PREDICTIVE DYNAMIC SPECTRUM ACCESS

T. Charles Clancy (Laboratory for Telecommunications Sciences, US Department of Defense, 8080 Greenmead Drive, College Park, MD 20740, clancy@ltsnet.net); Brenton D. Walker (Laboratory for Telecommunications Sciences, brenton@ltsnet.net)

ABSTRACT

This paper explores the idea of predictive dynamic spectrum access (PDSA). Modern spectrum resource allocation research typically divides users into two classes: primary users and secondary users. Primary users own licenses to particular frequency bands and are allowed to use it whenever they wish. Secondary users can reuse the frequencies when they are not being used by a primary user.

The goal of PDSA is to gather statistical information about a primary user in an effort to predict when the channel will be idle. This allows us to better plan secondary use of the spectrum without the cooperation of the primary user.

We explore two approaches to PDSA in this paper. The first uses cyclostationary detection on the primary users' channel access pattern to determine expected channel idle times. These techniques are simulated with both TDMA and CSMA networks. The second briefly examines the use of Hidden Markov Models (HMMs) for use in PDSA.

1. INTRODUCTION

Many dynamic and opportunistic spectrum access techniques search for gaps in frequency and time where spectrum is not being utilized by a primary license holder [2,4]. Typically they sense the RF spectrum, and transmit if it is idle. However, between the sense time and the end of the packet transmission, the licensed device could begin transmitting, resulting in interference to both the licensed signal and the unlicensed cognitive transceiver.

This paper proposes a learning technique for cognitive radios that will allow prediction of spectral vacancy called predictive dynamic spectrum access (PDSA). In order for it to be effective, the licensed signal must have periodic properties. An ideal signal would be TDMA, since it has inherent periodicity in its PHY. While CSMA networks are not natively periodic, many higher-layer protocols such as TCP induce periodicity into the traffic and consequently into its PHY.

The learning algorithm assumes frequency vacancy can be represented by a cyclostationary random process, and from historic data it searches for a period length that maximizes the autocorrelation function. Once the cognitive radio has measured the channel access statistics of the licensed signal, it can compute the expected length of channel vacancy conditioned on the current channel state. Using this data, cognitive radios can make intelligent decisions that minimize possible interference while maximizing achievable capacity.

We simulate these algorithms in the presence of generic TDMA and CSMA networks, measuring cognitive radio capacity as a function of interference. We also simulate predictive dynamic spectrum access sample recordings of an IEEE 802.11 signal. Overall, using predictive techniques can greatly increase our ability to coexist with bursty licensed signals that do multiplexing in the time domain.

Additionally, we provide an initial look at using Hidden Markov Models (HMMs) for the same task. The models can be trained using historic data, and then used to predict future behavior of the primary user.

Section two outlines the basic algorithms used in cyclostationary detection for predictive dynamic spectrum access. Section three simulates these algorithms in conjunction with a TDMA network. Section four simulates them again for a CSMA network. Section five provides some interesting study for future work on using Hidden Markov Models for predictive dynamic spectrum access. Section six concludes.

2. CYCLOSTATIONARY DETECTION

A cyclostationary random process $\mathcal{R}$ is one where the process statistics are stationary over some period τ. That is,

$$\mathcal{R}(t) \overset{d}{=} \mathcal{R}(t+\tau)$$

An example would be a Gaussian random process whose mean or variance changes periodically over time. For example, $\mathcal{N}(\sin(t),\sigma^2)$ is a Gaussian random process whose mean varies as $\sin(t)$. In this case the period is $\tau = 2\pi$.

Cyclostationary detection has frequently been used for feature detection in cognitive radios [3,5]. It can provide very powerful algorithms for identifying the presence of a modulated signal within noise, and can therefore be used to determine if a channel is busy or idle. Note that for our application, we assume such a detector already exists (possibly cyclostationary in nature as well). Our aim is to

*This work was funded by the Laboratory for Telecommunications Sciences, US Department of Defense. The opinions expressed in this document represent those of the authors, and should not be considered an official opinion or endorsement by the Department of Defense or US Federal Government.

perform further analysis on the primary users' channel accesses to locate periodic patterns that will allow us to predict when to transmit.

