OWLS: Opportunistic Wireless Link Scheduling with SINR Constraints

Xiaohua Xu[1(✉)], Yuanfang Chen[2], Shuibing He[3], and Patrick Otoo Bobbie[1]

[1] Department of Computer Science, Kennesaw State University, Kennesaw, USA
xxu6@kennesaw.edu

[2] School of Cyberspace, Hangzhou Dianzi University, Hangzhou, China

[3] College of Computer Science and Technology, Zhejiang University, Hangzhou, China

Abstract. We study a classical opportunistic wireless link scheduling problem in cognitive radio networks with Signal to Interference plus Noise Ratio (SINR) constraints. Consider a collection of communication links, assume that each link has a channel state. The state transitions follow a transition rule. The exact state information of each link is not available due to the uncertainty of primary users' activities. The expected channel state is predicted probabilistically by investigating its history and feedbacks when the channels are used. The objective is to pick communication links sequentially over a long time horizon to maximize the average reward. To the best of our knowledge, no prior work can satisfyingly provide solutions for the opportunistic wireless link scheduling problem when considering SINR constraints. In this work, we adopt the robust paradigm of restless multi-armed bandit for the problem and design an efficient algorithm. We analyze the performance via Lyapunov potential function and demonstrate that the proposed algorithm can achieve an approximation bound.

Keywords: Restless multi-armed bandit · Opportunistic wireless link scheduling · Signal to Interference Plus Noise Ratio

1 Introduction

The pressing issues such as spectrum scarcity and low spectrum utilization have motivated the study on Cognitive radio networks (CRNs) [17]. In CRNs, a cognitive radio can be programmed and configured sophisticatedly to explore the available spectrum opportunities. A CRN consists of two types of users, *i.e.*, primary users and secondary users. Primary users such as big corporations have guaranteed quality of service while secondary users must accept interference from primary users. Secondary users can detect available channels to allow more transmissions. A secondary user can transmit successfully only when its transmission does not conflict with any active primary user and other active secondary users.

E. S. Biagioni et al. (Eds.): WASA 2019, LNCS 11604, pp. 419–431, 2019.
https://doi.org/10.1007/978-3-030-23597-0_34

In cognitive radio networks, Opportunistic Wireless Link Scheduling problem (noted as **OWLS**) is a well-formulated problem in the literature [14]. In the problem **OWLS**, we assume there is a collection of communication links, each of them has a demand to transmit. For simplicity, we assume all nodes lie in a two-dimensional plane. Every communication link has a channel state to indicate whether the link is occupied by a primary user or not. In a discrete version, the channel state can be either occupied (busy) or idle (free). The channel state of a link will change according to some rules. The evolving process belongs to a Markov chain. Each link's channel state evolves independently. The problem **OWLS** aim to pick communication links sequentially over a long time horizon to achieve a maximum expected reward.

In a CRN, for each link, the true channel state is unknown unless we schedule it to transmit. This is due to unpredictable primary user activities. Thus, we have to predict the channel states by investigating their history and feedbacks when the channels are actually used. Let us use the channel state of a communication link to indicate the probability that the channel is free for that link. Thus, the channel state of a link directly reflects the throughput along the link or the link quality. We use a reward proportional to a link's channel state to denote the expected reward when the link is selected to transmit. In this work, we assume multiple communication links may transmit simultaneously subject to wireless constraints. Along with channel state predictions, we need to optimize the scheduling decision policy for the problem **OWLS**.

The problem **OWLS** has been studied extensively. However, most of the existing work simply assumes the protocol interference model [10] or an even simpler model. The protocol interference model can only capture the binary interference. In other words, if one communication link is transmitting, does another link conflict with this link? Generally, we can use a binary conflict graph to model the protocol interference model. However, the physical interference model [28,29] is well-accepted as it captures wireless interference more accurately. Under this model, the interference cause a cumulative effect on data transmissions. Thus, we have to take care of multiple signal to interference and noise ratio (SINR) constraints.

In this work, we consider the problem **OWLS** with SINR constraints. As observed in [7] and subsequent works, there is a gap between these two interference models. Generally, an algorithm under the protocol interference model does not imply any similar solution for another model. Even worse, the difficulty of scheduling problems may suddenly increase when considering SINR constraints. For example, maximum weighted independent set (MWIS) [19] is well-known and fundamental wireless scheduling problem. Such a problem under the protocol interference model was addressed extensively, and now one focus has been designing simple and fast algorithms [19] instead of deriving improving the approximation bounds. However, whether the counterpart under the physical interference model admits a constant approximation is still open. The only positive results under the physical interference model exist in some restricted settings of the MWIS problem such as monotone power control or linear power control settings.

