Statistical Modeling Framework of Live Spectrum Observation data for Opportunistic Spectrum Sharing

Tanim Taher¹, Kenneth Zdunek²
Electrical and Computer Engineering department
Illinois Institute of Technology, Chicago, IL, USA.
Email: tahetan@hawk.iit.edu¹, zdunek@iit.edu²

Dennis Roberson³
Computer Science department
Illinois Institute of Technology, Chicago, IL, USA.
Email: robersond@iit.edu³

Abstract—A statistical model of land mobile radio (LMR) voice traffic is developed from empirical RF Spectrum measurement data. This model builds upon previous work, and is used to generate synthetic voice traffic that closely follows the daily and weekly patterns of the measured traffic. The model is applied to a Dynamic Spectrum Access (DSA) simulation. A coexistence algorithm that takes advantage of the modeled channel statistics is presented that allows an opportunistic secondary user (SU) to share a channel with a primary user (PU). The algorithm shows a clear improvement compared to the basic Listen-before-talk scheme that has no knowledge of a PU’s statistical traffic characteristics. Spectrum Opportunity Accessed and collision rate are used as metrics to compare the DSA coexistence techniques. We demonstrate the utility of a spectrum observatory system as being the integral part of this DSA framework, where the observatory continually monitors and models PU channel activity in order to provide the SU with useful statistical information about the PU’s traffic characteristics.

Keywords-statistical model; Cognitive Radio; Dynamic Spectrum Access Networks (DSA); voice traffic; Land Mobile Radio (LMR); public safety radio; Listen-before-talk (LBT); Spectrum sharing; Spectrum Observatory.

I. INTRODUCTION

The Land Mobile Radio (LMR) band voice channels are widely used by commercial users. They also play a critical role for public safety communications where public safety agencies like police departments use LMR systems for communication between dispatch centers and mobile field agents, and also for direct mobile-to-mobile communications. In the United States, LMR channels are allocated in the VHF (148-174 MHz) and UHF (450-512 MHz) bands [KOB01].

In previous work [BAC10], it has been shown that the allocated LMR spectrum is sparsely utilized by some commercial user groups, while users like public safety commonly reach near full capacity – especially during emergencies [TAH11]. Hence, the LMR bands are candidates for studying the feasibility of Dynamic Spectrum Access (DSA) technology such that users in a high activity channel (like public safety) can gain extra capacity when needed by sharing spectrum with low activity channels [LEE10]. For example, DSA technology can be applied to increase the pool of public safety channels during an emergency situation when federal, state and city agencies may converge in a geographic locale [FCC08][FCC09].

In order to permit channel sharing, knowledge of the Primary User (PU) activity is essential towards the goal of limiting the likelihood of collisions between the PU and Secondary User (SU) [HUA09]. Additionally, the SU is able to use information about the PU to assess the level of spectrum opportunity available in that channel so it can decide whether to use it or to search for a more favorable channel. Knowledge of the PU has to be obtained from RF spectrum measurements – to this end, spectrum measurements conducted at the Illinois Institute of Technology (IIT) are used in this study.

Based on empirically measured LMR data, in [TAH12] a static long-term model and another time-varying statistical model for the voice traffic activity in an LMR channel were presented. The models in [TAH12] are applicable in developing algorithms for PU and SU coexistence. In this paper, an improved and more elaborate version of the time-varying statistical model of LMR channel traffic is presented.

Statistical models for LMR traffic are also applicable in network studies for simulating traffic [ASC09] – which, for example, could help in the understanding the dynamics of a public safety network. The model presented here is used to generate synthetic public safety traffic data which is then compared with the empirically measured traffic.

In [MCH07][MCH08], a DSA algorithm was designed and field-tested that used the Listen-before-talk (LBT) algorithm. This achieved good results, but did not use any statistical knowledge of the PU channels. In [HUA08], the authors describe a system where the SU is able to use statistical knowledge of the PU’s idle times in order to achieve optimal SU throughput subject to a collision limit constraint; the advantage of knowing the PU’s traffic characteristics was thus demonstrated. In the second part of this paper, a DSA framework is presented where a spectrum monitoring and analysis system continually measures LMR PU channels and then calculates the statistical parameters corresponding to the model presented in the first part of the paper. The model parameters are provided to an SU which is able to opportunistically access the temporal vacancies of the PU with minimal collisions. This DSA framework is compared to the basic LBT algorithm.

