A Novel Method for Dynamic Multichannel Access Based on Whittle's Index and KS Sensing Model

Group14: Bo Chen, Wenji Qian, Zizhong Cao

***Abstract*—In the study of cognitive radio, the spectrum access is one of the most important aspect. And every SU(Secondary User) should not only consider how to make its own benefit most, but also pay attention to the influence on other SUs in order to achieve the relatively large benefits of the whole group. Thus game-theory and corporation problems should be taken into consideration. In this paper, we provide a useful method, the whittle index for each SU to judge how to make the decision on sensing while considering other SUs and its own benefits. We use the Markovian Chain model and take the real difference between sensing and transforming time into consideration. And we assume that the sensing process obey the rule of KS(keep sensing) model. We build such model and get the proper sensing channels number one sensor should take in one sensing period by mathematical calculation and deduction. Through which we can predict the proper channels number under certain conditions. And then we will prove how this algorithm works by giving its upper and lower bound of the benefits and make comparison to the optimal solution. And we will further illustrate the simulation results.**

***Index Terms*—Markovian Chain, KS model, Whittle Index**

I. Introduction

Todays wireless networks are characterized by a fixed spectrum assignment policy. However, a large portion of the assigned spectrum is used sporadically and geographical variations in the utilization of assigned spectrum ranges from 15% to 85% with a high variance in time. The limited available spectrum and the inefficiency in the spectrum usage necessitate a new communication paradigm to exploit the existing wireless spectrum opportunistically. This new networking paradigm is referred to as NeXt Generation (xG) Networks as well as Dynamic Spectrum Access (DSA) and cognitive radio networks.

A. Cognitive Radio

Cognitive radio techniques provide the capability to use or share the spectrum in an opportunistic manner. Dynamic spectrum access techniques allow the cognitive radio to operate in the best available channel. More specifically, the cognitive radio technology will enable the users to (1) determine which portions of the spectrum is available and detect the presence of licensed users when a user operates in a licensed band (spectrum sensing), (2) select the best available channel (spectrum management), (3) coordinate access to this channel with other users (spectrum sharing), and (4) vacate the channel when a licensed user is detected (spectrum mobility).

B. Hierarchical Access Model

This model adopts a hierarchical access structure with primary and secondary users. The basic idea is to open licensed spectrum to secondary users while limiting the interference perceived by primary users (licensees). Two approaches to spectrum sharing between primary and secondary users have been considered: Spectrum underlay and spectrum overlay. The underlay approach imposes severe constraints on the transmission power of secondary users so that they operate below the noise floor of primary users. By spreading transmitted signals over a wide frequency band (UWB), secondary users can potentially achieve short-range high data rate with extremely low transmission power. Based on a worst-case assumption that primary users transmit all the time, this approach does not rely on detection and exploitation of spectrum white space. Spectrum overlay was first envisioned by Mitola under the term spectrum pooling and then investigated by the DARPA Next Generation (XG) program under the term opportunistic spectrum access. Differing from spectrum underlay, this approach does not necessarily impose severe restrictions on the transmission power of secondary users, but rather on when and where they may transmit. It directly targets at spatial and temporal spectrum white space by allowing secondary users to identify and exploit local and instantaneous spectrum availability in a non-intrusive manner.

C. Restless Multi-armed Bandit Problem

Restless Multi-armed Bandit Processes (RMBP) are generalizations of the classical Multi-armed Bandit Processes (MBP), which have been studied since 1930's. In an MBP, a player, with full knowledge of the current state of each arm, chooses one out of N arms to activate at each time and receives a reward determined by the state of the activated arm. Only the activated arm changes its state according to a Markovian rule while the states of passive arms are frozen. The objective is to maximize the long-run reward over the infinite horizon by choosing which arm to activate at each time.
Whittle generalized MBP to RMBP by allowing multiple ($K \geq 1$) arms to be activated simultaneously and allowing passive arms to also change states. Either of these two generalizations would render Gittins' index policy suboptimal in general, and finding the optimal solution to a general RMBP has been shown to be PSPACE-hard.

D. The Gilber-Elliot channel model.

Consider the problem of probing N independent Markov chains. Each chain has two states. .good. and .bad.. with different transition probabilities across chains (see Fig. 1). At each time, a player can choose K ($1 \leq K < N$) chains to probe and receives reward determined by the states of the probed chains.

