2020 54th Annual Conference on Information Sciences and Systems (CISS)

Learning Parallel Markov Chains over Unreliable Wireless Channels

Weichang Wang
School of Electrical, Computer and Energy Engineering
Arizona State University
Tempe, AZ, USA
wwang195@asu.edu

Lei Ying
Electrical Engineering and Computer Science
The University of Michigan, Ann Arbor
Ann Arbor, MI, USA
leiying@umich.edu

2020 54th Annual Conference on Information Sciences and Systems (CISS) 978-1-7281-4085-8/20/$31.00 ©2020 IEEE 10.1109/CISS48834.2020.1570614323

***Abstract*—This paper studies the problem of communications between aircraft and a control tower for aviation risk monitoring over wireless channels. The control tower needs to monitor the state of each aircraft in real time by receiving reports from the aircraft. Due to limited bandwidth, only a subset of aircraft can communicate with the control tower at the same time. This paper focuses on the problem of optimal scheduling of data transmissions to minimize the risk. We formulate the problem as learning states of parallel Markov chains where each Markov chain represents an aircraft, and the objective is to minimize the information entropy of all the aircraft. We propose an algorithm based on Whittle's index and study the indexability of the problem for both single-state wireless channels and multi-state wireless channels. Our numerical evaluations show that our algorithm improves the accuracy of the estimations compared with the heuristic scheduling methods such as greedy and Round& Robin.**

***Index Terms*—Whittle's Index, Restless Multi-Armed Bandit Problem, Multi-state channel**

I. Introduction

This paper considers the problem of monitoring parallel Markov chains over wireless networks. The problem is motivated by risk monitoring in aviation systems where a control tower needs to communicate with aircraft in its region to monitor their risk levels. The solution of this problem can also be applied to other risk monitoring applications. The major challenge in the problem is that the communication bandwidth is limited. For example, in aviation, automatic dependent surveillance – broadcast (ADS-B) is a current surveillance protocol in which each aircraft broadcasts its position periodically, enabling it to be tracked. The data bandwidth avaiable for ADS-B is about 1 Megabit/second. It has been shown in [1], the channel becomes very congested when multiple aircraft in an area broadcast their positions through the ADS-B channel, which lead to significant data loss.

In our problem, the central control needs to maintain an estimate of the risk level of each aircraft. When a report from an aircraft is successfully is received, the state of the Markov chain is know; otherwise, the estimate of the distribution of the risk level of an aircraft is updated based on a pre-defined Markov chain. In this paper, we assume simple two-state Markov chains. Due to limited bandwidth, the controller can only probe a subset of Markov chains each time. The objective is to develop a scheduling algorithm to minimize the total information entropy of the Markov chains.

This optimization problem is then formulated as a Multi-Armed Bandit (MAB) problem with the capacity of wireless channels as a hard constraint. The problem is similar to a restless bandit problem. The key difference is that the objective is to minimize the total information entropy of all bandits instead of finiding the optimal bandit. We adopt Whittle's Index to solve the problem. Whittle's Index was first proposed in [2] for restless bandit problems. Whittl's index has been used in wireless communication problems. For example, [3] consider sa delay minimization problem through a multi-state channels, and [4] studies the throughput maximization problem where transmitter in the system has dynamic multi-channel access.

In this paper, we consider both single-rate wireless channels and multi-rate wireless channels. We prove that the problem is indexable for single-rate wireless channels and establish a sufficient condition under which the problem is indexable with multi-rate wireless channels. Our numerical evaluations show that our algorithm outperform other heuristics such as the greedy policy and Round& Robin policies.

II. Problem Formulation

We consider a system consisting of M two-state Markov chains as shown in Fig.1. For simplicity, we assume that for the ith Markov chain, $p_{10} = p_{01} = p_i < 0.5$.

