A Robust Application Detector For Intelligent Wireless Collaboration

Kevin Pietsch and Sean Mason

Lockheed Martin Advanced Technology Laboratories (ATL)

Cherry Hill, New Jersey 08002

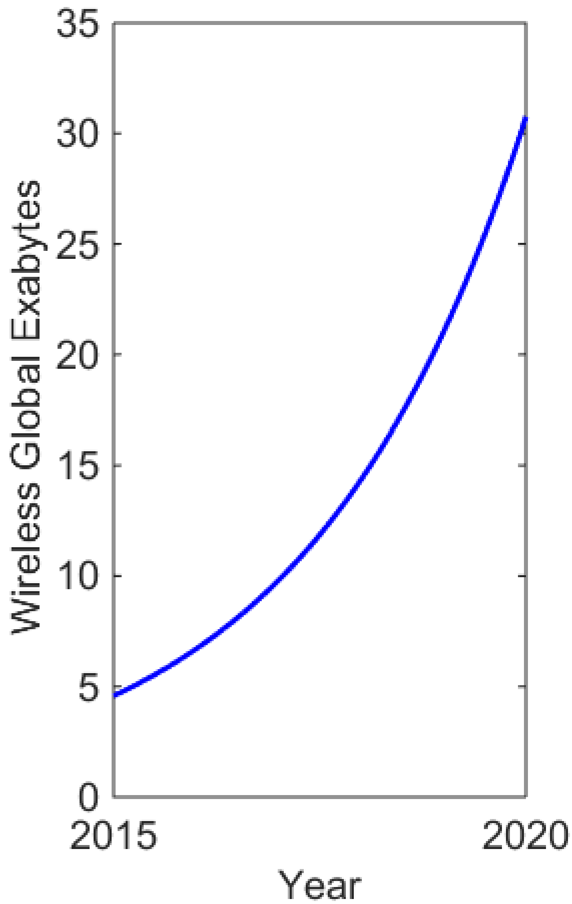
Presented by: Kevin Rigney of Lockheed Martin ATL

December 9, 2016

DISTRIBUTION A. Approved for public release: distribution unlimited. This research was developed partly under funding from the Defense Advanced Research Projects Agency's (DARPA) MTO Office under contract HR0011-11-C-0033. The views, opinions, and/or findings expressed are those of the author(s) and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

Spectrum Inference

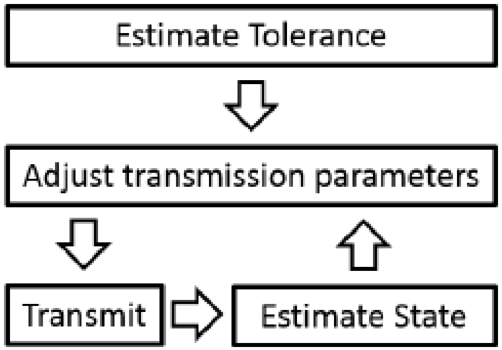
- **What is it?**
  - **Generally: the ability to derive information over a wireless link inexplicitly (i.e. from evidence and reason)**
  - **Practically this means without demodulation**
- **Why is it important?**
  - **Spectrum usage continues to grow exponentially**
  - **Current spectrum sharing paradigm of sense-and-avoid leaves capacity on the table**
  - **Just because a primary user radio (PU) occupies a channel doesn't mean the channel should be off limits to secondary users (SU)**
    - **SU's must be able to share without hurting the PU**



In February 2016 The Cisco Visual Networking Index Forecast predicted a growth in global data usage from 4.4 exabytes/month in 2015 to over 30 in 2020

Research Considerations

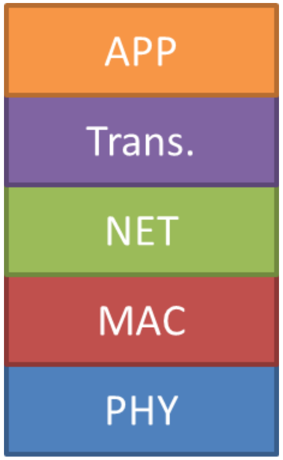
- **Where does it lead?**
  - **Radios that formulate strategies for coexistence without direct communication**
  - **Real-time feedback for spectrum sharing actions**
- **Practical considerations**
  - **How do you train?**
    - **What data is available for training?**
    - **What's the right amount of labeling?**
  - **How to build a radio around this inference ability?**