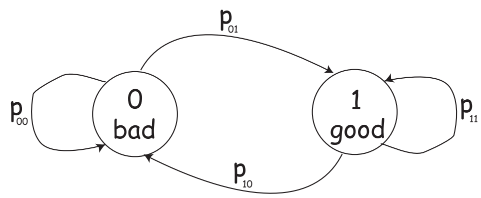

Fig. 1. The Gilber-Elliot channel model

E. Game Theory

Some concepts of game theory date back centuries, but modern game theory began in the mid-20th century. One of its earliest modern making by aggressive superpowers. A more enduring application has been as a powerful array of techniques for modeling economic behavior. The basic unit of game theory is, of course, the game. A game has three basic elements:

- A description of strategic interaction between players
- A set of constraints on the actions the players can take
- A specification of the interests of the players

Games are usually represented in one of two forms: the normal form and the extensive form. The normal form game for two players is represented as a bi-matrix. An extensive form game is depicted as a tree, where each node represents a decision point for one of the players. The normal form is easier to analyze, but the extensive form captures the structure of a real game in time.

F. KS Sensing Model

KS Scheme (Keep-Sensing-if-Busy): After a vacation, the SU(Secondary User) senses the channel. If the channel is idle, the SU transmits a packet and then starts vacation. If the SU senses the channel busy, it keeps sensing until the channel is idle. Then, the SU transmits a packet and starts a random vacation of length V2.

II. Related Work

Dynamic spectrum access among cognitive radios can be realized by an adaptive, game theoretic learning perspective. Spectrum-agile cognitive radios compete for channels temporarily vacated by licensed primary users in order to satisfy their own demands while minimizing interference. Reference [1] applies an adaptive regret based learning procedure which

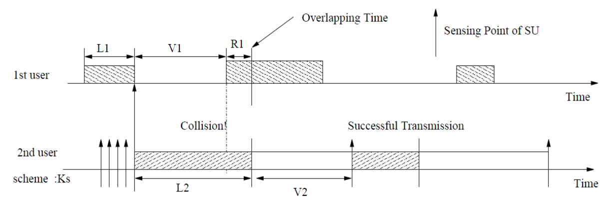

Fig. 2. KS sensing model

tracks the set of correlated equilibria of the game, treated as a distributed stochastic approximation. And this illustrates that by adding some degradation factor we can recalculate the value of the decision and make the predicted benefits of whole group largest by solving the differential equations. By the given degradation index we can get the most proper channel numbers to sense. In Reference [8], it shows the difference of both non-cooperative and cooperative game theory in static and dynamic settings. Careful attention is given to techniques for demonstrating the existence and uniqueness of equilibrium in non-cooperative games. And there are more about the game theory from Reference [16] about the applications of game theory to supply chain analysis and outlines game-theoretic concepts that have potential for future application.

Another aspect of the spectrum access includes the random process analysis. Reference [17] considers a scenario where secondary users can opportunistically access unused spectrum vacated by idle primaries. Supposing the PU's starting using one channel obeying poisson distribution, we can get the max transformation rate under certain limited collision rate by Probability Theory. And Reference [18] develops opportunistic scheduling policies for cognitive radio networks that maximize the throughput utility of the secondary (unlicensed) users subject to maximum collision constraints with the primary (licensed) users. It considers a cognitive network with static primary users and potentially mobile secondary users. The model assumes state whether the channel is idle is a kind of Markov Chain. The paper uses the technique of Lyapunov Optimization to design an online flow control, scheduling and resource allocation algorithm that meets the desired objectives and provides explicit performance guarantees.

In Reference [3],the spectrum access is optimal in that it strikes a balance between two conflicting needs: keeping spectrum assessment overhead low while increasing the likelihood of discovering spectrum opportunities. It study the effect of several network parameters, such as the primary traffic load, the secondary traffic load, and the collaboration level of the sensing method.

Reference [4] deal with multi-armed bandit problem for a gambler is to decide which arm of a K-slot machine to pull to maximize his total reward in a series of trials. It provides a preliminary empirical evaluation of several multi-armed bandit algorithms. It also describes and analyzes a new algorithm, Poker (Price Of Knowledge and Estimated Reward) whose performance compares favorably to that of other existing algorithms in several experiments.