For our application, we model the primary users' channel access patterns as cyclostationary random processes. Let $C(t)$ be a Bernoulli random variable such that

$$C(t) = \begin{cases} 1 & \text{if channel occupied at time } t \\ 0 & \text{if channel idle at time } t \end{cases}$$

Our goal is to determine the distribution of $C(t)$ with respect to period length τ, and use that information to determine optimal transmission times. Our first step is to compute τ. This can be accomplished using the autocorrelation function:

$$\tau = \arg\max_t \ \Re_{CC}(t)$$

Our next goal is to extract some useful statistics from $C(t)$. In particular, we need sufficient statistics that will help us decide optimal times to transmit such that we minimize the probability of interfering with the primary user. Depending on the type of statistic, differing amounts of data must be compiled to support it.

The first statistic we introduce is called $S_0(t)$, and let it be defined as

$$S_0(t) = \mathrm{P}[C(t) = 1] \qquad t \in [0, \tau)$$

This statistic gives us the probability that the channel is busy given a time index t, mod τ. Let `C(t)` be a recorded sample path from our channel with discrete samples `1` through `n`. Then `S0(t)` can be computed using the following algorithm:

```
for i=1:n
  S0(i mod tau) += C(i)*tau/n;
end
```

Given knowledge of $S_0(t)$, and we know our secondary transmission requires ℓ units of time, we can select:

$$t_{tx}(\ell) = \arg\min_{t \in [0,\tau]} \left(\int_t^{t+\ell} S_0(z)\, dz \right)$$

This will select a transmission time that minimizes the probability that the primary user will be transmitting during the secondary transmission.

One significant shortcoming of $S_0(t)$ is that it makes no use of the current time or channel state. In most radio transceivers, at any given time we can measure the current channel state to determine if it is busy or idle. Thus, we introduce our next statistic $S_1(t, t_0)$. First, let us define:

$$C^*(t_1, t_2) = 1 - \prod_{i=t_1}^{t_2} (1 - C(i))$$

Thus, it defines a function that is zero if and only if the channel is idle between times t_1 and t_2. And then:

$$S_1(t, t_0) = \mathrm{P}\left[C^*(t_0, t) = 1 \mid C(t_0) = 0, C(t_0 - 1) = 1\right]$$

In words, this statistic tells us, assuming the channel just became idle, the probability that during the time between now t_0 and some future time t, the channel will remain idle. This statistic can be created using the following algorithm:

```
idle_count=0;
i=2;
while (i<n)
  if C(i) && !C(i-1)
    idle_start(++idle_count)=i;
  end
  if !C(i) && C(i-1)
    d=i-idle_start(idle_count);
    idle_len(idle_count)=d;
  end
end

c=zeros(idle_count,1);
for i=1:idle_count
  s = idle_start(i) mod tau;
  t(s,++c(s)) = idle_len(i);
end

for i=1:tau
  p(:,i) = histogram(t(i,1:c(i)),[1:tau])/c(i);
  for j=1:tau
    S1(j,i) = sum(p(j:tau,i))
  end
  mean = mean(t(i,1:c(i)));
  std = sqrt(variance(t(i,1:c(i))));
  S1_ci(i) = mean – 2*std;
end
```

The output of our algorithm is the statistic `S1(t, t0)`, which gives us the desired statistic. Incidentally, it also gives us another useful piece of information, namely `S1_ci(t0)`. This value is a confidence interval telling us that with 98% probability, the channel will be idle for `S1_ci(t0)` units of time, assuming at time `t0` the channel became inactive.

In the next two sections, we simulate both a TDMA and a CSMA network, and apply the derived algorithms to those simulations.

3. COEXISTING WITH TDMA

TDMA is the perfect application for predictive dynamic spectrum access because of its inherent periodicity.