An outline of the collaborative spectrum sharing process

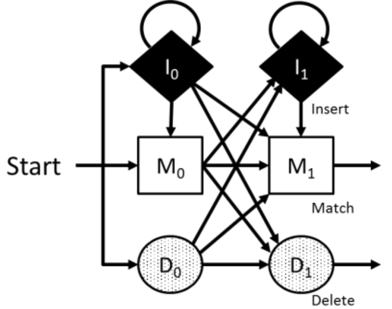
Our Problem Statement

- **Determine the protocol application in use by a PU over the wireless channel**
  - **SU cannot demodulate and only has a simple detector**
- **Why bother? This allows spectrum sharing rules based on <u>application</u>**
  - **Most cognitive radio research assumes rules based on radio type**

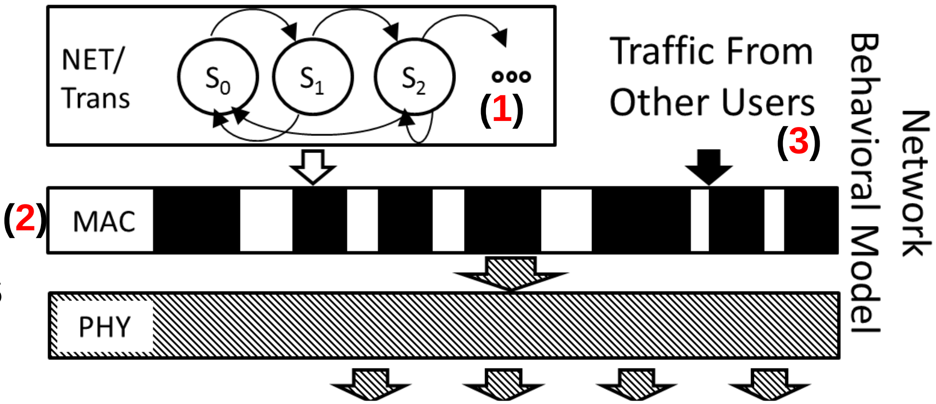


The open system interconnect model

Profile Hidden Markov Model (PHMM)



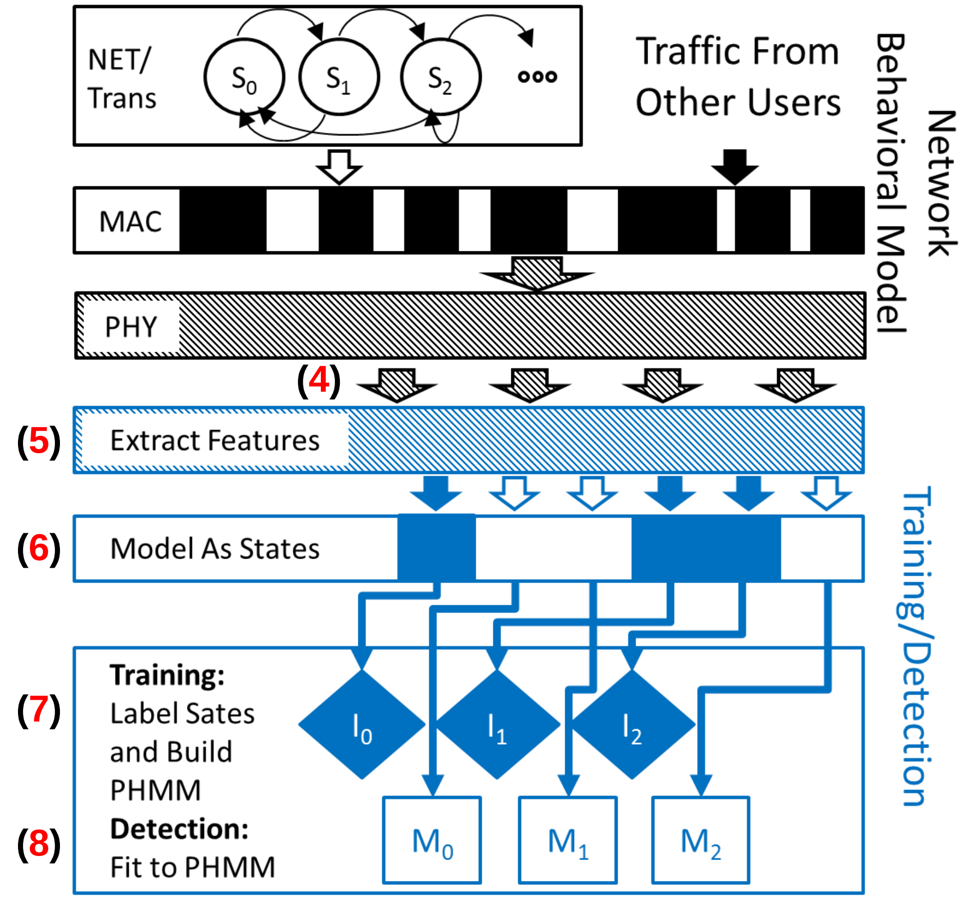
- **Extension of hidden Markov model (HMM), developed to find mutated genetic sequences**
  - **Genetic mutations (insertions and deletions) have close analogies to wireless traffic**
- **Models an application protocol as a sequence of hidden states (1)**
  - **Emissions (2) are protocol states after being encoded, modulated, and transmitted by the PU**
  - **Inserted (3) and deleted emissions are a consequence of shared, lossy medium**