We address the problem **OWLS** subject the SINR constraints. We consider the problem **OWLS** with the paradigm of Restless Multi-Armed Bandit (abbreviated to Restless MAB, Restless Bandit, or RMAB in the literature) [6]. The Restless MAB paradigm models sequential constrained resource allocations among several competing arms. The traditional MAB [6] focuses on the conflict of allocating constrained resources between arms that yield high current rewards and ones with better future rewards. However, we consider restless MAB which means that the state evolution of unprocessed arms. In other words, the arms selected are allowed to changes its state, at the same time, the rest of the arms are also allowed to change states instead of remaining frozen. Corresponding to the problem **OWLS** setting, this means that passive links (links that did not transmit last time) are also allowed to change channel states. Restless MAB theory has been applied to the problem **OWLS** in single-hop cognitive radio networks [14]. However, the existing work in the literature on Restless MAB does not consider the SINR constraints.

In this work, we design an efficient scheduling algorithm with an approximation guarantee for the opportunistic wireless link scheduling problem with SINR constraints. The proposed algorithm is low-complexity by using the *divide and conquer* paradigm. This paradigm only applies to the condition that the length of each link is at most $(1-\epsilon)$ fraction of the maximum link length where ϵ is a arbitrarily small constant. We first divide the whole plane into small cells and use a double partition and coloring scheme to ensure the links scheduled each time are well seperated. We then solve the problem **OWLS** in each cell and combine the individual solutions to obtain a feasible transmission schedule for the whole plane. The approximation ratio of this algorithm is bounded under an arbitrary setting. Under the length-monotone power control setting, the approximation bound is a constant. For each small cell, based on a semi-infinite programming and Lyapunov analysis which are challenging for the opportunistic wireless link scheduling problem even without interference constraint. we show the algorithm can achieve a solution with low time complexity.

The remaining of the paper is organized as follows. Section 2 formulates the problem **OWLS** under the physical interference model. Section 3 presents an approximation algorithm under the physical interference model. Section 4 details the performance analysis. Section 5 presents the related work on Restless MAB and cognitive radio scheduling. Section 6 outlines the conclusion and the future work.

2 Network Model

Given a collection $L = \{\ell_1, \ell_2, \cdots, \ell_N\}$ of communication links. Each communication link consists of a sender node and a receiver node. The Euclidean distance between two nodes u and v is denoted by $\|\overrightarrow{uv}\|$. Assume all nodes lie in a two-dimensional Euclidean plane.

Table 1. Notations

$\overrightarrow{uv}$	A link from node u to node v
$\Vert\overrightarrow{uv}\Vert$	Link length, Euclidean distance, of link $\overrightarrow{uv}$
L	Set of all links
P	Maximum power for a node
η	Reference path loss factor
κ	Path loss exponent
ξ	Noise power
σ	SINR threshold for successful reception

Under the physical interference model [28,29], each node u has a transmission power P_u upper-bounded by P. When a signal is sent from u to v over a Euclidean distance ℓ, there is a path loss of $\eta l^{-\kappa}$. Here η is called the *reference loss factor* and κ is called the *path loss exponent*. Let ξ be the background noise power. Let σ be the SINR threshold for successful reception. Let S_u be the set of other simultaneous senders with sender u. A transmission $\overrightarrow{uv}$ is able to occur successfully if and only if the following constraint (*i.e.*, Eq. (1)) is satisfied.

$$\frac{P_u \cdot \eta\ell^{-\kappa}}{\xi + \sum_{w\in S_u} P_w \cdot \Vert\overrightarrow{wv}\Vert^{-\kappa}} \geq \sigma. \tag{1}$$

Table 1 lists major notations of the physical interference model.

Given a collection of links L, a subset I of links in L is said to be interference free if and only if all links in I transmit without interference under the physical interference model.

Suppose each communication link has a channel state for transmitting along that channel. The channel state space for all links is {idle or good, busy or bad}. The channel state of each link will evolve independently across time. The states are evolving according to the probability transition matrix in Table 2.

