Link Throughput of Multi-Channel Opportunistic Access with Limited Sensing

Keqin Liu,   Qing Zhao
University of California, Davis, CA 95616
kqliu@ucdavis.edu, qzhao@ece.ucdavis.edu

Abstract—We aim to characterize the maximum link throughput of a multi-channel opportunistic communication system. The states of these channels evolve as independent and identically distributed Markov processes (the Gilbert-Elliot channel model). A user, with limited sensing and access capability, chooses one channel to sense and access in each slot and collects a reward determined by the state of the chosen channel. Such a problem arises in cognitive radio networks for spectrum overlay, opportunistic transmissions in fading environments, and resource-constrained jamming and anti-jamming. The objective of this paper is to characterize the optimal performance of such systems. The problem can be generally formulated as obtaining the maximum expected long-term reward of a partially observable Markov decision process or a restless multi-armed bandit process, for which analytical characterizations are rare. Exploiting the structure and optimality of the myopic channel selection policy established recently, we obtain a closed-form expression of the maximum link throughput for two-channel systems and lower and upper bounds when there are more than two channels. These results allow us to study the rate at which the optimal performance of an opportunistic system increases with the number of channels and to obtain the limiting performance as the number of channels approaches to infinity.

Index Terms—Opportunistic access, cognitive radio, spectrum overlay, dynamic channel selection, myopic policy.

I. INTRODUCTION

The fundamental idea of opportunistic communications is to adapt the transmission parameters (data rate, modulation, transmission power, *etc.*) according to the state of the communication environment including, for example, fading conditions, interference level, and buffer state. Since the seminal work by Knopp and Humblet in 1995 [1], the concept of opportunistic communications has found applications beyond transmission over fading channels. An emerging application is cognitive radios for spectrum overlay (also referred to as opportunistic spectrum access), where secondary users search in the spectrum for idle channels temporarily unused by primary users [2]. Another application is resource-constrained jamming and anti-jamming, where a jammer seeks channels occupied by users or a user tries to avoid jammers.

We take a simplified model of these opportunistic communication systems with N parallel channels. These N channels are modelled as independent and identically distributed Gilbert-Elliot channels [3] as illustrated in Fig. 1. The state of a channel — "good" (1) or "bad" (0) — indicates the desirability of accessing this channel and determines the resulting reward. With limited sensing and access capability, a user chooses one of the channels to sense and access in each slot, aiming to maximize its expected long-term reward (*i.e.*, , throughput). The objective of this paper is to characterize analytically the maximum throughput of such a system. In particular, we are interested in the relationship between the maximum throughput and the number of channels.
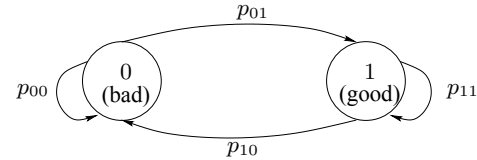
Fig. 1.   The Gilber-Elliot channel model.

This problem can be treated as a partially observable Markov decision process (POMDP) [4] or more specifically, a restless multi-armed bandit process [5] due to the independence across channels. The maximum throughput of the multi-channel opportunistic system is essentially the maximum expected total reward, or the value function, of a POMDP [6]. Unfortunately, obtaining optimal solutions to POMDPs, even numerically, is often intractable, and closed-form expressions for value functions are rare.

In this paper, we obtain a closed-form expression of the maximum throughput for two-channel opportunistic systems. For systems with more than two channels, we develop lower and upper bounds that monotonically tighten as the number N of channels increases. These results allow us to study the rate at which the optimal performance of an opportunistic system increases with N and to obtain the limiting performance as N approaches to infinity. They demonstrate that the optimal link throughput of a multi-channel opportunistic system with limited sensing quickly saturates as the number of channel increases.

Our analysis hinges on the structure and optimality of the myopic policy established in [7], [8]. The optimality of the myopic policy makes it sufficient to obtain the maximum throughput from the performance of the myopic policy, and the simple structure of the myopic policy makes it possible to characterize analytically its performance. Specifically, based on the structure of the myopic policy, we show that the performance of the myopic policy is determined by the stationary distributions of a higher-order countable-state Markov chain. For $N = 2$, we obtain the stationary distribution of this Markov chain in closed-form, leading to exact characterizations of the maximum throughput. For $N > 2$, we construct first-order Markov processes that stochastically dominate or are dominated by this higher-order Markov chain. The stationary distributions of the former, again obtained in closed-forms, lead to lower and upper bounds on the maximum throughput.