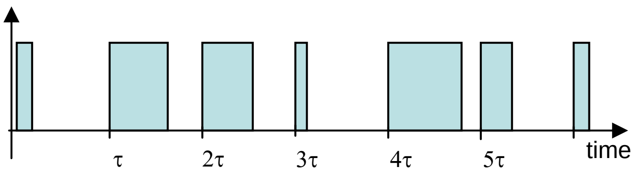

Fig 1. Diagram of TDMA channel accesses as a function of time.

As the above figure illustrates, TDMA is broken up into time slices. At the start of each time slice, a transmission occurs. The length of the transmission is a function of the amount of data that needs to be exchanged between devices in the network. More data means a longer channel access. The key feature that makes TDMA work well PDSA is that once the channel is idle, we know precisely how long it will be before the next transmission.

Due to the inherent periodicity, the autocorrelation function will yield the TDMA slice time as the time period τ. However, our time modulo τ might not be synchronized with the TDMA period. This generally is not a problem, as our algorithms don't require such synchronization.

Our TDMA simulation consists of *d* devices that have packets to transmit according to a rate *r* Poisson process that each require some fraction *p* of the time slice to transmit. During each time slice we compute the fraction of the slice necessary to transmit all the data, and from that determine the channel busy and idle times.

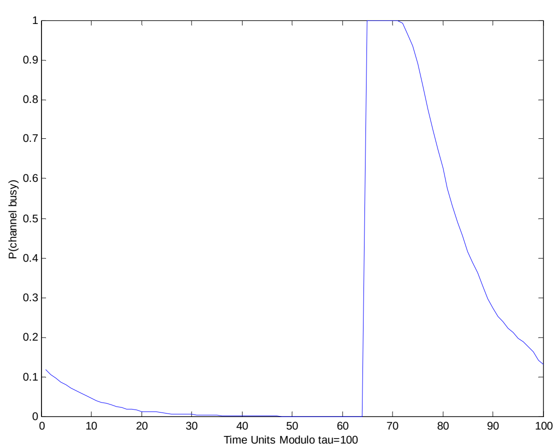

Fig 2. Plot of the basic statistic $S_0(t)$ for a TDMA network.

The above figure plots $S_0(t)$. We can see that as a function of time, modulo τ, the start of the TDMA time slice is at t = 65. This gives us valuable information about possible transmission times to minimize the probability of interference. In particular, from t = 40 to t = 64, we are nearly guaranteed to not interfere. However, as described in the last section, conditioning on the current channel state can further improve our performance.

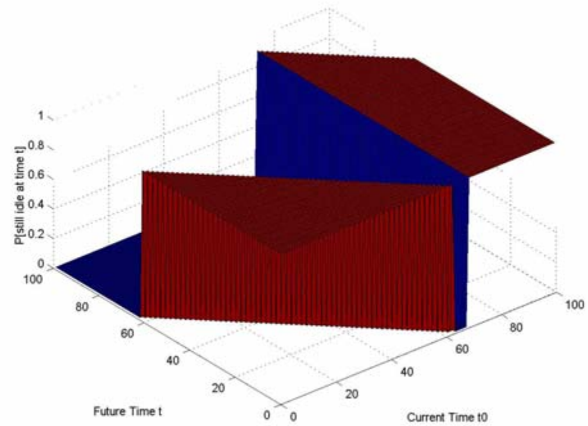

Fig 3. Plot of the extended statistic $S_1(t,t_0)$ for a TDMA network.

This figure plots $S_1(t,t_0)$, depicting the probability that the channel will still be idle at a future time t, given the current time is t_0 and the channel has just become idle. We can see this significantly increase our knowledge and ability to plan transmissions.

Using these techniques, with the exception of some overhead and guard time intervals, we can fully utilize all available secondary bandwidth, maximizing our capacity and causing absolutely no interference to the primary license holder.