This paper is organized as follows. Section II gives an overview of the measurement system. Section III describes the new model for LMR voice traffic; and Section IV applies this model to generate synthetic traffic data which is then compared
with empirical measurements. Section V describes the spectrum sharing framework between a primary LMR user and a secondary user, and compares the performance of this coexistence system with the basic LBT. Applications, conclusions and future work are presented in Section VI.

II. RF MEASUREMENT SYSTEM

A RF data acquisition and storage system was implemented to measure voice channels in the 450 MHz LMR band, as shown in Figure 1. The system consists of an omni-directional discone antenna set up on the roof of the 22 storey IIT tower, located approximately 3 miles south of downtown Chicago. The antenna is connected to a 450-474 MHz bandpass filter, and then to a USRP N200 software radio platform [ETT13] with a wideband frontend. During each measurement sweep, the USRP captures time-domain samples which are then used to estimate the Power Spectrum via Fast Fourier Transform (FFT). From the power spectrum, the power within every 12.5 kHz LMR channel span is found by integration. These power values are then stored in a database and also web-streamed live in real-time. A more detailed description of the system is found in [TAH12].

![System Block diagram](Image)

Figure 1. System Block diagram

For the purposes of this paper, this USRP based RF measurement system will be referred to as a “spectrum observatory” as it continually monitors the spectrum and channel usage in the 450 MHz LMR band on which this research is based. Although not the focus of this paper, parallel systems for measuring public safety narrowband traffic in the 700 MHz Block D spectrum, and the 851-854 MHz NPSPAC frequencies have also been implemented at IIT.

The 450 MHz LMR band spectrum observatory observes traffic activity in each of the 1,000 LMR channels within the 452.75-465.25 MHz frequency range with a time resolution of 250 ms. This sampling time of 250 ms captures most, if not all on-off voice traffic activity – this is because voice calls have a relatively slow switching time. Hence, the measurements obtained here have adequate time resolution necessary for modeling and characterizing the traffic in each channel.

For this paper, we analyzed data collected over a five week period starting from September 15th, 2011 in the 453-465 MHz frequency range. The channels studied included 21 Chicago Police Department (CPD) public safety city-wide dispatch channels and one business related voice channel used by IIT. These channels and the date range correspond to the same data used for modeling work in [TAH12]; this permits direct comparison with the newer channel model presented here.

III. CHANNEL MODELING

In [TAH12], statistical models for the “idle time” durations, “hold times”, and “inter-arrival times” were provided. Idle time is defined as the time period between transmissions, when there is no measurable RF power in the LMR channel. Hold time is the duration of transmissions, that is, the time when the channel has power that exceeds a threshold. Inter-arrival time [BAR97] is the time duration between the start of one transmission and the next, and is equivalent to the sum of two, consecutive hold and idle times.

In this paper, the term “conversation duration” is introduced which is the combined time duration of a group of calls that are spaced with very short gaps – so the series of calls are likely to be part of one relatively long conversation. A “idle period threshold time” of 10 s is chosen, which is used to group a series of calls into conversations – that is, if the idle period is lesser than 10 s between two consecutive calls (On-times), then the calls are part of the same conversation. Ten seconds was chosen since in voice communications, a pause longer than that generally indicates that the previous conversation has ended. However, for modeling purposes, it is possible to choose a different value for “idle period threshold time”.

Figure 2 illustrates the states in the new model. “Quiet state” refers to idle times longer than “idle period threshold time”. Within the “busy state”, several calls, with On-periods described as “hold times” and Off-periods between them labeled as “short gaps”, occur. “Short gaps” correspond to idle times less than 10 s. The model distribution for “busy state” is obtained from the histogram of “conversation durations”; a histogram of “quiet periods” (idle times > 10 s) corresponds to the distribution of the “quiet state”; and one histogram for “hold times” and another for “short gaps” describe the behavior of the measured traffic within the “busy state” conversations.