II. PROBLEM FORMULATION

We consider the scenario where a user is trying to access the wireless spectrum using a slotted transmission structure. The spectrum consists of N independent and statistically identical channels. The state $S_i(t)$ of channel i in slot t is given by a two-state discrete-time Markov chain shown in Fig. 1.

At the beginning of each slot, the user selects one of the N channels to sense. If the channel is sensed to be in the "good" state (state 1), the user transmits and collects one unit of reward. Otherwise

the user does not transmit (or transmits at a lower rate), collects no reward, and waits until the next slot to make another choice. The objective is to maximize the average reward (throughput) over a horizon of T slots by choosing judiciously a sensing policy that governs channel selection in each slot.

Due to limited sensing, the system state $[S_1(t), \cdots, S_N(t)] \in \{0, 1\}^N$ in slot t is not fully observable to the user. It can, however, infer the state from its decision and observation history. It has been shown that a sufficient statistic of the system for optimal decision making is given by the conditional probability that each channel is in state 1 given all past decisions and observations [4]. Referred to as the belief vector, this sufficient statistic is denoted by $\Omega(t) \triangleq [\omega_1(t), \cdots, \omega_N(t)]$, where $\omega_i(t)$ is the conditional probability that $S_i(t) = 1$. Given the sensing action a and the observation S_a in slot t, the belief vector for slot $t + 1$ can be obtained as follows.

$$\omega_i(t+1) = \begin{cases} p_{11}, & a = i, S_a = 1 \\ p_{01}, & a = i, S_a = 0 \\ \omega_i(t)p_{11} + (1 - \omega_i(t))p_{01}, & a \neq i \end{cases} \quad (1)$$

A sensing policy π specifies a sequence of functions $\pi = [\pi_1, \pi_2, \cdots, \pi_T]$ where π_t maps a belief vector $\Omega(t)$ to a sensing action $a(t) \in \{1, \cdots, N\}$ for slot t. Multi-channel opportunistic access can thus be formulated as the following stochastic control problem.

$$\pi^* = \arg \max_\pi \mathbb{E}_\pi \left[\sum_{t=1}^{T} R(\pi_t(\Omega(t))) | \Omega_o \right],$$

where $R(\pi_t(\Omega(t)))$ is the reward obtained when the belief is $\Omega(t)$ and channel $\pi_t(\Omega(t))$ is selected, and Ω_o is the initial belief vector. If no information on the initial system state is available, each entry of Ω_o can be set to the stationary distribution ω_o of the underlying Markov chain:

$$\omega_o = \frac{p_{01}}{p_{01} + p_{10}}. \quad (2)$$

III. STRUCTURE AND OPTIMALITY OF MYOPIC POLICY

A. The Value Function

Let $V_t(\Omega)$ be the value function, which represents the maximum expected total reward that can be obtained starting from slot t given the current belief vector Ω. Given that the user takes action a and observes S_a, the reward that can be accumulated starting from slot t consists of two parts: the immediate reward $R_a(\Omega) = \omega_a$ and the maximum expected future reward $V_{t+1}(\mathcal{T}(\Omega | a, s_a))$, where $\mathcal{T}(\Omega | a, s_a)$ denotes the updated belief vector for slot $t + 1$ as given in (1). Averaging over all possible observations S_a and maximizing over all actions a, we arrive at the following optimality equation.

$$\begin{aligned} V_T(\Omega) &= \max_{a=1,\cdots,N} \omega_a \\ V_t(\Omega) &= \max_{a=1,\cdots,N} (\omega_a + \omega_a V_{t+1}(\mathcal{T}(\Omega | a, 1))) \\ &\quad + (1 - \omega_a)V_{t+1}(\mathcal{T}(\Omega | a, 0)). \end{aligned} \quad (3)$$

In theory, the optimal policy π^* and its performance $V_1(\Omega_o)$ can be obtained by solving the above dynamic programming. Unfortunately, due to the impact of the current action on the future reward and the uncountable space of the belief vector Ω, obtaining the optimal solution using directly the above recursive equations is computationally prohibitive. Even when approximate numerical solutions can be obtained, they do not provide insight into system design or analytical characterizations of the optimal performance $V_1(\Omega_o)$.