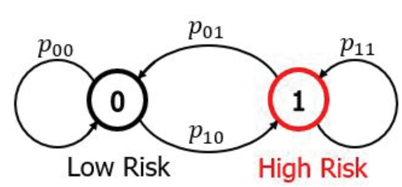

Fig. 1. A Two-State Markov Chain

We assume the controller can probe at most K $(K < M)$ of them at each time slot, and each probe succeeds with probability $r < 1$. Let $S_i(t)$ denote the state of the Markov chain i at time t, and $\theta_i(t) \in [0, 1]$ denotes the probability that

Markov chain i is in state "1" at time slot t given the most recent observation received at the controller.

$$\theta_i(t) = \begin{cases} S_i(t), & \text{if the probe is successful} \\ p_i + (1-2p_i)\theta_i(t-1), & \text{otherwise} \end{cases} \tag{1}$$

Given a Bernoulli distribution with parameter $\theta_i(t)$, the entropy of the distribution is

$$c_i(t) = -\theta_i(t)\log\theta_i(t) - (1-\theta_i(t))\log(1-\theta_i(t)).$$

The problem we are interested in is to minimize the overall entropy of the system, i.e.

$$\begin{aligned} &\min_{\pi\in\Pi} E\left[\sum_{t=0}^{\infty}\beta^t\sum_{i=1}^{M}c_i(t)\right] \\ &\text{subject to:} \sum_{i=1}^{M}A_i^{\pi}(t)\le K, \quad \forall t, \end{aligned} \tag{2}$$

where

$$A_i^{\pi}(t) = \mathbb{1}\{\text{Markov chain } i \text{ is probed at time } t\}$$

under a scheduling policy π, and Π is the set of all scheduling policies.

If we view $\theta_i(t)$ as the state of an arm, then the problem is related to restless bandit problems. A significant difference is that the reward $\sum_i c_i(t)$ depends on the states of all arms.

We use Whittle's index [2] to solve this problem. We will see that despite the fundamental difference in the cost function, the problem is an indexable problem. Following Whittle's index approach, we first relax the hard constraint per time slot to an average constraint, i.e., the number of Markov chains to be obsrved is at most K on average,

$$\sum_{t=0}^{\infty}\sum_{i=1}^{M}\beta^t A_i^{\pi}(t)\le\frac{K}{1-\beta}.$$

By introducing the Lagrange multiplier v to the problem, we have the following Lagrangian:

$$\mathcal{L}(v) = \min_{\pi\in\Pi} E\left[\sum_{t=0}^{\infty}\beta^t\sum_{i=1}^{M}c_i(t) + v\sum_{t=0}^{\infty}\beta^t\left(\sum_{i=1}^{M}A_i^{\pi}(t) - K\right)\right] \tag{3}$$

Note that the Lagrange multiplier v can be viewed as the penalty. Since the term $v\sum_{t=0}^{\infty}\beta^t K$ is a constant in the optimization problem, for a fixed v, the relaxed problem can be de-coupled into sub-problems associate with each individual Markov chain. In particular, we have

$$\min_{\pi\in\Pi} E\left[\sum_{t=0}^{\infty}\beta^t c_i(t) + v\sum_{t=0}^{\infty}\beta^t A_i^{\pi}(t)\right]. \tag{4}$$

Note that while we replace the hard constraint, the algorithm implemented can only probe K Markov chains. The Whittle index approach is to index the M Markov chains and then the algorithm picks the K ones with the highest indices.

The Whittle's Index Policy is a low-complexity heuristic that has been extensively used in the literature and performs well in practice.The challenge is that problems are not always indexable. In the following sections, we will prove the indexability and the conditioned indexability, i.e. Whittle's index is well defined.

III. Whittle's Index Approach

To solve the sub-problem Eq.(4) for each Markov chain. We consider the following Bellman equation:

$$\begin{aligned} V_i(\theta_i;v) = \min\Big\{ & c_i(\theta_i) + \beta V_i\big(p_i + (1-2p)\theta_i; v\big), v + \\ & c_i(\theta_i) + \beta\Big[r\theta_i V_i(1;v) + r(1-\theta_i)V_i(0;v) + \\ & (1-r)V_i\left(p_i + (1-2p_i)\theta_i; v\right)\Big]\Big\} \end{aligned} \tag{5}$$

where $V_i(\theta, v)$ is the value function of the ith Markov chain starting from state $\theta_i(t)$, and r is the message delivery ratio. The Whittle index in this problem is $v^*(\theta_i, r)$, the smallest value of v in the Eq. (4) that makes it equally desirable to observe and not to observe when the ith Markov chain is in state θ_i. The fundamental question in Whittle's index whether the problem is indexable. We will analyze the indexability in two different cases.

A. Single State Channels

We first consider the case the message delivery ratio r remains the same at all time, and have the following lemma.

Lemma 1. *$V_i(\theta_i; v)$ is a concave function in θ_i.*

The proof of this lemma can be found in the appendix [5].