In Table 2, α_i and β_i are defined in the Gilbert-Elliot channel model [5]. Specifically, α_i and β_i are associated with link i. α_i denotes the probability for link i to transit from busy state to idle state. β_i denotes the probability to transit from idle state to busy state. For simplicity, we only consider that for each link, the probability that each channel state remains the same is greater than the probability that the channel state jump to the other one. In this scenario, the state evolution is called positive correlated [14]. The scenario of a positive auto-correlation occurs when $\alpha_i + \beta_i < 1$ for each i.

Table 2. The state transition of link i is shown by a 2×2 probability transition matrix.

Begin	End	
	Idle	Busy
Idle	$1-\beta_i$	β_i
Busy	α_i	$1-\alpha_i$

Let r_i be the expected reward of the channel for a link i. We assume there is a communication request along the link all the time. Suppose a communication link transmits without any conflict along a channel which is in an idle state, a reward of r_i is obtained.

The complete state information of each channel is not known due to unpredictable activities of primary users. To address this difficulty, we estimate the channel state by researching the history and feedback from the activated links. Assume time is divided equally into time-slots. The scheduler maintains a parameter $\pi_i[T]$ for each communication link. $\pi_i[T]$ means the probability that link i's channel state is idle at time-slot T.

At each time-slot, we select a collection of links to sense and transmit. The accurate channel state is revealed via *Acknowledgement/Negative Acknowledgement (ACK/NACK)* feedbacks. Here ACK/NACK messages are generated only after the data are transmitted according to the *Automatic Repeat Request (ARQ)* mechanism. The value of $\pi_i[T]$ represents the expected reward collected if the link i is scheduled in time-slot T.

For simplicity, we assume that the the time horizon starts from slot $T = 0$. At the time-slot 0, we assume an arbitrary initial channel state for each communication link.

We define a scheduling algorithm as selecting a collection of links to transmit at each time-slot. For an algorithm, let $a_i[T]$ denote whether link i is scheduled ($a_i[T] = 1$) or not ($a_i[T] = 0$) at time-slot T. The problem **OWLS** aims to design an algorithm that maximizes the average reward over the long time horizon, *i.e.*,

$$\lim_{\mathcal{T}\to+\infty} E\Big\{\frac{\sum_{T=0}^{\mathcal{T}}\sum_{i=0}^{N} a_i[T]\cdot\pi_i[T]\cdot r_i}{\mathcal{T}}\Big\}.$$

The system is subject to SINR constraints such that the selected communication links at each time-slot do not cause any conflict.

We introduce a set of variables as defined in [9]. For any link i, let $u_i(t)$ be the probability that the channel remains in an idle state after t time-slots; let $v_i(t)$ be the probability that the channel moves to an idle state from a busy state after t time-slots. Both $u_i(t)$ and $v_i(t)$ have monotone properties. In other words, as t increases, $u_i(t)$ strictly decreases and $v_i(t)$ strictly increases.

For any algorithm, we define a performance measure $\{I_i(t), O_i(t)\}$ of link $i, i = 1, 2, \cdots, N$ as follows. $I_i(t)$ denotes the probability that the last time of selecting the link i is t time-slots ago, its last observed state is idle state, and link i is selected in the current slot. Note that the three events are independent from each other. Similar we define $O_i(t)$ denotes the probability that the last time of selecting the link i is t time-slots ago, its last observed state is busy state, and link i is selected in the current slot. Here, $I_i(t), O_i(t)$ take expected values and can capture the time distribution of different channel states of each link when executing an algorithm.

3 Algorithm Design

3.1 Seperating Links via Partition and Coloring

Under the condition that each link has a length at most $(1-\epsilon)\mathbf{r}$ where ϵ is a small constant, we design a fast and simple algorithm for the problem **OWLS** in the following. Our main idea is inspired by the fact that if transmissions are well-separated geographically, the interference on each link is bounded. After observing this geometric property, we schedule communication links such that at each time step, we use the well seperated links distributed in the plane.