PHMM Training & Detection

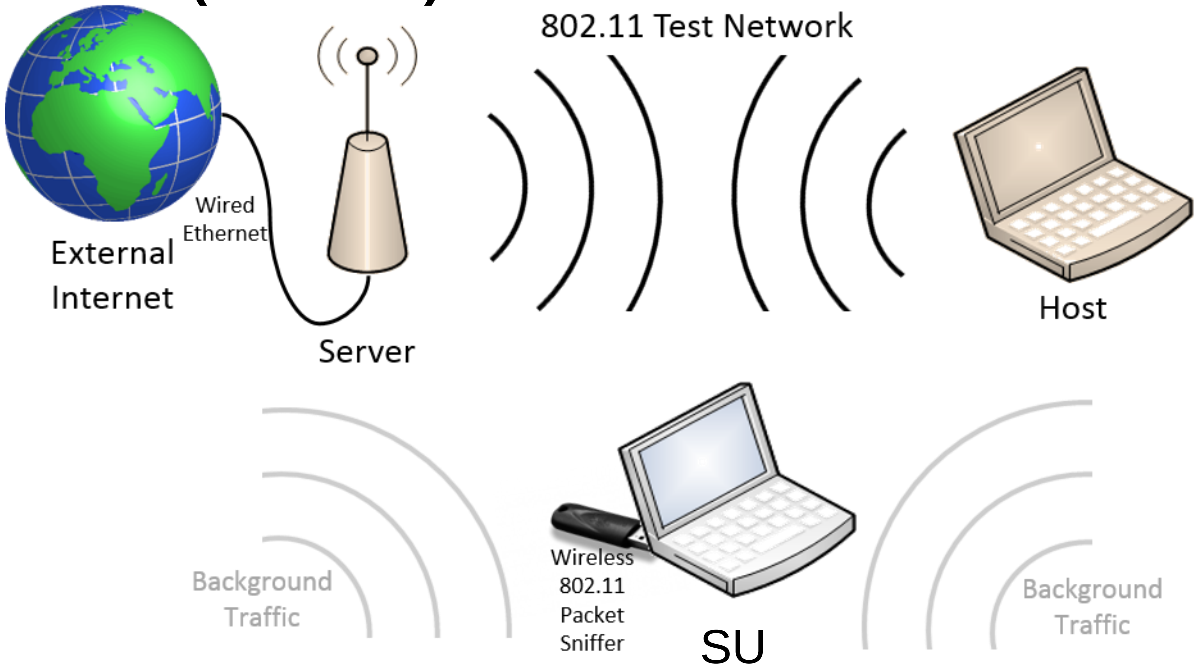
Common to Training and Detection

- **SU receives PU’s emissions over wireless channel (4)**
  - **Possibly with deletions and insertions from other users**
- **Feature Extraction (5) reduces dimensionality of observed traffic**
  - **E.g. convert raw voltage measurement to burst durations and interarrival times**
- **State modeling (6) defines discrete states in feature space (e.g. clustering)**
- **Training (7) creates a state transition model per application type**
- **Detection (8) fits observed state sequence to a model**



Experimental Setup

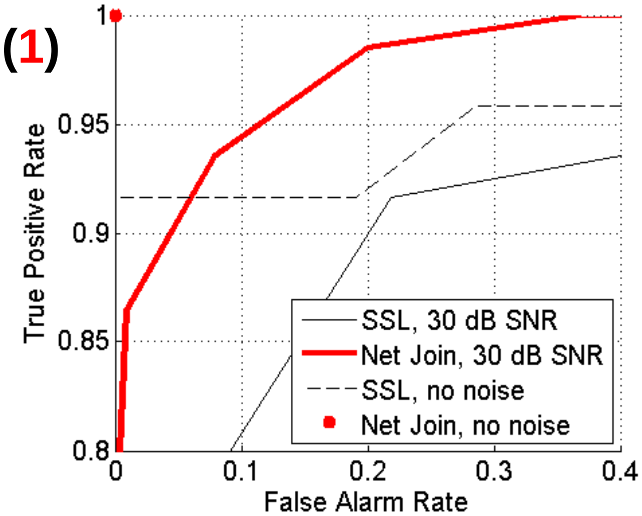
- **PU: IEEE 802.11g (WiFi)**
- **Test protocols**
  - **Network association using dynamic host control protocol (DHCP)**
  - **Secure socket layer (SSL) handshake**
  - **Background**
- **Feature extraction using Wireshark: packet size (bytes) and interarrival time (seconds)**