Letting the two terms in the minimization equal to each other, we obtain

$$\begin{aligned} c_i(\theta_i(t)) + \beta V_i\big(p_i + (1-2p_i)\theta_i(t); v\big) = v + c_i(\theta_i(t)) \\ \beta r\theta_i(t)V_i(1;v) + \beta r(1-\theta_i(t))V_i(0;v) + \\ \beta(1-r)V_i\left(p + (1-2p_i)\theta_i(t); v\right). \end{aligned} \tag{6}$$

Since the cost entropy function is symmetric in θ, we know that $V_i(0;v) = V_i(1;v)$, which yields

$$\beta r V_i\big(p_i + (1-2p_i)\theta_i(t); v\big) = v + \beta r V_i(0;v). \tag{7}$$

We next show that the problem is indexable when r is given. Let $D_i(v)$ be the set of values of θ_i for which Markov chain i will not be probed under the v-penalty policy, i.e.

$$D_i(v) = \{\theta\in[0,1] : \beta r V_i\big(p_i + (1-2p_i)\theta; v\big) < v + \beta r V_i(0;v)\}.$$

The problem is indexable if $D_i(v)$ increases monotonically from $\varnothing$ to the universe set as v increasing from 0 to ∞, as established in the following theorem.

Theorem 1. *The Markov chains are indexable.*

Proof. The indexable condition is equivalent to that Equation (7) has a unique solution $v^*(\theta)$ for each state θ. Because of the symmetry of the cost function, we know that $V_i(0;v) = V_i(1;v)$. Based on the concavity of $V_i(\theta; v)$, we have

$$V_i\big(p_i + (1-2p_i)\theta_i; v\big) > V_i(0;v)$$

because for every $\theta \in (0,1)$, the point $(\theta, V_i(\theta; v))$ on the graph of $V_i(\theta; v)$ is above the straight line joining the points $(0, V_i(0; v))$ and $(1, V_i(1; v))$, as shown in Fig. 2. So when $v = 0$, $D_i(0) = \varnothing$.

On the other hand, we know that $V_i(\theta, v)$ is upper bounded by $\frac{-\log(0.5)}{1-\beta}$, which is the discounted total cost when the state stays as $\theta = 0.5$, which occurs if no probing occurs. As a result, when $v > \frac{-\log(0.5)}{1-\beta}$, $D_i(v) = \{\theta_i : \theta_i \in [0,1]\}$.

Next we prove the monotonicity of $D_i(v)$. We know that the LHS of Equation (7) is a concave function in θ_i. So the LHS and RHS can be plotted as in Fig. 2. From the symmetry

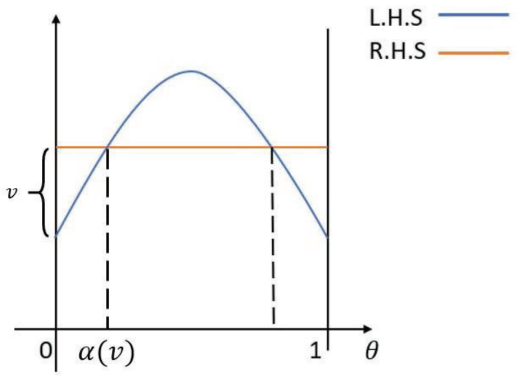

Fig. 2. LHS and RHS of Eq.(7)

of the cost function, we know that $V_i(\theta_i; v) = V_i(1-\theta_i; v)$, so in the remaining part of the proof, we assume that $\theta_i \in [0, 0.5]$ for simplicity.

According to Fig. 2, when $\theta_i \in [0, \alpha(v))$, the LHS is smaller than the RHS of Equation (7), $D_i(v) = \{[0, \alpha(v))\}$. Let

$$\begin{aligned} g_1(\theta_i(t), v) &= \beta r V_i\big(p_i + (1-2p_i)\theta_i(t); v\big) \\ g_2(\theta_i(t), v) &= v + \beta r V_i(0; v). \end{aligned}$$

We can prove the indexability by proving

$$\frac{\partial g_2}{\partial v} - \frac{\partial g_1}{\partial v} \geq 0$$

for any $\theta_i \in [0, \alpha(v))$. In other words, for fixed θ, as v increases, g_2 increases faster than g_1 if $\theta \in D(v)$. The condition can be written as:

$$1 + \beta r \frac{\partial V_i(0; v)}{\partial v} - \beta r \frac{\partial V_i(p_i + (1-2p_i)\theta_i; v)}{\partial v} \geq 0 \tag{8}$$

for any $\theta_i \in D_i(v)$.