In Reference [9],it considers a class of restless multi-armed bandit problems (RMBP). And it establishes indexability and obtain Whittle's index in closed-form for both discounted and average reward criteria. These results lead to a direct implementation of Whittle's index policy with remarkably low complexity. Furthermore, it has a semi-universal structure that obviates the need to know the Markov transition probabilities. In Reference [12] it provides an method to jointly detect the primary signals over multiple frequency bands rather than over one band at a time. By exploiting the hidden convexity in the seemingly nonconvex problems, optimal solutions can be obtained for multiband joint detection under practical conditions with certain constraints. To address this issue by exploiting the spatial diversity, a cooperative wideband spectrum sensing scheme refereed to as spatial-spectral joint detection is proposed, which is based on a linear combination of the local statistics from multiple spatially distributed cognitive radios. The cooperative sensing problem is also mapped into an optimization problem, for which suboptimal solutions can be obtained through mathematical transformation under conditions of practical interest.

And the most important paper we pay attention to is Reference [20], which make a combination of the Reference [9] and Reference [12]. When arms are stochastically identical, it shows that Whittle's index policy is optimal under certain conditions. Like Reference[9], it need some background knowledge of Markov transition probabilities. The optimality and the semi-universal structure result from the equivalency between Whittle's index policy and the myopic policy established in this work. For non-identical arms, it develops efficient algorithms for computing a performance upper bound given by Lagrangian relaxation.

Other References such as [2], [5], [6], [7], [10], [11], [13], [14], [15], [19] help to clarify the cognitive radio and dynamic spectrum access and offer more relevant background information and knowledge into our research.

III. Problem Statement and Formulation

A. Multi-channel Opportunistic Access

1) the classic Gilbert-Elliot channel model: Consider N independent Gilbert-Elliot channels, each with transmission rate B_i(i=1,$\cdots$,N). The classic model is shows in the [20],The state of channel i–"good"(1) or "bad"(0)– evolves from slot to slot as a Markov chain, as showed in Figure 1.

2) the advanced Gilbert-Elliot channel model: The basic Gilber-Elliot channel model is quite simple and is capable of representing the actual conditions to some extent. However, this basic model assumes that the time period of sensing equals that of transmission, which might not be the best way, comparing the loss of efficiency when keeping idle with the cost of keeping sensing. Thus, if we let the cognitive radio sense more frequently in the "bad" case, the total efficiency is likely to improve quite much. Based on this intension, we devide the time slots into smaller pieces and try to incorporate the KS sensing model as showed in Figure 2 into the Gilber-Elliot channel model.

Still we assume that the channel condition remains the same as the Gilbert-Elliot channel, the changes happen to the sensing mode. Assume that n represents the longest time for one data maker to transmit compared to one sensing period. And assume that before the data sending period n the SU will not sensing the channel even it has finished sending the data. Based on the assumptions above, we put forward an advanced model as showed in the Figure 3.

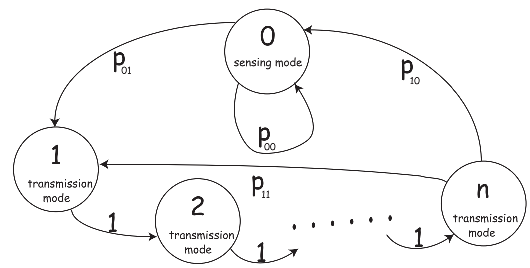

Fig. 3. the advanced Gilbert-Elliot model

At the beginning of slot t, if the state S_i(t)=1 of the sensed cannel is 1, the the SU transmits and collects B_i units of reward in this channel and the transmission period last for n_i time slot, which is no more than n. Otherwise, the user collects no reward in this channel.

Our objective is to maximize the expected long-run reward by designing a sensing policy that which channels are selected to sense in each slot, and prove that based on the advanced model a suitable number n will improve the performance of the channel significantly.