B. The Myopic Policy

A myopic policy ignores the impact of the current action on the future reward, focusing solely on maximizing the expected immediate reward $\mathbb{E}[R_a]$. Myopic policies are thus stationary. The myopic action $\hat{a}$ under belief state $\Omega = [\omega_1, \cdots, \omega_N]$ is simply given by

$$\hat{a}(\Omega) = \arg \max_{a=1,\cdots,N} \omega_a. \quad (4)$$

In general, obtaining the myopic action in each slot requires the recursive update of the belief vector Ω as given in (1), which requires the knowledge of the transition probabilities $\{p_{ij}\}$. Interestingly, it has been shown in [7] that the myopic policy has a simple structure that does not need the update of the belief vector or the precise knowledge of the transition probabilities.

Specifically, when $p_{11} \geq p_{01}$, the myopic action is to stay in the same channel if the channel in the current slot is in state 1. Otherwise, the user switches to the channel visited the longest time ago. The channel selection is thus in a round robin fashion as illustrated in Fig. 2: sense N channels in turn with a random switching time (when the current channel transits to state 0).
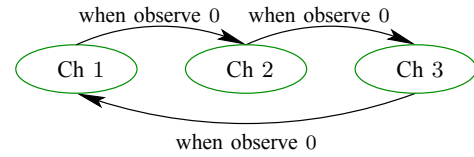

Fig. 2. The structure of the myopic policy for $p_{11} \geq p_{01}$ ($N = 3$).

When $p_{11} < p_{01}$, the myopic action is to stay in the same channel when the channel is in state 0 and switch otherwise. When a channel switch is needed, the user chooses, among those channels to which the last visit occurred an even number of slots ago, the one most recently visited. If there are no such channels, the user chooses the channel visited the longest time ago.

Note that the above simple structure of the myopic policy reveals that other than the order of p_{11} and p_{01}, the knowledge of the transition probabilities are unnecessary.

Surprisingly, the myopic policy with such a simple and robust structure achieves the optimal performance for $N = 2$ [7]. It has been conjectured in [7] (based on simulation results) that the optimality of the myopic policy holds for all N. In a recent work [8], the optimality of the myopic policy has been established for a general N under the condition of $0 \leq p_{11} - p_{01} \leq 0.5$.

IV. LINK THROUGHPUT LIMIT

The objective of this paper is to characterize the link throughput limit U of multi-channel opportunistic access with limited sensing. Specifically, we define the link throughput limit as

$$U(\Omega_o) \triangleq \lim_{T \to \infty} \frac{\hat{V}_1(\Omega_o)}{T},$$

where $\hat{V}_1(\Omega_o)$ is the expected total reward obtained in T slots under the myopic policy when the initial belief is Ω_o.

Our analysis hinges on the structure and optimality of the myopic policy given in Sec. III-B. The optimality of the myopic policy makes it sufficient to obtain U from the performance of the myopic policy, and the simple structure of the myopic policy makes it possible to characterize analytically its performance.

From the structure of the myopic policy we can see that the key to the throughput is how often the user switches channels, or equivalently, how long it stays in the same channel. When $p_{11} \geq p_{01}$,

the event of channel switch is equivalent to a slot *without* reward. The opposite holds when $p_{11} < p_{01}$: a channel switch corresponds to a slot *with* reward.
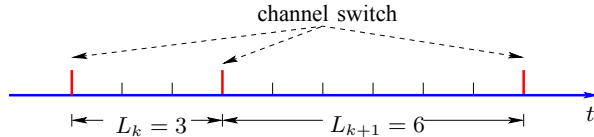

Fig. 3. The transmission period structure.

We thus introduce the concept of transmission period, which is the time the user stays in the same channel, as illustrated in Fig. 3. Let L_k denote the length of the kth transmission period. We thus have a discrete-time random process $\{L_k\}_{k=1}^{\infty}$ with a sample space of positive integers. Let $\bar{L} = \lim_{K \to \infty} \frac{\Sigma_{k=1}^{K} L_k}{K}$ denote the average length of a transmission period.

Lemma 1: The throughput limit U is given by

$$U(\Omega_o) = \begin{cases} 1 - \frac{1}{\bar{L}}, & p_{11} \geq p_{01} \\ \frac{1}{\bar{L}}, & p_{11} < p_{01} \end{cases}. \tag{5}$$

Proof: See [9]. ∎

For $N = 2$, the uniqueness and closed-form expressions of $\bar{L}$ can be established, which leads to a closed-form expression of the throughput limit $U(\Omega_o)$ (see Sec. IV-A). For $N > 2$, lower and upper bounds on $U(\Omega_o)$ are obtained (see Sec. IV-B).