Here we point out that $V(\theta; v)$ is not differentiable for some v. In particular, the function is not differentiable when

$$\beta r V_i\big(p_i + (1-2p_i)\theta; v\big) = v + \beta r V_i(0; v),$$

i.e. when θ is on the boundary of $D(v)$. When $V_i(\theta; v) = g_1(\theta, v) = g_2(\theta, v)$, probe or not does not make any difference, but the derivative of the two terms in the Bellman equation Equation (5) may be different. Since a boundary point is not included in $D(v)$ According to its definition, so we consider the derivative of the second term when it is not differentiable and defines it to be $\frac{\partial V(\theta; v)}{\partial v}$ here, because if Eq.(8) holds for all the differentiable points, it also holds for both left and right hand derivative at the non-differentiable points. For the non-differentiable point of $V(0; v)$, right hand derivative will be considered.

Let

$$h^0(\theta) = \theta,$$

and

$$h^t(\theta) = p_i + (1-2p_i)h^{t-1}(\theta)$$

for $t \geq 1$, represents the t step state transition without probe. For any $\theta \in D_i(v)$, let

$$k = \arg\max_k \{h^k(\theta) \in D_i(v)\}.$$

So for $\theta \in D_i(v)$ we have:

$$V_i(g^t(\theta); v) = \sum_{i=0}^{t} c_i(h^i(\theta)) + \beta^{t+1} V_i(h^{t+1}(\theta); v) \tag{9}$$

when $0 \leq t \leq k$. The costs $c_i(\theta_i)$ are independent with v, so we have:

$$\frac{\partial V_i(\theta; v)}{\partial v} = \beta^{k+1} \frac{\partial V_i(h^{k+1}(\theta); v)}{\partial v} \tag{10}$$

As a complement, we also point out that k is an integer related to v, $\frac{\partial V_i(h^{k+1}(\theta); v)}{\partial v}$ is not differentiable when $h^{k+1}(\theta)$ lies on the boundary of $D(v)$ for any $k \geq 0$. We will prove that the indexability holds for any $k \geq 0$, then both left-hand derivative and right-hand derivative are under consideration. So we simply regard k as a constant for all differentiable v.

Next we consider about the term $\frac{\partial V_i(h^{k+1}(\theta); v)}{\partial v}$, we have:

$$\begin{aligned} V_i(g^{k+t}(\theta); v) = v + c_i(g^{k+t}(\theta)) + \\ \beta r V_i(0; v) + \beta(1-r) V(g^{k+t+1}(\theta); v) \end{aligned} \tag{11}$$

for any $t \geq 1$. Let $\frac{\partial V(0; v)}{\partial v} = x$

$$\begin{aligned} \frac{\partial V_i(g^{k+1}(\theta); v)}{\partial v} &= 1 + \beta r x + \beta(1-r) \\ &\quad \Big\{1 + \beta r x + \beta(1-r)\Big(1 + \beta r x + \beta(1-r)\cdots\Big)\Big\} \\ &= \frac{1 + \beta r x}{1 - \beta(1-r)} \end{aligned}$$

so for $\theta \in D(v)$, we have:

$$\frac{\partial V_i(\theta; v)}{\partial v} = \beta^{k+1} \frac{1 + \beta r x}{1 - \beta(1-r)}$$

and

$$\frac{\partial V_i(p_i(1-2p_i)\theta_i(t); v)}{\partial v} = \beta^k \frac{1 + \beta r x}{1 - \beta(1-r)}$$

Then Eq.(8) becomes:

$$\begin{aligned} 1 + \beta r \left[x - \beta^k \frac{1 + \beta r x}{1 - (1-r)\beta}\right] &\geq 0 \\ x - \beta^k \frac{1 + \beta r x}{1 - (1-r)\beta} &\geq -\frac{1}{\beta r} \\ \left(1 - \frac{\beta^{k+1} r}{1 - \beta + \beta r}\right) x &\geq \frac{\beta^k}{1 - \beta + \beta r} - \frac{1}{\beta r} \end{aligned}$$

First it is easy to show that $\frac{\partial V(\theta;v)}{\partial v} > 0$, since as the penalty increase, the discounted total cost can not decrease, so we have $x \geq 0$, and the LHS of above equation is always positive for any $k \geq 0$.