4. COEXISTING WITH CSMA

Carrier Sense Multiple Access (CSMA) is another technique for multiplexing data from multiple users in the time domain. However, rather than having a very organized, slotted time system, devices have much more freedom in when they can transmit. The simplest technique is called "listen before talk" where you first make sure the channel is free, and if so you can transmit. Modern protocols such as IEEE 802.11 still use this basic philosophy, but have many added safeguards to minimize the probability of two people transmitting simultaneously, causing a collision.

For CSMA, our simulation is based on the IEEE 802.11b logs from IETF 62 [6]. Extracted from the logs were packet capture times and packet lengths. From that, channel access times were computed, and as a result we derived `idle_start` and `idle_len`.

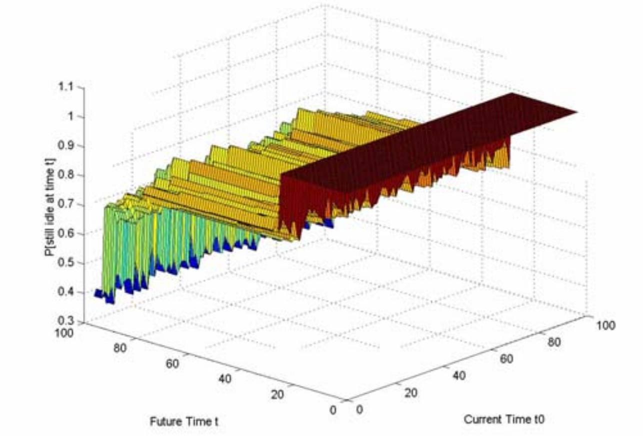

Fig 4. Plot of extended statistic for a CSMA network.

Using our algorithm, we can compute safe times directly after a transmission where a channel is guaranteed to be free. This is a common trait of modern CSMA networks because of their various timers to prevent collisions. As we can see, as time increases, the probability of the channel being idle decreases, as expected.

An interesting thing to notice in this example is we no longer see the diagonal cutting structure that we did for the TDMA network. The behavior is similar for any current time t_0. This is because at this level there is little periodicity in the network.

As we can see, our approach does not work as well with CSMA networks because they are inherently more random and cyclostationary analysis does not significantly increase our statistical understanding of the channel. However, the same algorithms can still find us good transmission times.

5. EXTENSIONS USING HIDDEN MARKOV MODELS

The analysis in this paper has focused mainly on the analysis of small-scale periodic patterns inside the length τ intervals. Another approach, that could be either an alternative or complementary to the approach in this paper, is to model the patterns of channel usage from one time period to another. This could involve allowing the period lengths to vary, which would lead to a more realistic model for non-TDMA protocols.

One approach that we expect to explore is using Hidden Markov Models (HMMs) to model channel usage patterns in a manner similar to that used to model speech production in speech processing applications [7]. HMMs have been used in speech processing for over a decade, and have already found many uses in the SDR world; for instance channel, agent, and burst error modeling and emulation [8,9].

Hidden Markov Models are statistical classification tools which model a process which is assumed to be approximated by a finite state Markov Model. The transition probabilities, states, and output probabilities are all assumed to be hidden. The Baum-Welch algorithm uses just the observable outputs of the system to compute values for the hidden model parameters which maximize:

$$P(\text{observed model} \mid \text{model})$$

5.1. Single model approach

In the simplest possible channel usage modeling situation we would have TDMA slots that either contain traffic or are essentially idle. In this case we could represent the channel usage as a stream of 1's and 0's, with 1's representing a busy slot and 0's an idle slot. We would want to train a Markov model whose states could output either a 0 or 1. Using the model in real time, in every slot we could determine which state the system is in. Some states with high probability of 0 (idle) would be considered safe to transmit in, while others with high probability of 1 (busy) would be considered unsafe.

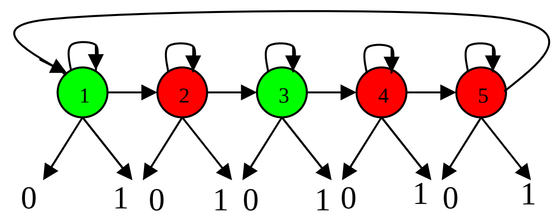

Fig 5. A simple circular feed-forward Markov model. Here states 1 and 3 would be considered safe to transmit, and states 2, 4, and 5 would be unsafe.