![State Diagram for LMR channel model](Image)

Figure 2. State Diagram for LMR channel model

Hence, the channel can be modeled with four histograms to be mapped and curve-fitted with common probabilistic distributions. The histograms being: (a) histogram of quiet periods; (b) histogram of conversation durations; (c) histogram of hold times; (d) and histogram of short gaps. Figure 3 shows the four histograms for a CPD LMR channel as measured empirically over a 24 hour time window. Notice that the quiet periods are all greater than 10 s, while the short gaps are all less than that due to the state assignments defined by Figure 2.

Each of the four call duration quantities was compared with exponential, generalized Pareto, lognormal, and Gamma distributions and maximum likelihood estimation was used to
calculate the parameters of the probability density function (pdf) each time. For the histogram of “quiet periods”, the time axis is first offset back to the origin before estimation is performed. For each of the four sets of quantities, the Kolmogorov-Smirnov (K-S) goodness of test [MAS51] with a 5% significance level was performed to compare the fit of measured data with the estimated cdf for each type of distribution (i.e., lognormal, Gamma, etc). The p-value of the K-S test statistic is an indicator of how well the distribution fits – the higher the p-value, better the fit. Table 1 shows the average p-values for the four call duration quantities when fitted with the four common probability distributions, across the 22 LMR channels.

From Table 1, the lognormal distribution is seen to best match with the histograms for both the “hold times” and “short gaps”. The Gamma distribution gives the highest p-values for the “quiet periods” and “conversation lengths”; meaning these two quantities can be closely modeled by a Gamma distribution. For simplicity, however, in this paper, for the 22 LMR channels measured, a single distribution function, that is the lognormal, is used to model the statistics for each of the four sets of random quantities.

In [TAH12], it was shown that the traffic in a channel is non-stationary. It was shown that the probability of the channel being stationary within an interval declines in an approximate linear fashion as the length of the time window increases. Thus, the parameters of the pdf distributions of the model presented above should be frequently updated to properly track the changes in traffic patterns within the LMR channel.

From analysis of the non-stationarity aspects of the data in [TAH12], it was found that the channel can be considered mostly stationary for window lengths of between 30 to 120 minutes. Hence, to track a channel over a long period of time, the 8 parameters corresponding to the 4 lognormal distributions of the model need to be updated at a frequency of anywhere between 30 minutes to 120 minutes. Figure 4a shows the channel activity over 3 days beginning September 15th, 2011 for the same CPD channel as in Figure 3. A smoothing filter was used for plotting purposes only, to reveal daily trends and dynamics. The model developed in this section was then applied to the channel activity data from the spectrum observatory, and every 120 minutes, the 8 parameters were re-estimated and 8 time series of parameters were obtained. From these, four selected time series of parameters are plotted in Figure 4b, to illustrate that the model is able to track a channel for long periods of time, even as the channels statistics keep changing over time. The fifth graph (green) in 4b is the plot of the ratio between the number of “quiet periods” observed and the number of “short gaps” within each 120 minute time window. This ratio, although not required for generating synthetic data (Section IV) and hence not part of the model, gives useful information that allows a SU in a DSA scenario to track how many usable “quiet periods” it is likely to find while attempting to share the channel with the PU.

### Table 1. Average p-value from K-S test to examine goodness of fit

<table>
<thead>
<tr>
<th></th>
<th>Log-normal</th>
<th>Gamma</th>
<th>Pareto</th>
<th>Exponential</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conversation</td>
<td>0.581</td>
<td>0.667</td>
<td>0.641</td>
<td>0.491</td>
</tr>
<tr>
<td>Quiet-periods</td>
<td>0.671</td>
<td>0.723</td>
<td>0.723</td>
<td>0.465</td>
</tr>
<tr>
<td>Hold Times</td>
<td>0.276</td>
<td>0.138</td>
<td>0.091</td>
<td>0.069</td>
</tr>
<tr>
<td>Short gaps</td>
<td>0.219</td>
<td>0.184</td>
<td>0.168</td>
<td>0.1332</td>
</tr>
</tbody>
</table>

