

Utility Maximization in Cognitive Radio Networks using POMDP Approximate Value Iteration methods

Bharath Keshavamurthy, Nicolò Michelusi

Abstract—

*Index Terms—*Hidden Markov Model, Viterbi Algorithm, Expectation-Maximization, Belief Space, POMDP, Value Iteration, and the PERSEUS Algorithm **NM-comment: [too many keywords, reduce to 2-3]**

I. INTRODUCTION

NM-comment: [Please take a look at the comments below. There are a lot of details on the model that are still missing. And some parts are unclear.]

II. SYSTEM MODEL

NM-comment: [You need to introduce the model by words first, then introduce mathematical form.] NM-comment: [Also, you should add a figure of the model (use power point to create it)] We consider a network consisting of

A. Observation Model

$$y(n) = \sum_{m=0}^{M-1} h(m)x(n-m) + v(n) \quad (1)$$

Here, $y(n)$ is the wideband signal observed at the SU receiver expressed as a convolution of the PU signal $x(n)$ with the channel impulse response $h(n)$, added with a noise term $v(n)$. Equation (1) can be written in the frequency domain by taking a K -point DFT which decomposes the observed wideband signal into K discrete narrow-band components as shown below,

$$Y_k(i) = H_k X_k(i) + V_k(i) \quad (2)$$

where,

$i \in \{1, 2, 3, \dots, T\}$ represents the index of the observation
 $k \in \{1, 2, 3, \dots, K\}$ represents the index of the channel
 $V_k(i) \sim \mathcal{CN}(0, \sigma_V^2)$ represents the circular symmetric additive complex Gaussian noise sample, i.i.d across channel indices and across time indices. These noise samples are assumed to be independent of the occupancy state of the channels.
 $H_k \sim \mathcal{CN}(0, \sigma_H^2)$ represents the k^{th} DFT coefficient of the impulse response $h(n)$ of the channel in between the PU and the SU receiver; **we model it as a another** circular symmetric complex Gaussian random variable, i.i.d across channel indices **NM-comment: [what about time?]** with variance σ_H^2 **NM-comment: [note that the model (1) introduces**

channel correlation in the frequency domain, hence assuming i.i.d. is not coherent with (1).]. These impulse response samples are also assumed to be independent of the occupancy state of the channels. We define the set **NM-comment: [use \mathcal{B} for sets]** as the set of channels obtained by discretizing the spectrum of interest, i.e. $B = \{b_1, b_2, b_3, \dots, b_K\}$ **NM-comment: [why do you need a new set when you already have the set of indices $\{1, \dots, K\}$ corresponding to channel indices?].** The PU occupancy behavior in each sub-band ~~$b_k \in B$~~ is modelled as $X_k \in \{0, 1\}$ taking two possible values 0 (*Idle*) and 1 (*Occupied*). Therefore, the PU occupancy behavior in the entire wideband spectrum of interest discretized into narrow-band frequency components can be modeled as ~~the vector a vector of size $|B| = K$ such that,~~

$$\vec{X} = [X_1, X_2, X_3, \dots, X_K]^T \in \{0, 1\}^K \quad (3)$$

NM-comment: [where is the time dependence?] NM-comment: [where do you incorporate the transmission power of the PU? Also, in this model you are assuming that the PU is using an OFDMA access strategy, so that indeed the spectrum occupancy vector has a "sparse" support. Need to clarify that.]

We model the spectrum occupancy dynamics as a Markov process with the following transition probabilities. Given \vec{X}_i at time i , \vec{X}_{i+1} is independent of the past $\vec{X}_j, j < i$, i.e.

$$\mathbb{P}(\vec{X}_{i+1} = \vec{x}_{i+1} | \vec{X}_j = \vec{x}_j, \forall j \leq i) = \mathbb{P}(\vec{X}_{i+1} = \vec{x}_{i+1} | \vec{X}_i = \vec{x}_i).$$

In addition, \vec{X}_i exhibits a Markov structure across frequency channels, according to the model

$$\begin{aligned} \mathbb{P}(\vec{X}(i+1) = \vec{x}(i+1) | \vec{X}(i) = \vec{x}(i)) \\ = \prod_{k=1}^K \mathbb{P}(X_k(i+1) = x_k(i+1) | X_k(i) = x_k(i), X_{k-1}(i) = x_{k-1}(i)). \end{aligned}$$

NM-comment: [It was my understanding that in the model we discussed $\vec{X}_k(i+1)$ (occupancy of channel k at time $i+1$) depends only on the occupancy at the PREVIOUS TIME i , channels k and $k-1$, see above. However, the model you present below is different. Please clarify.]

B. Correlation Model

The true states $\{\vec{X}_i, i = 1, \dots, T\}$ encapsulate the actual occupancy behavior of the PU **across frequency channels and across time**, and the measurements at the SU $\{Y_k(i), \forall k, i\}$ are noisy observations of these true states, which are modeled to be the observed states of a Hidden Markov Model.

This research has been funded by —.

Authors are with the School of Electrical and Computer Engineering, Purdue University. email: {bkeshava,michelusi}@purdue.edu.

NM-comment: [I moved this paragraph earlier] For some sub-band $j \in \{2, 3, 4, \dots, K\}$ and time index $i \in \{1, 2, 3, \dots, T\}$, the system is assumed to satisfy the Markov property as shown below,

$$\mathbb{P}(X_j(i)|X_{j-1}(i), X_{j-2}(i), \dots, X_1(i)) = \mathbb{P}(X_j(i)|X_{j-1}(i))$$

NM-comment: [What is the dependence across time? I don't understand, are you satisfying the Markov property over frequency j but not over time i ? I thought the model we discussed was different, i.e. Markov across both time and frequency. See the model I wrote above, and please let me know if what you are doing is different.] NM-comment: [You should also provide a figure of this Markov structure. Nodes represent states at different times and arrows Markov dependence.] And, we will use $\mathbb{P}(X_1(i))$ for $j = 1$. Now, let's expand on the previously discussed observation model.

