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

Bharath Keshavamurthy, Nicolò Michelusi

***Abstract*—**

***Index Terms*—Hidden Markov Model, POMDP, and the PERSEUS Algorithm**

I. Introduction

II. System Model

A. The Observation Model

We consider a network consisting of one licensed user termed the Primary User (PU) and one cognitive radio node termed the Secondary User (SU) which is equipped with a spectrum sensor. The SU should learn to intelligently access spectrum holes (white-spaces) in order to maximize its throughput while maintaining strict non-interference compliance with incumbent transmissions. The wideband signal observed at the SU receiver is denoted as $y(n)$ and is given by,

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

Here, $y(n)$ is 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,

$$Y_k(i) \ = \ H_k X_k(i) + V_k(i) \tag{2}$$

where,
$i \in \{1, 2, 3, ..., T\}$ represents the index of the observation
$k \in \{1, 2, 3, ..., K\}$ 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 zero-mean circular symmetric complex Gaussian random variable with variance σ_H^2. These impulse response samples are assumed to be independent of the occupancy state of the channels. The PU occupancy behavior in each sub-band 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 at time index i can be modelled as a vector as shown below.

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

We assume that the PU employs an OFDMA access strategy and therefore, this spectrum occupancy vector has a sparse support.

B. The Correlation Model

We model the spectrum occupancy dynamics as a Markov process with the following transition model. Given $\vec{X}(i)$, the spectrum occupancy state at time index i, $\vec{X}(i+1)$ is independent of the past, $\vec{X}(j),\ j < i;\ j, i \in \{1, 2, 3, ..., T\}$, i.e.

$$\mathbb{P}(\vec{X}(i+1)|\vec{X}(j),\ \forall j \leq i) \ = \ \mathbb{P}(\vec{X}(i+1)|\vec{X}(i)) \tag{4}$$

Additionally, the spectrum occupancy vector $\vec{X}(i)$ exhibits Markovian correlation across the sub-bands as,

$$\mathbb{P}(\vec{X}(i+1)|\vec{X}(i)) \ = \ \prod_{k=1}^{K} \mathbb{P}(X_{k+1}(i+1)|X_{k+1}(i), X_k(i)) \tag{5}$$

The true states encapsulate the actual occupancy behavior of the PU and the measurements at the SU are noisy observations of these true states which are modelled to be the observed states of a Hidden Markov Model. Owing to physical design limitations at the SU's spectrum sensor, not all sub-bands in the discretized spectrum can be sensed. Given this constraint, we model the emission process of the HMM as shown below. The observation vector at time index i is given by,

$$\vec{Y}(i) \ = \ [y_1(i), y_2(i), \phi, y_4(i), ..., \phi, y_{K-1}(i), y_K(i)]^T \tag{6}$$

where, $y_k(i) = \phi$ indicates that the SU did not sense sub-band k at time index i. Therefore, the observation probability termed as the emission probability of the HMM is given by,

$$m_r(y_k(i)) \triangleq \mathbb{P}(y_k(i)|x_k(i) = r)$$

where,

$$\begin{aligned} m_r(y_k(i)) \ &\sim \mathcal{CN}(0, \sigma_H^2 r + \sigma_V^2),\ \text{if } y_k(i) \neq \phi \\ m_r(y_k(i)) \ &= \ 1,\ \text{if } y_k(i) = \phi \end{aligned} \tag{7}$$

Now, we model the spectrum access scheme of the SU as a Partially Observable Markov Decision Process (POMDP) wherein the goal of the POMDP agent is to devise an optimal sensing and access policy in order to maximize its throughput while maintaining strict non-interference compliance with incumbent transmissions.

This research has been funded by —–.

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

C. The POMDP Agent Model

The agent's limited observational capabilities coupled with its noisy observations result in an increased level of uncertainty at the agent's end about the occupancy state of the spectrum under consideration and the exact effect of executing an action on the radio environment. The transition model of the underlying MDP as described in the Correlation Model of this paper, is denoted by A and is learnt by the agent by interacting with the radio environment. The emission model B is given by (7). We model the POMDP as a tuple $(\mathcal{X}, \mathcal{A}, \mathcal{Y}, \mathcal{B}, A, B)$ where, $\mathcal{X}$ represents the state space of the underlying MDP with states $\vec{x}$ which are realizations of the spectrum occupancy vector as given by (3) and $|\mathcal{X}| = 2^K$, $\mathcal{A}$ represents the action space of the agent considering the imposed sensing limitations, $\mathcal{Y}$ represents the observation space of the agent based on the Observation Model outlined earlier in the paper, and $\mathcal{B}$ represents the belief space of the agent.
The run-time or interaction time of the agent is quantized into discrete time-steps termed as episodes. At the beginning of each episode, the agent executes an action $a \in \mathcal{A}$, observes $\vec{y} \in \mathcal{Y}$, and updates it belief $\forall \vec{x}'$ as follows.