Since $\frac{\beta^k}{1-\beta+\beta r} < \frac{1}{\beta r}$, the RHS of the above equation is always smaller than 0 for any $k \geq 0$. So Eq.(8) always holds when $\theta_i(t) \in D_i(v)$, and the problem is indexable.

As we mentioned before, $V_i(h^{k+1}(\theta), v)$ is not differentiable in v when $h^{k+1}(\theta)$ is on the boundary of $D(v)$, since k is a piece-wise constant on v. Both of left derivative and right derivative can be support by Eq.(8), since it holds for any $k \geq 0$. □

The value $v^*(\theta_i, r)$ is defined as the penalty on probing to balance the two terms in the Bellman equation Eq.(5). Whittle's index based policies are known to have good performance in practice, see [6] and [3].

We next summarize the calculation of Whittle's index. According to Eq.(9) and Eq.(11), θ is on the boundary of $D(v^*(\theta))$, in other words, $D(v^*(\theta)) = [0, \theta)$ for any $\theta > 0$. So we can get:

$$\begin{aligned}V_i(0, v^*(\theta)) &= \sum_{j=0}^{L_0} \beta^j c_i(h^j(0)) + \beta^{L_0}\Big(v+ \\ &\beta\Big[rV_i(0;v^*(\theta)) + (1-r)V_i\left(h^{L_0+2}(0);v^*(\theta)\right)\Big]\Big) \\ &= \sum_{j=0}^{L_0} \beta^j c_i(h^j(0)) + \beta^{L_0}\frac{v+\beta r V(0;v^*(\theta))}{1-\beta(1-r)} \\ &+ \sum_{j=1}^{\infty} \beta^{L_0+j}(1-r)^j c(h^{L_0+j}(0)) \end{aligned} \tag{12}$$

where

$$\begin{aligned}L_0 &= \arg\max_k \{h^k(0) \in D(v^*(\theta))\} \\ &= \lceil \log_{1-2p_i}(1-2\theta) \rceil.\end{aligned}$$

On the other hand, Eq.(7) holds for $v = v^*(\theta)$, we have

$$\beta r V_i\big(h^1(\theta); v^*(\theta)\big) = v + \beta r \cdot V_i(0; v^*(\theta))$$

$$\begin{aligned}\sum_{j=1}^{\infty} r\beta^j(1-r)^{j-1}c(h^j(\theta)) + \frac{\beta r v + \beta^2 r^2 V_i(0;v^*(\theta))}{1-\beta(1-r)} \\ = v + \beta r \cdot V_i(0; v^*(\theta))\end{aligned} \tag{13}$$

Combine Eq.(12) and Eq.(13), Whittle's index $v^*(\theta)$ of the ith aircraft at state θ can be solved.

B. Multi-State Channel

We now consider multi-state channel case, assume that the channel states of the ith Markov chain r_i is an i.i.d. random variable such that $r_i \in \mathcal{R}_i = \{r_{i,1}, r_{i,2}, \cdots, r_{i,n}\}$ with $r_{i,1} > r_{i,2} > \cdots > r_{i,n}$ for any i. Each channel state occurs with probabilities $\rho_{i,1}, \rho_{i,2}, \cdots, \rho_{i,n}$ respectively, and satisfying $\sum_j \rho_{i,j} = 1$ for any i. Also we assume that the channel states at current time is known for all Markov chains, but the future channel states are unknown. This setting is similar to the multi-state channel in [3].

In Multi-State channel, for the ith Markov chain, the tuple (θ_i, r_i) where $\theta_i \in [0,1]$, and $r_i \in \mathcal{R}_i$ consists the state, since decision depends on both θ_i and r_i, and the state space is $[0,1] \times \mathcal{R}_i$. Still, let $D_i(v)$ be the set of states where the ith Markov chain would not to be probed under v-penalty policy. The Bellman Equation (5) becomes:

$$\begin{aligned}V_i(\theta_i(t), r_i; v) = \min\Big\{& c_i(\theta_i(t)) + \beta\overline{V}_i\big(p_i + (1-2p_i)\theta_i(t); v\big), \\ & v + c_i(\theta_i(t)) + \beta r_i \overline{V}_i(0;v) + \beta(1-r_i)\cdot \\ & \overline{V}_i\left(p_i + (1-2p_i)\theta_i(t); v\right)\Big\}\end{aligned} \tag{14}$$

