Using Classification to Protect the Integrity of Spectrum Measurements in White Space Networks

Omid Fatemieh, Ali Farhadi, Ranveer Chandra*, Carl A. Gunter

University of Illinois at Urbana-Champaign
*Microsoft Research

Network and Distributed System Security Symposium (NDSS)
Feb 17, 2011
Opportunistic Spectrum Access

• Spectrum crunch
  – Increased demand
  – Limited supply
  – Inefficiencies of fixed and long term spectrum assignment (*licenses*)

• Emerging solution: opportunistic access to *unused portions of licensed bands*
Opportunistic Spectrum Access

- Spectrum crunch
  - Increased demand
  - Limited supply
  - Inefficiencies of fixed and long term spectrum assignment (*licenses*)
- Emerging solution: opportunistic access to *WHITE SPACES*

- Cognitive Radio: A radio that interacts with the environment and changes its transmitter parameters accordingly
White Space Networks

• Allowed by FCC in Nov 2008 (and Sep 2010)
  – TV White Spaces: unused TV channels 2-51 (54 MHz-698MHz)
  – Much spectrum freed up in transition to Digital Television (DTV) in 2009
  – Excellent penetration and range properties

• Applications
  – Super Wi-Fi
  – Campus-wide Internet (e.g. Microsoft)
  – Rural broadband (e.g. Claudville, VA)
  – Advanced Meter Infrastructure (AMI) [FatemiehCG – ISRCS ‘10]
How to Identify Unused Spectrum?

- **Spectrum Sensing – Energy Detection**
  - Requires sensing-capable devices → cognitive radios
  - Signal is variable due to terrain, shadowing and fading
  - Sensing is challenging at low thresholds

- **Central aggregation of spectrum measurement data**
  - Base station (e.g. IEEE 802.22)
  - Spectrum availability database (required by the FCC)
Problem: Detecting Malicious Misreporting Attacks

- Malicious misreporting attacks
  - **Exploitation**: falsely declare a frequency occupied
  - **Vandalism**: falsely declare a frequency free

- Why challenging to detect?
  - *Spatial variations* of primary signal due to signal attenuation
  - *Natural differences* due to shadow-fading, *etc.*
  - *Temporal variations* of primary
  - Compromised nodes may *collude* and employ smart strategies to hide under legitimate variations
Setting and Attacker Model

- Network of cognitive radios (nodes) in large area
- Node $i$ periodically reports measurement $p_i$ to aggregation center to build a spectrum availability map
- End-to-end secure channel between nodes and aggregation center
- Geo-location for nodes
- Problem: How to protect against malicious attackers that may perform exploitation or vandalism
  1. Uncoordinated
  2. Coordinated
  3. Omniscient

\[ p_i \text{ higher than threshold} \]
\[ p_i \text{ lower than threshold} \]
Limitations of Existing Work

• [ChenPB – INFOCOM ‘08] [KaligineediKB – ICC ‘08] [MinSH – ICNP ‘09]
  – Consider detection in a small area with a common ground truth
  – Attackers constitute a small fraction of nodes (e.g. up to 1/3 [MinSH 09])
  – Not designed to detect areas dominated by attackers
  – Attackers use unsophisticated misreporting strategies

• [FatemiehCG – DySPAN ‘10]
  – Arbitrary assumptions about models and parameters of signal propagation
  – Rely on outlier detection threshold parameters that
    • Depend on propagation models and parameters
    or
    • Must be manually tuned
Solution Idea and Overview

- let data speak for itself
- Use natural signal propagation patterns to train a (machine learning) classifier
- Subsequently use classifier to detect unnatural propagation patterns -> attacker-dominated cells
• Widely used in spam detection, fraud detection, etc.

• Identifying patients with high risk of heart attack
  – Represent each patient as an example = \langle \text{label}, \text{features} \rangle
  
  – Goal: predict label for examples with known features (test examples) using examples with known features and labels (training examples)
  
  – Approach: building a classifier using training examples

• How to build classifiers? Winnow, Decision Trees, Naïve Bayes, Support Vector Machines (SVM), etc.

• Important factors: data representation, feature selection, choice of classifier
The local neighborhood of any cell $A$: $N_A$

Neighborhood (feature) representation of $A$

How to get training examples?
- Negative (normal): A one-time process using war-driving or a trusted set of sensors
- Positive (attacker-dominated): Randomized approach to inject uncoordinated, coordinated, and omniscient attacks

To build a unified classifier for each region, we use SVM with quadratic kernels

$$
\min \frac{1}{2} \| \hat{W} \|^2 + \gamma \sum_{i=1}^{N} \xi_i \\
\text{subject to } y_i (\hat{W} \cdot \Phi(\vec{x}) + \hat{w}_0) \geq 1 - \xi_i \quad \forall i
$$
Evaluation

Flat East-Central Illinois

Hilly Southwest Pennsylvania (Stress Test)

- TV transmitter data from FCC
- Terrain data from NASA
- House density data from US Census Bureau
- Ground truth: predicted signal propagation using empirical Longley-Rice model
Pennsylvania (Stress Test) Results

- 20km by 20km area
- Data from 37 transmitters within 150km
- Train classifier using data from 29
- Test classifier on data from 8
- Represent unaccounted uncertainties by Gaussian variations with mean 0 and std dev (σ) up to 6 (dB-spread) only to test data
- Worst-case results (σ=6)
  - Attacker detection rate
    - Uncoordinated: 97%
    - Coordinated: 95%
    - Omniscient: 94%
  - False positive rate: 7%
Conclusions and Future Work

• Motivated and formulated exploitation and vandalism attacks
• Showed how to build a classification-based defense using location-tagged signal propagation data
• Showed the effectiveness of approach against uncoordinated, coordinated, and omniscient attacks