$$b_a^{\vec{y}}(\vec{x}') \;=\; \mathbb{P}(\vec{x}'|\vec{y}, a, \vec{b}) \;=\; \frac{\mathbb{P}(\vec{y}|\vec{x}', a)}{\mathbb{P}(\vec{y}|a, \vec{b})} \sum_{\vec{x} \in \mathcal{X}} \mathbb{P}(\vec{x}'|\vec{x}, a) b(\vec{x}) \tag{8}$$

where, $\mathbb{P}(\vec{y}|a, \vec{b})$ is the normalization constant given by,

$$\mathbb{P}(\vec{y}|a, \vec{b}) \;=\; \sum_{\vec{x}' \in \mathcal{X}} \mathbb{P}(\vec{y}|\vec{x}', a) \sum_{\vec{x} \in \mathcal{X}} \mathbb{P}(\vec{x}'|\vec{x}, a) b(\vec{x}) \tag{9}$$

$\vec{b} \in \mathcal{B}$ represents the belief vector of the agent, i.e. a probability distribution over all states, in the previous time-step, $b(\vec{x}) \in \vec{b}$ is termed the belief and it represents the degree of certainty assigned to world state $\vec{x} \in \mathcal{X}$ by the belief vector $\vec{b}$. The belief, by definition being a probability measure, has to satisfy the Kolmogorov's axioms, i.e.

$$\begin{aligned} \sum_{\vec{x} \in \mathcal{X}} b(\vec{x}) \;&=\; 1 \\ 0 \leq b(\vec{x}) &\leq 1 \end{aligned} \tag{10}$$

Considering sub-band k, the set of available actions to the agent in an episode i is given by,

$$a_k(i) \;=\; \begin{cases} 1, & \text{sense and access sub-band } k, \\ 0, & \text{do nothing with respect to sub-band } k \end{cases} \tag{11}$$

We define $\kappa < K$ to be the number of the channels the SU can sense simultaneously. Based on this sensing constraint, the size of the action space is given by, $\mathcal{A} = K^\kappa$. The reward to the agent is modelled as follows based on the number of truly idle sub-bands found which accounts for the throughput maximization aspect of our end-goal and a penalty for missed detections which accounts for the incumbent non-interference constraint.

$$R(\vec{x}(i), a(i)) \;=\; (1 - P_{FA}(i)) + \lambda P_{MD}(i) \tag{12}$$

where, $P_{FA}(i)$ represents the False Alarm Probability across all channels in episode i, $P_{MD}(i)$ represents the Missed Detection Probability across all channels in episode i, and $\lambda < 0$ represent the cost term penalizing the agent for missed detections, i.e. interference with the incumbent. The action policy of the agent $\pi : \mathcal{B} \rightarrow \mathcal{A}$ maps the belief vectors $\vec{b} \in \mathcal{B}$ to actions $a \in \mathcal{A}$ and is characterized by a Value Function,

$$V^\pi(\vec{b}) \;=\; \mathbb{E}_\pi\Big[\sum_{i=0}^{\infty} \gamma^i R(\vec{b}_i,\ \pi(\vec{b}_i))|\vec{b}_0 = \vec{b}\Big] \tag{13}$$

where, $0 < \gamma < 1$ is the discount factor, $\pi(\vec{b}_i)$ is the action taken by the agent in episode i under policy π, and $\vec{b}_0$ is the initial belief vector. The optimal policy π^* specifies the optimal action to take in the current episode assuming that the agent behaves optimally in future episodes as well. It is evident from equation (13) that we have an infinite-horizon discounted reward problem formulation and in order to solve for the optimal policy we need to solve the modified Bellman equation given as follows. $\forall \vec{b} \in \mathcal{B}$,

$$V^*(\vec{b}) \;=\; \max_{a \in \mathcal{A}} \Big[\sum_{\vec{x} \in \mathcal{X}} R(\vec{x}, a) b(\vec{x}) + \gamma \sum_{\vec{y} \in \mathcal{Y}} \mathbb{P}(\vec{y}|a, \vec{b}) V^*(\vec{b}_a^{\vec{y}})\Big] \tag{14}$$

Given the high dimensionality of the spectrum sensing and access problem, i.e. the number of states of the underlying MDP scales exponentially with the number of sub-bands, solving equation (14) using Exact Value Iteration and Policy Iteration algorithms is computationally infeasible. Additionally, solving for the optimal policy from equation (14) requires prior knowledge about the underlying MDP's transition model. Therefore, in this paper we present a framework to estimate the transition model of the underlying MDP and then utilize this learned model to solve for the optimal policy by employing Randomized Point-Based Value Iteration techniques, namely, the PERSEUS algorithm.

III. Prepare Your Paper Before Styling

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections III-A–III-E below for more information on proofreading, spelling and grammar.