B. Basic Analyze of The Advanced Gilbert-Elliot Channel

Obviously if n equals 1, the advanced model will have the same performance as the basic model do and since the channel's performance will remain the same as it only related to the other users while has no relationship with the sensing period. Then the Transfer Matrix of the Markov chain will change according to different n. Assume that the when n=1, the transfer time from one state to another in the Markov chain is T, the Transfer Matrix $\begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}$ can be the index of the channel. Then when $n > 1$, which means transfer time changes from T to $\frac{T}{n}$, Transfer Matrix will change to $\begin{pmatrix} p_{00}^n & p_{01}^n \\ p_{10}^n & p_{11}^n \end{pmatrix}$ obviously the two matrix have the following relation:

$$\begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} = \begin{pmatrix} p_{00}^n & p_{01}^n \\ p_{10}^n & p_{11}^n \end{pmatrix}^n \quad (1)$$

We can see that

$$\begin{cases} p_{00} + p_{11} - 1 = (p_{00}^n + p_{11}^n - 1)^n, \\ \frac{1-p_{00}}{1-p_{11}} = \frac{1-p_{00}^n}{1-p_{11}^n}. \end{cases} \tag{2}$$

We may wish to set up n is odd so that

$$\begin{cases} p_{11}^n = \frac{p_{01}+(1-p_{11})\sqrt[n]{p_{11}-p_{01}}}{1+p_{01}-p_{11}}, \\ p_{01}^n = \frac{p_{01}(1-\sqrt[n]{p_{11}-p_{01}})}{1+p_{01}-p_{11}}. \end{cases} \tag{3}$$

The channel states are not directly observable before the sensing action is made. The user can, however, infer the channel states from its decision and observation history. Assume $\omega(t)$ is the conditional probability that the state of the channel is 1.Refer to as the belief vector or information state, the belief state in the time slot t+1 can be obtained recursively as follows:

$$\omega(t+1) = \begin{cases} p_{01}^n, & S(t) = 0 \\ p_{11}, & S(t) = n \\ \tau(\omega(t)), & not - sensed \end{cases} \tag{4}$$

where

$$\tau(\omega(t)) \triangleq \omega(t)p_{11}^n + (1-\omega(t))p_{01}^n \tag{5}$$

denotes the operator for the one-step belief update for unobserved channels.

If no information on the initial system state is available, the initial belief vector can be set to the stationary distribution ω_0 of the underlying Markov chain:

$$\omega_0 = \frac{p_{01}^n}{p_{10}^n + p_{01}^n} = \frac{p_{01}}{p_{01} + p_{10}} \tag{6}$$

We suppose to work on the model and to find the long-run reward based on the benefit from the data transmitted and the penalty from the sensing cost.

IV. Conclusion and Future Work

In this period, we get familiar with the basic concepts and knowledge of the cognitive radio and spectrum access, and we also learn some methods and models in this research field. And we study the Whittle Index, the Markovian Chain, KS model. Then we find some interesting point to do further researches. It is because that the time is not enough and we spend most our time reading papers and consider the researching direction, thus we do not finish the model construction. We consider the difference between sensing and transmitting time. We use a practical Markovian Chain structure to build our model. We successfully make some mathematic deduction to get the Markovian Chain matrix.

In the next period, we will focus on consummating the model, getting mathematic results from the model, doing computational simulations, and offering related proofs. Especially, we would like to propose limitations to both the largest number of channels K that can be sensed simultaneously for a single cognitive radio and the minimal subsidy m of keeping the channels idle which is a key factor in the utilization of Whittle's index. Hardware constraints will be considered when determining the largest number of channels K. Power consumption factors and elements of the Game Theory will be applied to the definition of the minimal subsidy m. Also, we will learn more about the Lagrangin multiplier and stochastic process so as to get the exact performance or the upper and lower bounds of this algorithm. Furthermore, we can decide the proper transmitting time versus sensing time ratio, and may get the best transmitting strategy getting the most benefits.