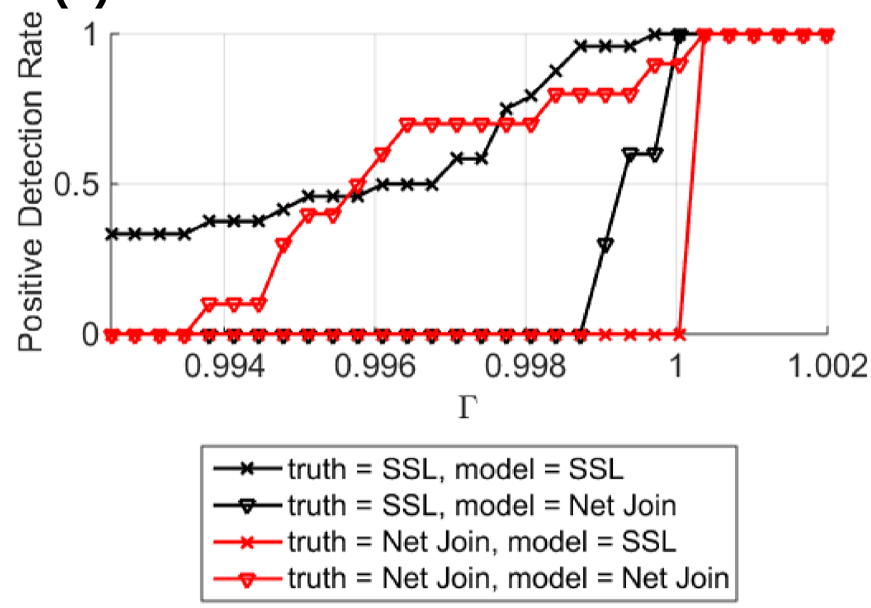
Experimental Results

- **Results include**
  - **Discrimination of protocol from background (1)**
  - **Discrimination of protocols from each other (2)**
- **Number of training samples per model**
  - **DHCP (Net Join): 9**
  - **SSL: 23**
  - **Background (single state HMM): 20**

(1)



(2)



Future/Ongoing Work

- **New detectors**
  - **Esp. ones that require less training**
- **More primary user radio types**
- **More challenging, realistic experimental setups**

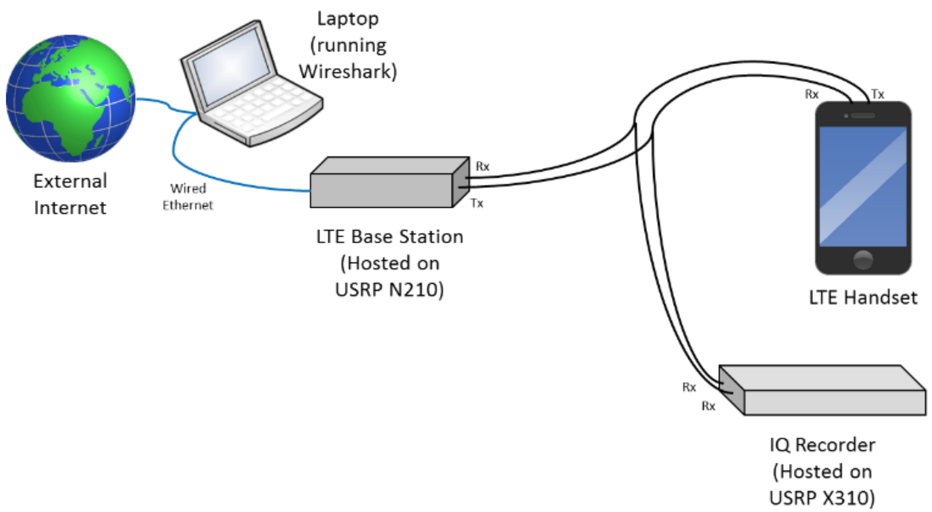
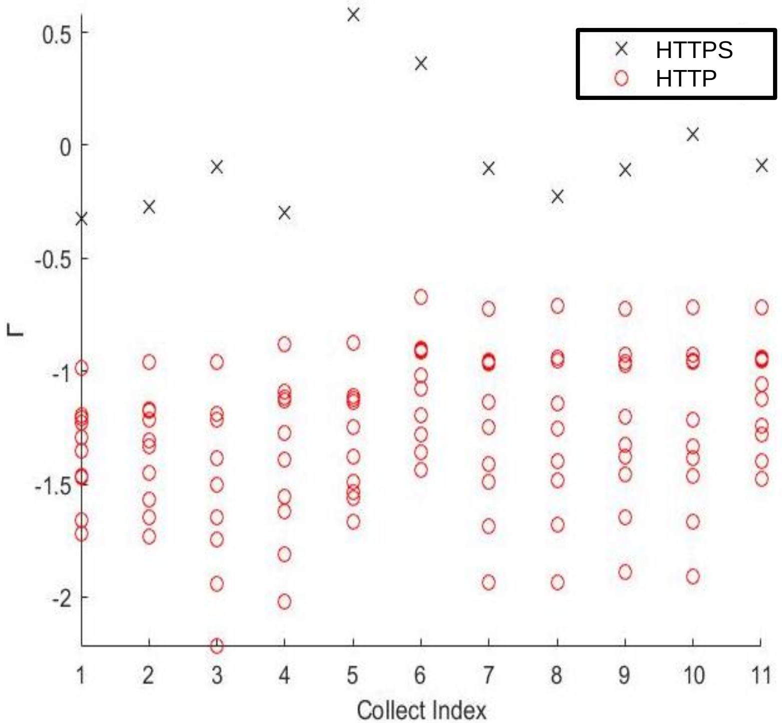