From Eq. (1), if no other link is transmitting concurrently, *i.e.*, S_u is empty, the SINR is reduced to *signal to noise ratio* (SNR): $\frac{P_u\cdot(\eta\cdot\|uv\|^{-\kappa})}{\xi} \geq \sigma$, thus, we have $\|uv\| \leq \sqrt[\kappa]{\frac{\eta P_u}{\sigma\xi}} \leq \sqrt[\kappa]{\frac{\eta P}{\sigma\xi}}$. Let $\mathbf{r} = \sqrt[\kappa]{\frac{\eta P}{\sigma\xi}}$ be the *maximum transmission radius.*

The plane is partitioned into a grid, where each cell has a diagonal length of $\mathbf{r}$. Let $\gamma = \mathbf{r}/\sqrt{2}$ denote the side length of a cell. Let $\mathbb{Z}$ represent the integer set. We use the vertical line segments $\{x = i\cdot\gamma : i \in \mathbb{Z}\}$ and horizontal line segments $\{y = j\cdot\gamma : j \in \mathbb{Z}\}$ to partition the plane. After the partition, the plane consists of half-open, half-closed grids of side length γ: $[i\gamma,(i+1)\gamma) \times [j\gamma,(j+1)\gamma) : i,j \in \mathbb{Z}$. We call the plane area $[i\gamma,(i+1)\gamma) \times [j\gamma,(j+1)\gamma)$ as cell $g_{i,j}$. In other words, the plane is partitioned into square cells $g_{i,j}$ for $i,j \in \mathbb{Z}$, with side length γ. After this partition of the plane, we define a large-block as a square which consists of $\big(K+1\big) \times \big(K+1\big)$ cells. The value of K is set as follows to based on the parameters of the physical interference model.

$$K = \lceil\sqrt{2}\left((4\tau)^{-1}(\sigma^{-1} - \xi(\eta)^{-1}R^{\kappa-\beta})\right)^{-1/\kappa} + \sqrt{2}\rceil \tag{2}$$

For each cell, we first derive an upper bound on the number of links that transmit concurrently without interference. Based on the parameters of the SINR constraints, this upper bound is $\omega = \lceil\frac{2^\kappa P}{\sigma^2\xi} + 1\rceil$.

3.2 Decouple Links in a Cell via Lagrange Relaxation

We then consider the problem **OWLS** in a cell c. We formulate the restricted version of the problem **OWLS** as a Whittle's Linear Programming (LP) [25]. Consider one cell c, let L' be the collections of communication links in this single cell. We have the following LP.

$$\begin{aligned}
&\max \sum_{i\in L'}\sum_{t\geq 1} r_i I_i(t)\\
&s.t. \ \sum_{i\in L'}\sum_{t\geq 1}\big(I_i(t)+O_i(t)\big)\leq \omega\\
&\sum_{t\geq 1} t\big(I_i(t)+O_i(t)\big)\leq 1, \forall i\\
&\sum_{t\geq 1} O_i(t)v_i(t)=\sum_{t\geq 1} I_i(t)(1-u_i(t)), \forall i\\
&I_i(t), O_i(t)\geq 0, \forall i, t\geq 1
\end{aligned} \tag{3}$$

The maximum value of the above LP, `OPT`, is greater than or equal to the maximum value for the problem **OWLS** in the cell c under the physical interference model.

For any $\lambda_c \geq 0$, the Lagrange relaxation of the LP in Eq. (3) is as follows.

$$\begin{aligned}
&\max \sum_{i\in L'}\sum_{t\geq 1} r_i I_i(t)\\
&\quad + \lambda_c\Big(\omega - \sum_{i\in L'}\sum_{t\geq 1}\big(I_i(t)+O_i(t)\big)\Big)\\
&= \max\ \omega\lambda_c + \sum_{i\in L'}\sum_{t\geq 1}\Big(r_i I_i(t) - \lambda_c(I_i(t)+O_i(t))\Big)
\end{aligned} \tag{4}$$

In the Lagrange relaxation, the links are decoupled. Thus, we decouple an $|L'|$-dimensional LP to a collection of one-dimensional LPs. Moreover, these one-dimensional LP are not interleaved and can be solved independently.

3.3 Semi-infinite Programming of Each Link Independently

For each link in the cell c, we identify the Semi-Infinite Program (SIP) as follows.

$$\begin{aligned}
\mathtt{SIP}_i^{\lambda_c}:\quad &\max \sum_{t\geq 1}\Big(r_i I_i(t) - \lambda_c(I_i(t)+O_i(t))\Big)\\
&s.t. \ \sum_{t\geq 1} t\big(I_i(t)+O_i(t)\big)\leq 1\\
&\sum_{t\geq 1} O_i(t)v_i(t)=\sum_{t\geq 1} I_i(t)(1-u_i(t))\\
&I_i(t), O_i(t)\geq 0, t\geq 1.
\end{aligned} \tag{5}$$