Keep your text and graphic files separate until after the text has been formatted and styled. Do not number text heads—LaTeX will do that for you.

A. Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, ac, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

B. Units

- Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as "3.5-inch disk drive".
- Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.
- Do not mix complete spellings and abbreviations of units: "Wb/m^2" or "webers per square meter", not "webers/m^2". Spell out units when they appear in text: ". . . a few henries", not ". . . a few H".
- Use a zero before decimal points: "0.25", not ".25". Use "cm^3", not "cc".)

C. Equations

Number equations consecutively. To make your equations more compact, you may use the solidus (/), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in:

$$a + b = \gamma \tag{15}$$

Be sure that the symbols in your equation have been defined before or immediately following the equation. Use "(15)", not "Eq. (15)" or "equation (15)", except at the beginning of a sentence: "Equation (15) is . . ."

D. LaTeX-Specific Advice

Please use "soft" (e.g., `\eqref{Eq}`) cross references instead of "hard" references (e.g., `(1)`). That will make it possible to combine sections, add equations, or change the order of figures or citations without having to go through the file line by line.

Please don't use the `{eqnarray}` equation environment. Use `{align}` or `{IEEEeqnarray}` instead. The `{eqnarray}` environment leaves unsightly spaces around relation symbols.

Please note that the `{subequations}` environment in LaTeX will increment the main equation counter even when there are no equation numbers displayed. If you forget that, you might write an article in which the equation numbers skip from (17) to (20), causing the copy editors to wonder if you've discovered a new method of counting.

BibTeX does not work by magic. It doesn't get the bibliographic data from thin air but from .bib files. If you use BibTeX to produce a bibliography you must send the .bib files.

LaTeX can't read your mind. If you assign the same label to a subsubsection and a table, you might find that Table I has been cross referenced as Table IV-B3.

LaTeX does not have precognitive abilities. If you put a `\label` command before the command that updates the counter it's supposed to be using, the label will pick up the last counter to be cross referenced instead. In particular, a `\label` command should not go before the caption of a figure or a table.

Do not use `\nonumber` inside the `{array}` environment. It will not stop equation numbers inside `{array}` (there won't be any anyway) and it might stop a wanted equation number in the surrounding equation.

E. Some Common Mistakes

- The word "data" is plural, not singular.
- The subscript for the permeability of vacuum μ_0, and other common scientific constants, is zero with subscript formatting, not a lowercase letter "o".
- In American English, commas, semicolons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
- A graph within a graph is an "inset", not an "insert". The word alternatively is preferred to the word "alternately" (unless you really mean something that alternates).
- Do not use the word "essentially" to mean "approximately" or "effectively".
- In your paper title, if the words "that uses" can accurately replace the word "using", capitalize the "u"; if not, keep using lower-cased.
- Be aware of the different meanings of the homophones "affect" and "effect", "complement" and "compliment", "discreet" and "discrete", "principal" and "principle".
- Do not confuse "imply" and "infer".
- The prefix "non" is not a word; it should be joined to the word it modifies, usually without a hyphen.
- There is no period after the "et" in the Latin abbreviation "et al.".
- The abbreviation "i.e." means "that is", and the abbreviation "e.g." means "for example".

An excellent style manual for science writers is [7].

F. Authors and Affiliations

The class file is designed for, but not limited to, six authors. A minimum of one author is required for all conference articles. Author names should be listed starting from left to right and then moving down to the next line. This is the author sequence that will be used in future citations and by indexing services. Names should not be listed in columns nor group by affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

G. Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is "Heading 5". Use "figure caption" for your Figure captions, and "table head" for your table title. Run-in heads, such as "Abstract", will require you to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced.

H. Figures and Tables

a) Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables.

TABLE I
TABLE TYPE STYLES

Table Head	**Table Column Head**		
	Table column subhead	***Subhead***	***Subhead***
copy	More table copy[a]		

[a]Sample of a Table footnote.

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity "Magnetization", or "Magnetization, M", not just "M". If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write "Magnetization (A/m)" or "Magnetization {A[m(1)]}", not just "A/m". Do not label axes with a ratio of quantities and units. For example, write "Temperature (K)", not "Temperature/K".

ACKNOWLEDGMENT

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression "one of us (R. B. G.) thanks ...". Instead, try "R. B. G. thanks...". Put sponsor acknowledgments in the unnumbered footnote on the first page.

REFERENCES

Please number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use "Ref. [3]" or "reference [3]" except at the beginning of a sentence: "Reference [3] was the first ..."

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors' names; do not use "et al.". Papers that have not been published, even if they have been submitted for publication, should be cited as "unpublished" [4]. Papers that have been accepted for publication should be cited as "in press" [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

REFERENCES

[1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955.
[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
[3] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
[4] K. Elissa, "Title of paper if known," unpublished.
[5] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.
[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
[7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.