where $\overline{V}_i(\theta_i; v) = \mathbb{E}_{r_i}[V_i(\theta_i, r_i; v)]$ is the expected value over r_i, that is:

$$\begin{aligned}\overline{V}_i(\theta_i(t); v) = \mathbb{E}_{r_i}\Big[& \min\Big\{c_i(\theta_i(t)) + \beta\overline{V}_i\big(p_i + (1-2p_i)\theta_i(t); v\big), \\ & v + \beta r_i \overline{V}_i(0;v) + \beta(1-r_i)\overline{V}_i\left(p_i + (1-2p_i)\theta_i(t); v\right)\Big\}\Big]\end{aligned} \tag{15}$$

Similarly, because of the symmetric property of the Markov process, we only consider about $\theta \in [0, 0.5]$. From the previous section, we know that the concavity still holds. Let the two terms in the minimum are equal to each other, we have:

$$\beta r_i \overline{V}_i\big(p_i + (1-2p_i)\theta_i(t); v\big) = v + \beta r_i \overline{V}_i(0;v) \tag{16}$$

For agent i, the space of θ_i can be divided into $n+1$ parts $\{\Phi_{i,l}\}$ where $l = 0, 1, 2, \cdots, n$, and $\Phi_{i,0}$ satisfying:

$$\beta r_{i,j}\overline{V}_i\big(p_i + (1-2p_i)\theta_i; v\big) < v + \beta r_{i,j}\overline{V}_i(0;v) \text{ for all } r_{i,j} \in \mathcal{R}_i$$

$\Phi_{i,1}$ satisfies:

$$\begin{cases} \beta r_{i,1}\overline{V}_i\big(p_i + (1-2p_i)\theta_i; v\big) \geq v + \beta r_{i,1}\overline{V}_i(0;v) \\ \beta r_{i,j}\overline{V}_i\big(p_i + (1-2p_i)\theta_i; v\big) \\ \qquad < v + \beta r_{i,j}\overline{V}_i(0;v) \quad \text{for all } j > 1 \end{cases}$$

$\Phi_{i,l}$ satisfies:

$$\begin{cases} \beta r_{i,j}\overline{V}_i\big(p_i + (1-2p_i)\theta_i; v\big) \\ \qquad \geq v + \beta r_{i,j}\overline{V}_i(0;v) \quad \text{for all } j \leq l \\ \beta r_{i,j}\overline{V}_i\big(p_i + (1-2p_i)\theta_i; v\big) \\ \qquad < v + \beta r_{i,j}\overline{V}_i(0;v) \quad \text{for all } j > l \end{cases}$$

$\Phi_{i,n}$ satisfies:

$$\beta r_{i,j}\overline{V}_i\big(p_i(1-2p_i)\theta_i; v\big) \geq v + \beta r_{i,j}\overline{V}_i(0;v) \text{ for all } r_{i,j} \in \mathcal{R}_i$$

From the concave property, we have $\overline{V}_i\big(p_i(1-2p_i)\theta_i; v\big) > \overline{V}_i(0;v)$ for any θ_i, by moving the $\beta r_{i,j}\overline{V}(0,v)$ term to the left, it is easy to show that any $\theta_i \in \Phi_{i,l-1}$ is smaller than $\theta_i' \in \Phi_{i,l}$. Then the rested set $D_i(v)$ of the ith Markov chain can be described as $D_i(v) = \{(\theta_i, r_i) : \theta_i \in \Phi_{i,0}, r_i \in \mathcal{R}_i \text{ or } \theta_i \in \Phi_{i,l}, r_i < r_{i,l} \text{ for } 0 < l \leq n\}$.

Similarly, to prove the indexability of iid situation, we need to prove

$$1+\beta r_i\frac{\partial\overline{V}_i(0;v)}{\partial v}-\beta r_i\frac{\partial\overline{V}_i(p_i+(1-2p_i)\theta_i;v)}{\partial v}\geq 0 \quad (17)$$

holds for any $(\theta_i, r_i)\in D_i(v)$ for any i.

Theorem 2. *The mult-state channel Markov chains are indexable when*

$$\beta<\frac{1}{1+(1-\rho_{i,1})r_{i,1}}, \quad (18)$$

holds for all i.