For $\mathtt{SIP}_i^{\lambda_c}$ with a Lagrange multiplier λ_c, the maximum value of $\mathtt{SIP}_i^{\lambda_c}$ is

$$H_i^{\lambda_c} = \frac{(r_i-\lambda_c)v_i(t_i)-\lambda_c\beta_i}{v_i(t_i)+t_i\beta_i}, \quad t_i = \arg\max_t \frac{(r_i-\lambda_c)v_i(t)-\lambda_c\beta_i}{v_i(t)+t\beta_i}.$$

Since we have $\forall\lambda_c \geq 0\ \omega\lambda_c + \sum_i H_i^{\lambda_c} \geq OPT$, and $H_i^{\lambda_c}$ decreases with λ_c, there is a λ_c such that $\lambda_c = \sum_i H_i^{\lambda_c} \geq \frac{OPT}{\omega+1}$.

Algorithm 1. Divide and Conquer Algorithm for the problem **OWLS** with SINR Constraints

Input : Set of links L.
Compute the value of K under the physical interference model;
Implement the cell partition and coloring scheme;
for *each cell* **do**
  Find the links restricted to the cell and formulate the problem as Whittle's LP;
  Calculate the Lagrange relaxation of Whittle's LP;
  Solve the decoupled SIP for each link in the cell;
  Let $H_i^{\lambda_c} \leq 0$ be the maximum value for link i and let t_i be the corresponding t such that $H_i^{\lambda_c}$ is achieved;
  Remove any link i with $H_i^{\lambda_c} \leq 0$;
For each color k, let R_k denote the summation of rewards of all cells with color index k;
We return the largest reward $R_{k'}$ with a color k';
$t \leftarrow 0$;
while *TRUE* **do**
  $t \longrightarrow t + 1$;
  for $i, j \in \mathbb{Z}$ *and the cell* $g_{i,j}$ *contains links from* L *;* **do**
    if *the cell* $g_{i,j}$ *has a color* k' **then**
      Find the links whose sender lies within $g_{i,j}$, assume they form $L_{i,j}$;
      We select a link from $L_{i,j}$ based on the following priority: if there is a link last observed in idle state, then we select it for this time-slot;
      Else, we select a link last observed in busy state $t \geq t_i$ time-slots ago;
      If we cannot find any link, we skip the selection for this cell.
    All the selected links among all cells form a set S_t;
if $S_t \neq \emptyset$ **then**
  Transmit all links in S_t at time-slot t;
return $S_1, S_2, \cdots, S_t : t \to \infty$.

3.4 Overall Algorithm

The proposed algorithm is described in Algorithm 1. At each time, a link with (g, t) is called *idle ready*. A link with $(b, t), t \geq t_i$ is called *busy ready*. Here, R_k is the summation of rewards of all cells with color index k. At the initial time-slot, we simply assume that the initial states are all busy for all links. This assumption will not modify the long term's average reward.

4 Performance Analysis

We first prove the correctness of Algorithm 1. Then, we compute its approximation ratio. The analysis is based on a carefully defined Lyapunov potential function.

Theorem 1. *Algorithm 1 outputs an independent set of links at each time-slot.*

Proof. For all $i, j \in \mathbb{Z}$, let L_{ij} be the set of links in L whose senders lie in cell $g_{i,j}$. For each large block, I contains up to one link from the cell of the same color. As each link set S_T contains well-separated links, the link set is an independent set according to the selection of K. This proof is similar to Theorem 1 of [26].

Next, we derive the approximation bound of Algorithm 1. The main idea of the proof is that at each time-slot, the scenario of collecting rewards may be different. We employ a potential function to derive the total reward. The average reward is derived by dividing the total reward by the total time used.

Theorem 2. *Algorithm 1 achieves a $(K+1)^2 \cdot (\omega+1)$ approximation bound for the problem* ***OWLS*** *subject to SINR constraints.*

Proof. For a cell of the color k', assume L' consists of links in the cell. We first define a potential function for each link $i \in L'$. At time-slot T, if link i moved to busy channel state y time-slots ago, the potential $\phi_i^T = H_i^{\lambda_c}\big(\min(y, t_i) - 1\big)$. If link i was last observed at idle channel state, the potential $\phi_i^T = \frac{\lambda_c + t_i H_i^{\lambda_c}}{v_i(t)}$. When a link is observed in a busy state, after that, the link will become blocked if the link has not been selected to transmit for more than t_i time-slots.