Figure 3. Histograms of state durations (measured data for CPD channel)

Figure 4. (a) Occupancy vs time in the CPD channel over 3 days. (b) Variation of model parameters in 120 minute intervals over the 3 days

### IV. SYNTHETIC TRAFFIC GENERATED FROM MODEL

In this section, the model developed in Section III is used to generate synthetic traffic data for a CPD channel. The traffic generated is then compared with the empirical data in order to validate the model. The 8 time-series of parameters estimated from the LMR spectrum observatory measurements for the CPD channel over a 3 day period are applied to a simulation algorithm to generate 3 days worth of synthetic voice traffic data that mimics the activity in the CPD channel. The algorithm runs as follows:

1) Update the 8 statistical parameters for the model’s four lognormal distributions in the current time window with data from LMR spectrum observatory.
2) Generate a random number from the pdf of “conversation durations”. This will be the duration of one conversation.

3) Generate calls within this conversation:
   a. Generate a random number for call hold duration from the pdf of “hold times”.
   b. Generate a random number for the short idle time before the next call from the pdf of “short gaps”.
   c. If the series of On-Off calls fill up the conversation duration from step 2, advance to step 4; otherwise iterate to 3a.

4) Generate a random number to be used as the long quiet state after the conversation ends from the pdf of “quiet periods”.

5) Simulation time-limit reached? If No, advance to step 6. If Yes, then End simulation.

6) Check if the simulation time has reached the end of the current time window. If Yes, advance to the next time window and iterate to step 1; if No, iterate to step 2.

The same scheme is easily extendable to any of the 22 LMR channels and also for any time period spanning weeks and even months. However, for illustration purposes, only 3 days of synthetic traffic data is plotted in Figure 5. Figure 5 plots the empirically measured traffic activity in the CPD channel alongside the computer-generated traffic where the model parameters were updated (algorithm step 1) every 30 minutes. Both the plots are smoothed with a 20 minute moving average filter to reveal daily features and traffic dynamics. The synthetic traffic is seen to follow the empirical traffic very closely in Figure 5, thus validating the modeling procedure developed here.

Further model validation is demonstrated by the plots in Figure 6. In this case, 5 days’ worth of synthetic data was generated using a time window length of 45 minutes for the same CPD channel. Figure 6a counts the total number of calls within each 45 minute time window and compares the values obtained for synthetic and empirical traffic data. Close match is again observed, despite the random nature of the statistical model used to generate the synthetic traffic. The count of calls within a time window is a good metric for model validation as it does not follow directly from the statistical state distributions, but rather is a consequence of the model’s dynamics. Similarly, 6b uses another high-level metric – that is the count of separate conversations within each time window. Again, good correspondence in the number of conversations counted is seen between the synthetic and empirical traffic datasets. Thus, the model has been demonstrated to closely track the LMR channel for arbitrarily long periods of time, given that good statistical estimates for the model parameters are available from a spectrum observatory.

Figure 6. Model validation by comparing call statistics of Computer-generated traffic with that of traffic measured by the spectrum observatory in a CPD channel. (6a) Total number of calls in each 45 minute time window. (6b) Total number of separate conversations in each 45 minute time window.

V. FRAMEWORK FOR PU AND SU COEXISTENCE & SIMULATION RESULTS

A simulation environment for spectrum sharing was created consisting of PU traffic (either empirical or synthetic) in a channel. A LMR spectrum observatory constantly monitors the PU traffic and obtains model parameters to describe the channel-traffic’s statistics once every 30 minutes. This observatory is an integral part of the DSA framework. An SU attempts to make use of the Spectrum Opportunity available – that is, the fraction of time that the PU is vacant from the channel. Two test cases are studied – in case 1, the SU has no knowledge of the PU and receives no information from the spectrum observatory; in case 2, the observatory provides the SU with channel model parameters once every 30 minutes, and the SU assumes that these parameters describe the channel condition sufficiently well for the following 30 minutes.