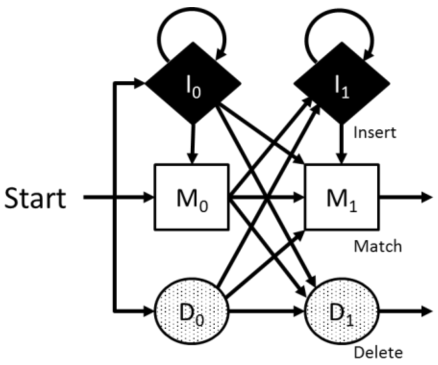
Diagram of an experimental setup with a Long Term Evolution (LTE) primary user



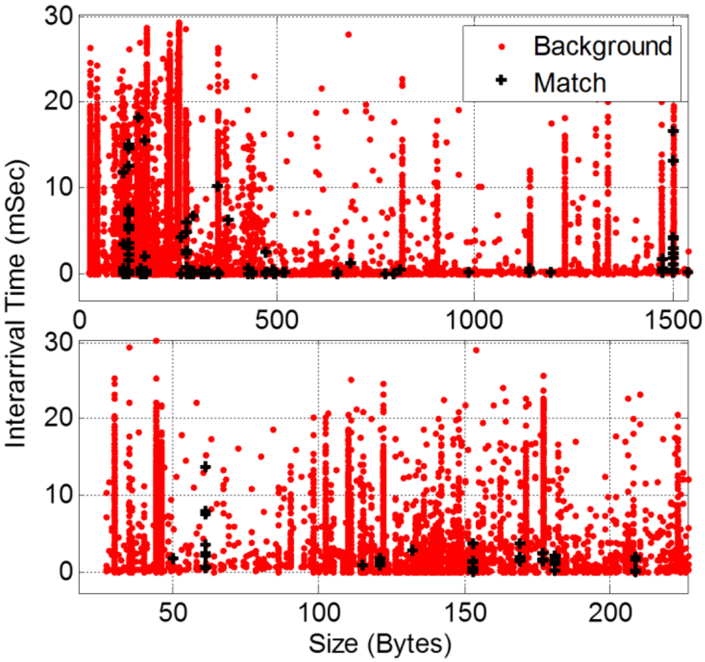
Detector output from LTE experiment. A PHMM trained to detect SSL consistently produces a higher detection score on HTTPS collects (which contain SSL) than plain HTTP (which does not)

Thank You

Profile Hidden Markov Model (PHMM) Detector



- **Extension of hidden Markov model (HMM), developed to find mutated genetic sequences**
  - **Genetic mutations (insertions and deletions) have close analogies to wireless traffic**
- **The HMM models an application protocol as a sequence of hidden states**
  - **Emissions are protocol states after being encoded, modulated, and transmitted by the PU**
  - **Inserted and deleted emissions are a consequence of shared, lossy medium**



These are the features used to classify radio type. A non temporal classifier would need many degrees of freedom (and therefore many training examples) to separate background and match