The proof of this Theorem can be found in the appendix. [5]

Similarly, the we would choose to probe the K Markov chains with higher indexes. However the indexes now depend on both state estimation $\theta_i(t)$ and channel state r_i. In multi-state channel, the explicit format of $v_i^*(\theta, r_i)$ is hard to solve, especially when the number of state of channel is large. However, we can use binary search to get an approximation. As for an example, we will show brief process to derive the two-state channel index as an example.

For the ith Markov chain, to solve the index $v^*(\theta, r_{i,1})$, let $v^*=v^*(\theta, r_{i,1})=v^*(\theta', r_{i,2})$ $(0<\theta<\theta'<0.5)$, and θ' is temporarily unknown. From the proof above, let $\overline{V}_i(0, v^*)=x$, $(\theta, r_{i,1})$ and $(\theta', r_{i,2})$ is on the boundary of $D_i(v^*)$, we have the following equations:

$$\beta r_{i,1}\overline{V}_i(p_i+(1-2p_i)\theta;v^*)=v^*+\beta r_{i,1}x \quad (19)$$

$$\beta r_{i,2}\overline{V}_i(p_i+(1-2p_i)\theta';v^*)=v^*+\beta r_{i,2}x \quad (20)$$

On the other hand, $x=\overline{V}_i(0,v^*)$ can be expressed as:

$$\begin{aligned}x=&\sum_{j=0}^{L_0}\beta^j c_i(h_i^j(0))+\sum_{j=0}^{L_1-1}\beta^{L_0+1+j}(1-\rho_1 r_1)^j\\&\left(c_i(h_i^{L_0+1+j}(0))+\rho_{i,1}v^*+\beta\rho_{i,1}r_{i,1}x\right)\mathbb{1}\{L_1>0\}+\\&\sum_{j=0}^{\infty}\beta^{L_0+L_1+1+j}(1-\rho_{i,1}r_{i,1})^{L_1}(1-\overline{r}_i)^j\\&\left(c_i(h_i^{L_0+L_1+1+j}(0))+v^*+\beta\overline{r}_i x\right) \quad (21)\end{aligned}$$

where $L_0=\max_k\{h_i^k(0)<\theta\}$, and $L_1=\max_k\{h^k(0)<\theta'\}-k_0$. Combine Eq.(19)(20)(21) the index value of v^*, θ', $\overline{V}_i(0,v^*)$ for the ith Markov chain can be estimated by using binary search.

IV. Simulations

We consider a scenario where a control tower is monitoring aircrafts in the area. Each aircraft has two states: “low risk” and “high risk”. The transition probability p from one state to another is assumed to be 0.05.

We assume that there are 500 aircrafts in the system, based on the channel bandwidth, the control tower can require information from 150 of them at each time slot. Assume that there are two types of aircraft, each types has 250 aircrafts. The first one has transition probability $p_1=0.05$, and has transmission success probability $r_1=0.5$. The second type of aircraft has transition probability $p_2=0.02$ and transmission success probability $r_2=0.7$.

We can plot the index of these two types of aircraft as in Fig.3. We compare the Whittle's index approach with a greedy

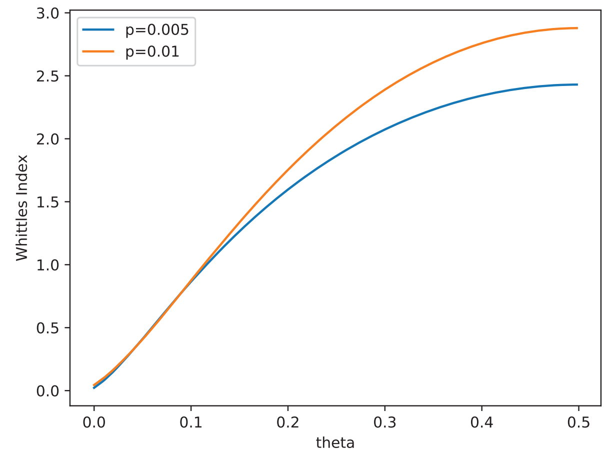

Fig. 3. Index of two types of aircraft

method that aircraft with larger $c(\theta)$ will be selected, and the Round Robin method, where all the aircrafts are selected periodically with same frequency. And the simulation results of total information entropy of all these 500 aircrafts are shown in Fig.4. As we can see, the information entropy level of Whittle's Index Approach is lower than both the greedy or Round & Robin methods.