In case 1, the SU uses a simple LBT algorithm to access the channel. The SU senses the channel first; if the PU is absent or if the channel is observed to transition from busy to vacant, then a countdown timer of random duration $T_c$ is started. If the PU does not return during this countdown interval, the SU transmits for a random duration $T_x$. During the transmission state, the SU cannot do channel sensing, and hence collisions can occur if the PU returns. $T_c$ is a uniform random variable between $[T_{slot} \times T_{max}]$, and $T_x$ is uniform between $[T_{slot} \times T_{max}]$, where $T_{slot}$ is the time-slot used in the simulation and $T_{max}$ and $T_{max}$ are, respectively, the maximum countdown timer and the
maximum transmit duration. The $T_{C_{\text{max}}}$, $T_{X_{\text{max}}}$ are constant during every time window throughout the simulation duration, and are not functions of the PU traffic’s model. For the simulation $T_{\text{dum}}=0.25$ s, $T_{C_{\text{max}}}=4$ s and $T_{X_{\text{max}}}=2$ s. The latter two parameters were chosen rather arbitrarily. The only design consideration was that $T_{C_{\text{max}}}>T_{X_{\text{max}}}$ to minimize the likelihood of interference that occurs when the SU collides with the PU; since the maximum sensing duration was selected to be larger than the maximum transmit duration.

The secondary opportunistic spectrum access user in simulation case 2 employs a smart LBT algorithm. Here the maximum sensor countdown timer, $T_{C_{\text{max}}}$, and the maximum transmission duration, $T_{X_{\text{max}}}$, are recalculated during every 30 minute interval using the PU traffic model parameters provided to the SU by the LMR spectrum observatory. Many different functions with varying levels of sophistication can be used to obtain the two SU parameters, $T_{C_{\text{max}}}$ and $T_{x_{\text{max}}}$. However, the goal of this short paper is to demonstrate the utility of modeling PU traffic and the advantages of incorporating, within the DSA framework, a spectrum observatory that measures empirical PU traffic and calculates the model parameters. Hence, for simplicity, non-complex equations are used to calculate $T_{C_{\text{max}}}$ and $T_{X_{\text{max}}}$ from the PU model parameters. Equations (1) and (2) are used to calculate $T_{C_{\text{max}}}$ and $T_{X_{\text{max}}}$. Equation (2) is the lognormal cdf function where $\mu$ and $\sigma$ correspond to the channel model parameters for the distribution of “short gaps” in the PU traffic. For both $T_{C_{\text{max}}}$ and $T_{X_{\text{max}}}$ $p=0.6$ was used as inputs to (1) and (2). As stated above, both these LBT control parameters are recalculated at the start of each new window.

$$T = F^{-1}(p|\mu, \sigma),$$

$$p = F(T|\mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \int_0^T e^{-\frac{(\ln(\tau) - \mu)^2}{2\sigma^2}} d\tau,$$  \hspace{1cm} (1)

Two metrics are used to test the performance of the coexistence techniques employed by the SU – Spectrum Opportunity Accessed (SOA) and Fraction of Collisions (FoC) affecting the PU. These metrics are obtained by (3) and (4), respectively. For both case 1 and case 2 DSA scenarios, SOA and FoC are calculated and plotted.

$$SOA = \frac{\text{SU Transmit Duration} - \text{Collision Duration}}{\text{total time}},$$ \hspace{1cm} (3)

$$\text{FoC} = \frac{\text{Collision Duration}}{\text{PU Transmit Duration}}.$$ \hspace{1cm} (4)

Figure 7 displays the results of the DSA coexistence simulations for cases 1 and 2. The smart SU (case 2) which utilizes the PU traffic’s model parameters achieves higher throughput than the simple LBT (case 1) that has no knowledge about the PU’s statistics. Figure 8 shows how the SOA and FoC values change over the two day period when the smart case 2 SU operates in the channel. Overall, the case 2 smart SU is able to achieve an SOA of 53.3% while the PU experiences a FoC of only 3.6%. In contrast, the simple LBT (case 1) SU algorithm attains a lower SOA of 43.6% but causes higher interference to the PU (FoC is 6.2%). Hence, for this particular channel, the smart SU easily outperforms the simple LBT coexistence technique. This shows the advantage of having a spectrum observatory within the DSA framework to measure and model the PU channel statistics for appropriate use by intelligent SUs.