• Future work
  – Additional features used for classification, *e.g.* elevation, *building density/height*
  – Building a crowdsourced nationwide spectrum availability map using *participatory sensing* data
  – Use a small subset of attestation-capable nodes as trust foundation [submitted to SECON ‘11]
Thanks
Illinois Results

- Train a **unified classifier** with WEIU-TV (PBS) and KTVI (Fox)
- Test on the following four

<table>
<thead>
<tr>
<th></th>
<th>WAOE</th>
<th>WCIA</th>
<th>WICS</th>
<th>WQAD-TV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>D.A. (%)</td>
<td>F.P. (%)</td>
<td>D.A.</td>
<td>F.P.</td>
</tr>
<tr>
<td>$P &gt; -65$</td>
<td>100</td>
<td>0</td>
<td>99.8</td>
<td>0</td>
</tr>
<tr>
<td>$-65 \geq P &gt; -85$</td>
<td>100</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>$-85 \geq P &gt; -105$</td>
<td>100</td>
<td>0</td>
<td>99.9</td>
<td>0</td>
</tr>
<tr>
<td>$-105 \geq P &gt; -114$</td>
<td>99.1</td>
<td>.9</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$-114 \geq P$</td>
<td>97.3</td>
<td>3.2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td><strong>99.3</strong></td>
<td><strong>.8</strong></td>
<td><strong>99.9</strong></td>
<td><strong>0</strong></td>
</tr>
</tbody>
</table>
Pennsylvania (Stress Test) Results

- 20km by 20km area
- Data from 37 transmitters within 150km
- Train classifier using data from 29
- Test classifier on data from 8
- Represent unaccounted uncertainties by adding Gaussian variations with mean 0 and std. dev (\(\sigma\)) up to 6 (dB-spread) only to test data

<table>
<thead>
<tr>
<th>False Positive Rates</th>
<th>Standard Deviation of Added Variations in Test Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\sigma=0)</td>
</tr>
<tr>
<td>(P &gt; -65)</td>
<td>0</td>
</tr>
<tr>
<td>(-65 \geq P &gt; -85)</td>
<td>0</td>
</tr>
<tr>
<td>(-85 \geq P &gt; -105)</td>
<td>.5</td>
</tr>
<tr>
<td>(-105 \geq P &gt; -114)</td>
<td>6.8</td>
</tr>
<tr>
<td>(-114 \geq P)</td>
<td>9</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td><strong>2.9</strong></td>
</tr>
</tbody>
</table>
Related Work – White Space Networks

• **Limitations of existing work**
  – Consider detection in a **small region** with a common ground truth
  – Attackers constitute a **small fraction** of nodes (e.g. up to 1/3 [MinSH 09])
  – Not able to detect **regions dominated** by attackers
  – Attackers use **unsophisticated** misreporting strategies

• [ChenPB – INFOCOM ‘08]
  – Weighted likelihood ratio test using similarity to final outcome as reputation
  – Uses 0/1 results: low overhead but Ignores measurement details
  – Bases the decisions on accurate knowledge of $P_{FA}$ and $P_{MD}$

• [KaligineediKB – ICC ‘08]
  – Assign (low) trust factors based on (an arbitrary) outlier detection
  – Use trust factors as weights in the averaging

• [MinSH – ICNP ‘09]
  – Shadow-fading correlation filters to exclude abnormal reports
Related Work – Sensor Networks (1)

• Major differences with sensor networks
  – More capable nodes
  – Long communication ranges

• Differences enable:
  – Centralized solutions with global view
  – Attestation, primary emulation, etc.
Related Work – Sensor Networks (2)

- Resilient data aggregation
  - [Wagner 04] Statistical analysis techniques for various aggregators
    - (+) Could be used to analyze our grid-based scheme
    - (-) Limited to small regions
  - [HurLHY 05] A trust-based framework in a grid: each sensor builds trust values for neighbors and reports them to the local aggregator
    - (sim) Similar to our grid-based scheme
    - (diff) No global view for a centralized aggregator
    - (-) Cannot identify compromised regions
    - (-) Does not consider statistical propagation / uncertainties
  - [ZhangDL 06] Identifies readings not statistically consistent with the distribution of readings in a cluster
    - (-) Local: only works for a small region
    - (+) Considers statistical distribution for readings
    - (-) Assumes data comes from distribution in the time domain
Related Work – Sensor Networks (3)

• Reputation/trust frameworks
  – [GaneriwalBS 04 & 08] A general reputation-based trust framework, where each sensor maintains a local reputation and trust for its neighbors
    • (diff) Local and P2P: reputation based on the quality of each interaction/report
    • (diff) Very general framework, focused on local decision making at each sensor

• Insider attacker detection
  – [LiuCC 07] Each node builds a distribution of the observed measurements around it and flags deviating neighbors as insider attackers
    • (diff) Local and P2P: voting among neighboring sensors to detect insiders
    • (-) Does not work in areas with more than 25% attackers

• Event region detection
  – [Krishnamachari 04] Fault tolerant event region detection
    • (diff) Only considers faulty nodes (not malicious); uniformly spread
    • (-) Node itself participates in detection
A Small Subset of Trusted Nodes

• Previous solutions
  – Used reported sensor measurements for inferring (dis)trust

• Remote attestation: A technique to provide certified information about software, firmware, or configuration to a remote party
  – Detect compromise
  – Establish trust

• Root of trust for remote attestation
  – Trusted hardware: TPM on PCs or MTM on mobile devices
  – Software on chip [LeMay, Gunter - ESORICS ‘09]

• Why a subset?