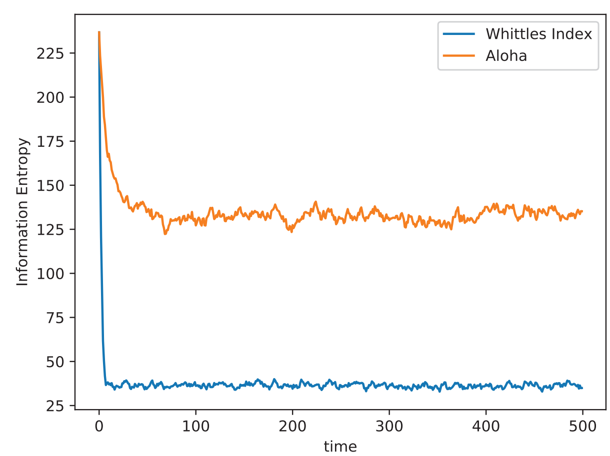

Fig. 4. Information Entropy Simulation

We next consider multi-state channels such that $\Pr(r=0.9)=0.4$, $\Pr(r=0.7)=0.3$, and $\Pr(r=0.5)=0.3$.

The simulation results are plotted in Fig.5. Again we can observe that the Whittle's index outperforms the other two algorithms.

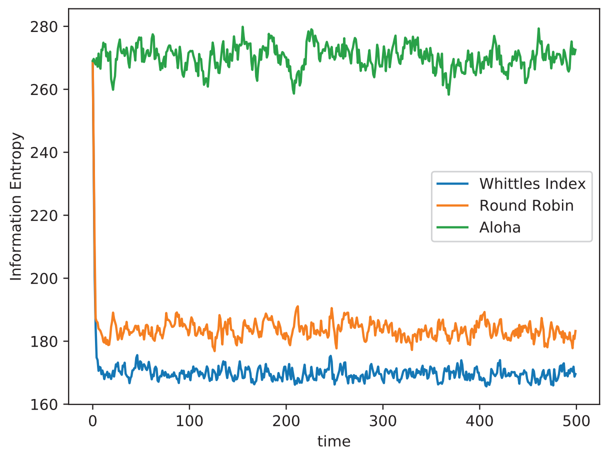

Fig. 5. Index of three states channel aircraft

V. Conclusion

In the paper, we studied the problem of learning the states of parallel Markov chains over unreliable wireless networks. The solution to this problem has applications in risk monitoring such as monitoring the states of aircrafts (or UAVs) from a control tower. We first proved that for single state wireless channels, the problem based on Whittle's index is indexable, and the index can be derived explicitly. For multi-state channels, the indexability can be proved under a sufficient condition, and the index value can be calculated numerically. And simulation shows that the proposed Whittle's index approach can have better performance than greedy policy and Round & Robin policy.

Acknowledgement

Research supported in part by NSF ECCS 1609202, ECCS 1739344, CNS 2002608 and CNS 2001687, and NASA University Leadership Initiative program (Contract No. NNX17AJ86A)

References

[1] G. Galati, E. G. Piracci, N. Petrochilos, and F. Fiori. 1090 mhz channel capacity improvement in the air traffic control context. In *2008 Tyrrhenian International Workshop on Digital Communications - Enhanced Surveillance of Aircraft and Vehicles*, pages 1–5, Sep. 2008.

[2] P. Whittle. Restless bandits: activity allocation in a changing world. *Journal of Applied Probability*, 25(A):287–298, 1988.

[3] Arjun Anand and Gustavo de Veciana. A whittle's index based approach for qoe optimization in wireless networks. *Proc. ACM Meas. Anal. Comput. Syst.*, 2(1):15:1–15:39, April 2018.

[4] K. Liu and Q. Zhao. Indexability of restless bandit problems and optimality of whittle index for dynamic multichannel access. *IEEE Transactions on Information Theory*, 56(11):5547–5567, Nov 2010.

[5] W. Wang and L. Ying. Learning parallel markov chains over unreliable wireless channels. https://www.dropbox.com/s/arcpcxdy3ri7pvg/Learning%20Parallel%20Markov%20Chains%20over%20UnreliableWireless%20Channels.pdf?dl=0.

[6] Samuli Aalto, Pasi Lassila, and Prajwal Osti. Whittle index approach to size-aware scheduling for time-varying channels with multiple states. *Queueing Systems*, 83(3):195–225, Aug 2016.