The set of simulations was repeated for each of the 22 LMR channels, over a longer 14 day time-span. This was done with both real empirically measured traffic (Table 2) and synthetic computer-generated traffic (Table 3). Tables 2 and 3 present the average SOAs achieved by the simple LBT and the smart SU during the 22 channel simulation run. A higher SOA value means higher SU throughput. Tables 2 and 3 also show the average FoC values obtained across the 22 channels. A higher FoC value indicates higher interference experienced by the PU.

Both with real and synthetic data, it is seen every time that the smart SU with PU traffic model information outperforms the simple LBT algorithm and attains a higher SOA. The average FoC is approximately the same for both algorithms. So for similar interference levels, the smart SU achieves higher throughput and higher spectrum opportunity utilization for the 22 LMR channels. This validates the performance of the smart SU algorithm, and demonstrates the advantages of the SU possessing statistical knowledge of PU traffic.
Table 2. Performance of DSA coexistence techniques where real empirical data is used for PU Traffic; all 22 LMR channels tested in simulation

<table>
<thead>
<tr>
<th>Average across 22 LMR Channels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectrum Opportunity Accessed (simple LBT)</td>
</tr>
<tr>
<td>Spectrum Opportunity Accessed (smart SU)</td>
</tr>
<tr>
<td>Collisions experienced by PU (simple LBT)</td>
</tr>
<tr>
<td>Collisions experienced by PU (smart SU)</td>
</tr>
</tbody>
</table>

Table 3. Performance of DSA coexistence techniques where Computer-generated data is used for PU Traffic; all 22 LMR channels tested

<table>
<thead>
<tr>
<th>Average across 22 LMR Channels</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOA (simple LBT)</td>
</tr>
<tr>
<td>SOA (smart SU)</td>
</tr>
<tr>
<td>FoC (simple LBT)</td>
</tr>
<tr>
<td>FoC (smart SU)</td>
</tr>
</tbody>
</table>

VI. CONCLUSIONS AND FUTURE WORK

An improved model for the voice traffic in public safety channels in the LMR band was presented. This model consists of busy and quiet states; the busy state in turn is subdivided into two sub-states, namely holds and gaps. The histograms of the dwell time in each state were compared to common distributions, and lognormal distribution was selected to model each state. Time-series’ of parameters for the four distributions allow the model to track the traffic in any channel for long time periods. The model was validated by employing it to generate synthetic traffic whose characteristics and dynamics matched closely with empirical traffic measurements. For future work, the model needs to be refined by identifying distributions that correspond more closely to the histogram of “conversation durations”.

The model was then applied in a dynamic spectrum sharing scenario where a secondary user with knowledge of the PU traffic’s model parameters was able to coexist with minimal interference and at the same time make high utilization of the available spectrum opportunity. This smart SU was shown to attain a higher performance level than the simple LBT coexistence technique that employs no knowledge of the PU’s traffic characteristics. The key point in this work was to demonstrate an application for continuous spectrum monitoring systems. The spectrum monitor played an integral part of the spectrum sharing framework by measuring the PU’s traffic characteristics, modeling them and then reporting the parameters to the SU. Without the spectrum observatory, the smart SU would not have performed any differently than the simple LBT SU.

As part of ongoing and future work, we are investigating spectrum sharing algorithms that make better use of the PU’s channel statistics, and that aims to achieve even higher secondary user throughput with low interference. Also, a new method to classify multiple voice LMR channels based on the model parameter values is being investigated. This would allow an SU to maximize throughput by selecting PU channels that are predicted to provide the highest spectrum opportunities.

REFERENCES