Let r_i^T denote the expected reward (throughput) accrued from link i until time T.

$$\Delta\phi_i^T = \phi_i^{T+1} - \phi_i^T, \quad \Delta r_i^T = r_i^{T+1} - r_i^T.$$

Next, we consider a link i at any time-slot $T+1$. Based on whether the link is blocked or not, we have the following cases:

If the link i has been observed in a busy channel state for consecutive $t < t_i$ time-slots, the potential function is increased by $H_i^{\lambda_c}$ at this time-slot, the reward obtained by this link is zero, and we have

$$E[\Delta r_i^T + \Delta\phi_i^T] = H_i^{\lambda_c}$$

If the link i is selected and the link is last observed in busy channel state, assume the link is last observed y time-slots ago, then $y \geq t_i$. Since $v_i(t)$ is monotonically increasing with t, with probability $q \geq v_i(t_i)$, the observed channel state is idle and the reward $\Delta r_i^T = r_i$ and the change of potential is $\frac{\lambda_c + t_i H_i^{\lambda_c}}{v_i(t)} - H_i^{\lambda_c}(t_i - 1)$. With probability $1-q$ the observed channel state is busy and the change of potential is $H_i^{\lambda_c}(t_i-1)$ and the reward remains the same. Since we have $q \geq v_i(t_i)$ and $\frac{\lambda_c + t_i H_i^{\lambda_c}}{v_i(t)} \geq 0$, we have

$$E[\Delta r_i^T + \Delta\phi_i^T] = q\left(r_i + \frac{\lambda_c + t_i H_i^{\lambda_c}}{v_i(t)}\right) - H_i^{\lambda_c}(t_i - 1) \geq \lambda_c + H_i^{\lambda_c}$$

If the link i's last observed state is idle, then the link is selected at last time-slot. There are two cases. The first case is that the reward is increased by r_i

and the potential remains the same. The probability of the first case is $1 - \beta_i$. The second case is that the potential function is decreased by $\frac{\lambda_c + t_i H_i^{\lambda_c}}{v_i(t)}$. The probability of the second case is β_i. We have

$$E[\Delta r_i^T + \Delta \phi_i^T] = (1 - \beta_i) r_i - \beta_i \frac{\lambda_c + t_i H_i^{\lambda_c}}{v_i(t)} \geq \lambda_c + H_i^{\lambda_c}$$

If link i is blocked, then the reward from this link is zero, and the potential function of the link is not changed, *i.e.*, $\phi_i^T = \phi_i^{T+1} = H_i^{\lambda_c}(t_i - 1)$. Then we have $E[\Delta r_i^T + \Delta \phi_i^T] = 0$.

For all of the four cases, we have obtained the expected reward of a link. Next we consider the links in the cell together at time T. Let $\Phi_T = \sum_{i \in L'} \phi_i^T$.

We use $R_T = \sum_{i \in L'} r_i^T$ to denote the total reward collected from the begining to the time-slot T.

$$\Delta \Phi_T = \Phi_{T+1} - \Phi_T = \sum_{i \in L'} \Delta \phi_i^T, \quad \Delta R_T = R_T - R_{T-1} = \sum_{i \in L'} \Delta \phi_i^T.$$

The following equality holds no matter whether blocked links exist or not. $E[\Delta R_T + \Delta \Phi_T | \Phi_T] \geq \frac{OPT}{\omega + 1}$

In the first scenario, there exists blocked link(s). A link is blocked because of some other link in the same cell. The total reward of the selected links at this time-slot is λ_c. Note that the values of $\Delta r_i^T + \Delta \phi_i^T$ of other selected links are non-negative.

$$E[\Delta r_i^T + \Delta \phi_i^T] \geq \frac{OPT}{\omega + 1}$$

If no link is blocked, then each link is either last observed $t < t_i$ time-slots ago or selected. In either case, we have

$$E[\Delta r_i^T + \Delta \phi_i^T] \geq H_{\lambda_c}^i.$$

To sum up, for all N links, we have

$$E[\Delta R_T + \Delta \Phi_T | \Phi_T] \geq \sum_{i \in L'} H_{\lambda_c}^i \geq \frac{OPT}{\omega + 1}.$$

Thus, no matter whether there is a link blocked or not, at each time-slot we have

$$E[\Delta R_T + \Delta \Phi_T | \Phi_T] \geq \frac{OPT}{\omega + 1}$$

Since we have $\Phi_T = \sum_{i \in L'} H_i^{\lambda_c}(\min(y, t_i) - 1 \leq \sum_{i \in L'} H_i^{\lambda_c}(t_i - 1)$, the value of Φ_T is bounded. We have

$$\lim_{t \to \infty} \frac{R_T}{T} \geq \frac{OPT}{\omega + 1}$$

Note that there is one cell with color k' among every $(K+1)^2$ cells. Based on the definition of R_k and the fact that k' corresponds to the largest reward $R_{k'}$ among all $(K+1)^2$ colors, Algorithm 1 outputs a $(K+1)^2 \cdot (\omega + 1)$ approximation solution.