References

[1] Michael Maskery, Vikram Krishnamurthy and Qing Zhao *"Decentralized Dynamic Spectrum Access for Cognitive Radios: Cooperative Design of a Non-cooperative Game,"* New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN

[2] Qing Zhao and Brian M. Sadler], *"A Survey of Dynamic Spectrum Access,"* IEEE Signal Processing Magazine, 2007

[3] Bechir Hamdaoui, *"Adaptive Spectrum Assessment for Opportunistic Access in Cognitive Radio Networks,"* IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, VOL. 8, NO. 2, FEBRUARY 2009

[4] Joann'es Vermorel and Mehryar Mohri, *"Multi-Armed Bandit Algorithms and Empirical Evaluation,"* Courant Institute of Mathematical Sciences, 719 Broadway, New York, NY 10003, USA

[5] Ian F. Akyildiz *, Won-Yeol Lee, Kaushik R. Chowdhury, *"CRAHNs: Cognitive radio ad hoc networks,"* Volume 7, Issue 5, July 2009, Pages 810-836

[6] Zhu Ji and K. J. Ray Liu, University of Maryland, *"Dynamic Spectrum Sharing: A Game Theoretical Overview,"* Z Ji, KJR Liu - IEEE Communications Magazine, 2007 - cspl.umd.edu

[7] P. S. Ansell, K. D. Glazebrooky, J. Nina-Mora and M. O keeffe, *"Whittles index policy for a multi-class queueing system with convex holding costs,"* PS Ansell, KD Glazebrook, J Nino-Mora, MO'Keeffe - Mathematical Methods of Operations Research (ZOR), 2003 - Springer

[8] Gerard P. Cachon, *"GAME THEORY IN SUPPLY CHAIN ANALYSIS,"* GP Cachon, S Netessine - INTERNATIONAL SERIES IN OPERATIONS RESEARCH AND MANAGEMENT, 2004

[9] Keqin Liu, Qing Zhao, *"Indexability of Restless Bandit Problems and Optimality of Whittle's Index for Dynamic Multichannel Access"* J Nino-Mora - Advances in Applied Probability, 2001

[10] Yunxia Chen, Student Member, IEEE, Qing Zhao, Member, IEEE, and Ananthram Swami, Fellow, IEEE, *"Joint Design and Separation Principle for Opportunistic Spectrum Access in the Presence of Sensing Errors",* IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 54, NO. 5, MAY 2008

[11] Ian F. Akyildiz, Won-Yeol Lee, Mehmet C. Vuran *, Shantidev Mohanty, *"NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey"* Computer Networks 50 (2006) 2127C2159

[12] Zhi Quan, Student Member, IEEE, Shuguang Cui, Member, IEEE, Ali H. Sayed, Fellow, IEEE, and H. Vincent Poor, Fellow, IEEE, *"Optimal Multiband Joint Detection for Spectrum Sensing in Cognitive Radio Networks"* IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 57, NO. 3, MARCH 2009

[13] D. Bertsimas and J. Nino-Mora, *"Restless Bandits, Linear Programming Relaxations and a Primal-Dual Heuristic"* D Bertsimas, J Ni?o-Mora - Operations Research, 2000

[14] Qing Zhao and Brian M. Sadler, *"A Survey of Dynamic Spectrum Access"*Q Zhao, BM Sadler - IEEE Signal Processing Magazine, 2007

[15] Qing Zhao, Lang Tong, Ananthram Swami, and Yunxia Chen, *"Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework"* IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, VOL. 25, NO. 3, APRIL 2007

[16] Gerard P. Cachon, *"GAME THEORY IN SUPPLY CHAIN ANALYSIS"* GP Cachon, S Netessine - INTERNATIONAL SERIES IN OPERATIONS RESEARCH AND MANAGEMENT , 2004

[17] Senhua Huang, Xin Liu, and Zhi Ding, *"Opportunistic Spectrum Access in Cognitive Radio Networks"*

[18] Rahul Urgaonkar, Student Member, IEEE, Michael J. Neely, Member, IEEE, *"Opportunistic Scheduling with Reliability Guarantees in Cognitive Radio Networks"* R Urgaonkar, MJ Neely - ieeexplore.ieee.org

[19] Qing Zhao, Bhaskar Krishnamachari, Keqin Liu, *"On Myopic Sensing for Multi-Channel Opportunistic Access: Structure, Optimality, and Performance"* IEEE Transactions on Wireless Communications, 2007

[20] Keqin Liu, Qing Zhao, *"Indexability of Restless Bandit Problems and Optimality of Whittle's Index for Dynamic Multichannel Access"* IEEE Transactions on Information Theory , 2008