The choice of model (number of states) and any restrictions on topology could be crucial to the success of a system like this and is something that should be studied. Realistically one would not build a model with binary state outputs. More realistic outputs would be multi-dimensional feature vectors containing (not necessarily discrete) information extracted from monitoring the channel during some time period. If, as in the TDMA analysis in this paper, the activity during a time slot has an exponential distribution, the state outputs might correspond to parameters of an exponential distribution associated with the state.

5.2. Multiple model approach

Another approach that would be more analogous to speech processing would be to have several different models, some representing a mostly busy channel and some representing a mostly idle channel. The models could be pre-trained on training data and adapted in-service [1], or randomly initialized and trained using some unsupervised clustering method. Just as in continuous speech recognition, the most likely state in the most likely model at any given moment would be computed using the Viterbi algorithm. This approach would be easier to process in real-time than a large HMM, but would probably not work as well because it treats the busy and idle periods in isolation. It would be a generalization on the method evaluated in this study in the sense that it allows the periods τ to vary with each transmission, which is a more reasonable assumption for non-TDMA systems.

6. CONCLUSION

In this paper, we have introduced the idea of predictive dynamic spectrum access. The basic goal is to predict when the channel will be idle based on observations of the primary channel user. If we can predict idle periods, we can better plan our secondary channel accesses.

We developed algorithms for selecting channel access times using cyclostationary detection. Subsequent simulations using actual network data show that cyclostationary techniques work very well for protocols with inherent periodicity in their channel access patterns, such as TDMA. Other more random channel access techniques like CSMA can still benefit, but not to the extent of TDMA.

Lastly, we introduce the concept of using HMMs for modeling large-scale changes in channel access patterns. With further development, these could significantly increase our ability to predict channel accesses in CSMA-like networks. For example, many higher-level network protocols such as TCP operate on the basis of a state machine, the behavior of which could be learned by a HMM. Once the behavior is learned, it can be better predicted.

Overall, predictive techniques to dynamic spectrum access offer an interesting new research field. The ideas presented here offer a starting point for significant further research.

REFERENCES

[1] U. Bub, J. Kohler, B. Imperl, In-Service Adaptation of Multilingual Hidden Markov Models, ICASSP 1997, vol. 2, pp 1451–1454.

[2] D. Cabric, S. Mishra, D. Willkomm, R. Brodersen, and A. Wolisz. “A Cognitive Radio Approach for Usage of Virtual Unlicensed Spectrum.” IST Mobile Summit 2005.

[3] D. Cabric, S. Mishra, R. Brodersen, “Implementation issues in spectrum sensing for cognitive radios, Proceedings of the 38th Asilomar Conference on Signals, Systems and Computers.” November 2004, pp. 772–776.

[4] K. Challapali, S. Mangold, and Z. Zhong, “Spectrum Agile Radio: Detecting Spectrum Opportunities.” IEEE GlobeCOM 2004.

[5] A. Fehske, J. Gaeddert, J. Reed, “A new approach to signal classification using spectral correlation and neural networks.” Proceedings IEEE DySPAN, November 2005, pp. 144–150.

[6] A. Jardosh, K. Ramachandran, K. Almeroth, E. Belding-Royer. “Understanding Congestion in IEEE 802.11b Wireless Networks.” Proceedings of the Internet Measurement Conference, October 2005.

[7] L. Rabiner, “A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition”, Proceedings of the IEEE, Vol 77, No2, February 1989.

[8] T. Rondeau, C. Rieser, T. Gallagher, C. Bostian, “Online Modeling of Wireless Channels with Hidden Markov Models and Channel Impulse Responses for Cognitive Radios”, 2004 IEEE MTT-S Digest.

[9] A. Umbert, P Diaz, A Radio Channel Emulator for WCDMA Based on the Hidden Markov Model (HMM), IEEE VTC, vol. 5, pp 2173–2179, September 2000.

