7</td>
<td>Markov</td>
<td>Pre-learned, continuously updated through exploration</td>
<td>0.3</td>
<td>980</td>
<td>0.42</td>
</tr>
</tbody>
</table>
Flight Testing: Policy vs Random Performance

**Sweeping Jammer**
- Random: 0.2
- Pre-learned w/exploration: 0.56

**Markov Jammer**
- Random: 0.27
- Pre-learned no exploration: 0.32
- Pre-learned w/exploration: 0.42
Conclusions

• Results show that the developed WACR approach is an effective anti-jamming tool, regardless of learning type and channel selection algorithms are used.

• Reinforcement learning aided cognitive anti-jamming communications policy significantly outperforms the random channel selection policy, both in terms of the average unjammed time in a channel as well as the fraction of time in a channel without being jammed.

• Performance is consistent regardless of the type of the jammer: Sweep or Markov.

• Allowing learning-based policy update and policy exploration during actual RF environment will lead to better performance with cognitive anti-jamming communications.

• Best possible performance improvements with the CAJ communications policy can expected to be higher than what is observed in these tests since these tests only allowed a very short learning period length, and parameters of the algorithms (i.e. learning rate, and forgetting factor, etc.) were unoptimized arbitrary values.
Contact Info:

jayaweera@bluecomsystems.com
505.615.1807

dale.mortensen@nasa.gov
216.433.6823