5 Related Work

MAB can be classified into three categories, *i.e.*, stochastic bandits, adversarial bandits, and restless MAB. Here we only conduct the literature review for the third class *i.e.*, restless MAB, which is the focus of this paper. **Index Structure:** In [25], the pioneer work of Whittle suggested extending the multi-armed bandit [6,12] to a restless version, where an arm not selected to be active can also change its state. The investigation on Restless MAB has led to "Whittle's index" policies or algorithms. In such policies, an index called Whittle's index which is a mapping from a channel state to a real value is computed, each time we simply select the channels whose indices are above some threshold. In [14], Liu and Zhao studied the optimality of Whittle's index policy for multiple channels scheduling and computed a closed form of Whittle's index.

Myopic Algorithms: Ahmad *et al.* [2] proposed a myopic algorithm for stochastically identical and independent channels. They assume that one arm is activated. In [1], the myopic algorithm was extended to select multiple channels each time. Wang *et al.* [21–24] addressed the Restless MAB from the perspective of myopic algorithm for both two and multiple state Restless MAB. Comprehensive surveys on Restless MAB are available in [4,15,16]. However, their work either cannot provide a performance guarantee for opportunistic wireless scheduling or simply assume all links have the same transition probabilities.

Approximation Algorithm Perspective: Guha and Munagala [8,9] initiated the study on the Restless MAB problem from an approximation algorithm perspective. Wan and Xu [20] designed two approximation algorithms for Restless MAB. They considered both a weighted version and a multiple constrained version of the problem respectively. In [27], Xu and Song focused on the protocol interference model and proposed an interference-aware approximation algorithm for Restless MAB. The difference between their work and this study is that they either did not address wireless interference or use a over-simplified graph-based interference model.

In terms of the applications, the restless MAB has been applied to opportunistic downlink fading channel scheduling [18,20], network utility maximization [13], unmanned aerial vehicle routing [11], and energy harvesting [3].

6 Conclusion

We designed an efficient scheduling algorithm with an approximation guarantee for opportunistic link scheduling problem with SINR constraints. Some questions are left open in this area. The first one is to design a simple algorithm with idle performance guarantees for the simple but common case where all links have the same state transition matrix. The second one is to explore the connection between the opportunistic wireless link scheduling problem and the well studied maximum weighted independent set problem. The third one is that the state transition dynamics may be unknown. Under such a setting, it is critical to take a learning approach to estimate the state transition dynamics and design a scheduling policy. Thus, how to optimize the performance through learning is the key.

Acknowledgements. The research of authors is supported in part by KSU OVPR grant, NSFC Grant No. 61802097 and No. 61572377, and the Project of Qianjiang Talent (Grant No. QJD1802020).