Taking the expectation operator on both sides of equation (2) given X_k has realized as x_k , we have,

$$\begin{aligned} \mathbb{E}[Y_k(i)|X_k(i) = x_k] &= \mathbb{E}[H_k x_k] + \mathbb{E}[V_k(i)] \\ \mathbb{E}[Y_k(i)|X_k(i) = x_k] &= \mathbb{E}[H_k] \mathbb{E}[x_k] + \mathbb{E}[V_k(i)] \\ \mathbb{E}[Y_k(i)|X_k(i) = x_k] &= 0 \end{aligned} \quad (4)$$

because, $V_k(i) \sim \mathcal{CN}(0, \sigma_V^2)$ and $H_k \sim \mathcal{CN}(0, \sigma_H^2)$. **NM-comment: [No need for this in the paper.]**

Furthermore, the variance of $Y_k(i)$ given X_k at observation cycle i has realized as x_k , is calculated to be,

$$\begin{aligned} \text{Var}(Y_k(i)|X_k(i) = x_k) &= \mathbb{E}[|H_k X_k(i) + V_k(i)|^2 | X_k(i) = x_k] \\ \text{Var}(Y_k(i)|X_k(i) = x_k) &= \mathbb{E}[|H_k X_k(i)|^2 + |V_k(i)|^2 + 2\Re(H_k X_k(i) V_k^*(i)) | X_k(i) = x_k] \\ \text{Var}(Y_k(i)|X_k(i) = x_k) &= \sigma_H^2 x_k + \sigma_V^2 \end{aligned} \quad (5)$$

For the first part of our paper **NM-comment: [what about the second part?]**, we assume that the temporal dynamics of the PU Occupancy are slower than the SU's process times. **We approximate this scenario by assuming that** ~~In other words, for the first part of our paper, we assume that~~ the PU is static during our evaluation period. We can incorporate the above assumption into our correlation model as shown below by eliminating the time dependence.

$$\mathbb{P}(X_j|X_{j-1}, X_{j-2}, \dots, X_1) = \mathbb{P}(X_j|X_{j-1}), \text{ for } j > 1,$$

And, we will continue to use $\mathbb{P}(X_1)$ for $j = 1$. Now, we know that,

$$\vec{X} = [X_1, X_2, X_3, \dots, X_K]^T$$

which realizes as $\vec{x} = [x_1, x_2, x_3, \dots, x_K]^T$, so,

$$\mathbb{P}(\vec{X} = \vec{x}) = \mathbb{P}(X_1 = x_1) \prod_{k=2}^K \mathbb{P}(X_k = x_k | X_{k-1} = x_{k-1}) \quad (6)$$

Since $x_k \in \{0, 1\}$, for $k \in \{1, 2, 3, \dots, K\}$, let,

$$\mathbb{P}(X_k = 1) \triangleq \Pi, \forall k$$

NM-comment: [why is Π the same for all k ? Not necessarily, it depends on how you choose p and q ...] Furthermore, let,

$$\mathbb{P}(X_k = 1 | X_{k-1} = 0) \triangleq p, \forall k$$

And,

$$\mathbb{P}(X_k = 0 | X_{k-1} = 1) \triangleq q, \forall k$$

NM-comment: [This is part of your model, should be presented way earlier when you introduce PU dynamics.] From the above definitions, we have,

$$\mathbb{P}(X_k = 1) = \Pi = \frac{p}{p+q}, \forall k$$

Moreover, we also assume that the Markov Property is satisfied when we traverse the spectrum in the descending order of the channel indices, i.e, the reverse direction. **NM-comment: [This assumption constrains your Markov chain and you need stronger properties on p and q in order to get this.. please clarify. Technically, you need to prove that this property holds in our model, rather than stating it as an assumption, i.e, show that**

$$\mathbb{P}(\vec{X} = \vec{x}) = \mathbb{P}(X_K = x_K) \prod_{k=1}^{K-1} \mathbb{P}(X_k = x_k | X_{k+1} = x_{k+1})$$

given the assumptions that we already have.] Mathematically,

$$\mathbb{P}(\vec{X} = \vec{x}) = \mathbb{P}(X_K = x_K) \prod_{k=1}^{K-1} \mathbb{P}(X_k = x_k | X_{k+1} = x_{k+1})$$

We now proceed with our discussions assuming that there is only one Primary User (PU), i.e. licensed incumbent in the wideband spectrum of interest and that there is only one Secondary User (SU) learning to intelligently access the *spectrum holes* or *white spaces* both spatially and temporally.

NM-comment: [this discussion should go at the beginning of the system model, not here, since it is a general description of the problem we are trying to solve.] In the next section of our paper, assuming a static PU, model knowledge, and complete observations, **NM-comment: [didn't we say we can handle incomplete observations, i.e. we have a constraint in the number of channels we can sense? Just go straight with that, and introduce in the system model.. This is a paper, there is no need to go through step-by-step model improvements, but we should tackle the complete model directly (unless we have special reason to not do so, such as special algorithms that we need in certain conditions).]** we discuss an algorithm to estimate the PU occupancy behavior. Later, we methodically relax these assumptions and detail occupancy behavior estimation results and optimal POMDP policy search methods for cases where we have incomplete observations, dynamic PU occupancy behavior, and no model information.