References

1. Ahmad, S., Liu, M.: Multi-channel opportunistic access: a case of restless bandits with multiple plays. In: Allerton, pp. 1361–1368 (2009)
2. Ahmad, S.H.A., Liu, M., Javidi, T., Zhao, Q., Krishnamachari, B.: Optimality of myopic sensing in multichannel opportunistic access. IEEE Trans. Inf. Theory **55**(9), 4040–4050 (2009)
3. Blasco, P., Gündüz, D.: Multi-access communications with energy harvesting: a multi-armed bandit model and the optimality of the myopic policy. IEEE JSAC **33**(3), 585–597 (2015)
4. Bubeck, S., Cesa-Bianchi, N., et al.: Regret analysis of stochastic and nonstochastic multi-armed bandit problems. Found. Trends Mach. Learn. **5**(1), 1–122 (2012)
5. Gilbert, E., et al.: Capacity of a burst-noise channel. Bell Syst. Tech. J **39**(9), 1253–1265 (1960)
6. Gittins, J., Glazebrook, K., Weber, R.: Multi-armed Bandit Allocation Indices. Wiley, Hoboken (2011)
7. Goussevskaia, O., Wattenhofer, R., Halldórsson, M.M., Welzl, E.: Capacity of arbitrary wireless networks. In: IEEE INFOCOM, pp. 1872–1880 (2009)
8. Guha, S., Munagala, K.: Approximation algorithms for partial-information based stochastic control with Markovian rewards. In: IEEE FOCS, pp. 483–493 (2007)
9. Guha, S., Munagala, K., Shi, P.: Approximation algorithms for restless bandit problems. J. ACM (JACM) **58**(1), 3 (2010)
10. Gupta, P., Kumar, P.R.: The capacity of wireless networks. IEEE Trans. Inf. Theory **46**(2), 388–404 (2000)
11. Le Ny, J., Dahleh, M., Feron, E.: Multi-UAV dynamic routing with partial observations using restless bandit allocation indices. In: American Control Conference, pp. 4220–4225. IEEE (2008)
12. Li, B., Yang, P., Li, X.-Y., Tang, S., Liu, Y., Wu, Q.: Almost optimal dynamically-ordered multi-channel accessing for cognitive networks. In: IEEE INFOCOM, pp. 3081–3085. IEEE (2012)
13. Li, C.-P., Neely, M.J.: Network utility maximization over partially observable markovian channels. Perform. Eval. **70**(7), 528–548 (2013)
14. Liu, K., Zhao, Q.: Indexability of restless bandit problems and optimality of whittle index for dynamic multichannel access. IEEE Trans. Inf. Theory **56**(11), 5547–5567 (2010)
15. Lunden, J., Koivunen, V., Poor, H.V.: Spectrum exploration and exploitation for cognitive radio: recent advances. IEEE Signal Process. Mag. **32**(3), 123–140 (2015)
16. Mahajan, A., Teneketzis, D.: Multi-armed bandit problems. Found. Appl. Sens. Manag., 121–151 (2008)
17. Mitola, J., Maguire, G.Q., et al.: Cognitive radio: making software radios more personal. IEEE Pers. Commun. **6**(4), 13–18 (1999)

18. Ouyang, W., Murugesan, S., Eryilmaz, A., Shroff, N.B.: Exploiting channel memory for joint estimation and scheduling in downlink networks–a Whittle's indexability analysis. IEEE Trans. Inf. Theory **61**(4), 1702–1719 (2015)
19. Wan, P.-J., Jia, X., Dai, G., Du, H., Frieder, O.: Fast and simple approximation algorithms for maximum weighted independent set of links. In: IEEE INFOCOM, pp. 1653–1661 (2014)
20. Wan, P.-J., Xu, X.: Weighted restless bandit and its applications. In: IEEE ICDCS, pp. 507–516 (2015)
21. Wang, K., Chen, L.: On optimality of myopic policy for restless multi-armed bandit problem: an axiomatic approach. IEEE Trans. Signal Process. **60**(1), 300–309 (2012)
22. Wang, K., Chen, L., Liu, Q., Wang, W., Li, F.: One step beyond myopic probing policy: a heuristic lookahead policy for multi-channel opportunistic access. IEEE Trans. Wirel. Commun. **14**(2), 759–769 (2015)
23. Wang, K., Chen, L., Yu, J., Zhang, D.: Optimality of myopic policy for multistate channel access. IEEE Commun. Lett. **20**(2), 300–303 (2016)
24. Wang, K., Liu, Q., Li, F., Chen, L., Ma, X.: Myopic policy for opportunistic access in cognitive radio networks by exploiting primary user feedbacks. IET Commun. **9**(7), 1017–1025 (2015)
25. Whittle, P.: Restless bandits: activity allocation in a changing world. J. Appl. Probab. **25**, 287–298 (1988)
26. Xu, X., Li, X.-Y., Song, M.: Efficient aggregation scheduling in multihop wireless sensor networks with SINR constraints. IEEE Trans. Mob. Comput. **12**(12), 2518–2528 (2013)
27. Xu, X., Song, M.: Approximation algorithms for wireless opportunistic spectrum scheduling in cognitive radio networks. In: IEEE INFOCOM (2016)
28. Yu, D., Ning, L., Zou, Y., Yu, J., Cheng, X., Lau, F.C.: Distributed spanner construction with physical interference: constant stretch and linear sparseness. IEEE/ACM Trans. Netw. **25**(4), 2138–2151 (2017)
29. Yu, J., Huang, B., Cheng, X., Atiquzzaman, M.: Shortest link scheduling algorithms in wireless networks under the SINR model. IEEE Trans. Veh. Technol. **66**(3), 2643–2657 (2017)