Proposition 1: : $U(\Omega_o) \equiv U$ is independent of the initial belief Ω_o for (i) $N = 2$; (ii) $N > 2$ and $p_{11} \geq p_{01}$.

Proof: : See [9]. ∎

A. Link Throughput Limit for $N = 2$

For $N = 2$, $\{L_k\}_{k=1}^{\infty}$ is a first-order Markov chain. We have the following lemma.

Lemma 2: $\{L_k\}_{k=1}^{\infty}$ is an irreducible and aperiodic first-order Markov chain with the following unique stationary distribution (the limiting distribution).

- Case 1: $p_{11} \geq p_{01}$

$$\lambda_l = \begin{cases} 1 - \bar{\omega}, & l = 1 \\ \bar{\omega} p_{11}^{l-2} p_{10}, & l \geq 2 \end{cases}, \tag{6}$$

where $\bar{\omega}$ is the expected probability that the channel we switch to is in state 1, *i.e.,* , the expected belief value of the channel we switch to. It is given by

$$\bar{\omega} = \frac{p_{01}^2}{1 + p_{01}^2 - A}, \tag{7}$$

where $A = \frac{p_{01}}{1 + p_{01} - p_{11}}(1 - \frac{(p_{11} - p_{01})^3(1 - p_{11})}{1 - (p_{11})^2 + p_{11}p_{01}})$.

- Case 2: $p_{11} < p_{01}$

$$\lambda_l = \begin{cases} \bar{\omega}', & l = 1 \\ (1 - \bar{\omega}') p_{00}^{l-2} p_{01}, & l \geq 2 \end{cases}, \tag{8}$$

where $\bar{\omega}'$ is the expected probability that the channel we switch to is in state 1. It is given by

$$\bar{\omega}' = \frac{B}{1 - p_{11}^2 + B}, \tag{9}$$

where $B = \frac{p_{01}}{1 + p_{01} - p_{11}}(1 + \frac{(p_{11} - p_{01})^3(1 - p_{11})}{1 - (1 - p_{01})(p_{11} - p_{01})})$.

Proof: See [9]. ∎

From Lemma 1 and 2, we obtain the closed-form expression of the throughput limit U as follows.

Theorem 1: For $N = 2$, the throughput limit U is given by

$$U = \begin{cases} 1 - \frac{1 - p_{11}}{1 + \bar{\omega} - p_{11}}, & p_{11} \geq p_{01} \\ \frac{p_{01}}{1 - \bar{\omega}' + p_{01}}, & p_{11} < p_{01} \end{cases}, \tag{10}$$

where $\bar{\omega}$ and $\bar{\omega}'$ are given, respectively, in (7) and (9).

Proof: See [9]. ∎

B. Link Throughput Limit for $N > 2$

For $N > 2$, $\{L_k\}_{k=1}^{\infty}$ is a higher-order Markov chain. It is difficult to obtain its stationary distributions in closed form. Our objective is to develop lower and upper bounds on U.

The approach is to construct first-order Markov chains that stochastically dominate or are dominated by $\{L_k\}_{k=1}^{\infty}$. The limiting distributions of these first-order Markov chains, which can be obtained in closed-form, thus lead to lower and upper bounds on U according to Lemma 1. Specifically, for $p_{11} \geq p_{01}$, a lower bound on U is obtained by constructing a first-order Markov chain whose limiting distribution is stochastically dominated by the stationary distributions of $\{L_k\}_{k=1}^{\infty}$. An upper bound on U is given by a first-order Markov chain whose limiting distribution stochastically dominates the stationary distributions of $\{L_k\}_{k=1}^{\infty}$. Similarly, bounds on U for $p_{11} < p_{01}$ can be obtained.

Theorem 2: For $N > 2$, we have the following lower and upper bounds on the throughput limit U.

- Case 1: $p_{11} \geq p_{01}$

$$\frac{C}{C + (1 - D + C)(1 - p_{11})} \leq U \leq \frac{\omega_o}{1 - p_{11} + \omega_o}, \tag{11}$$

where ω_o is given by (2), $C = \omega_o(1 - (p_{11} - p_{01})^N)$,

$D = \omega_o(1 - \frac{(p_{11} - p_{01})^{N+1}(1 - p_{11})}{1 - (p_{11})^2 + p_{11}p_{01}})$.

- Case 2: $p_{11} < p_{01}$

$$\frac{p_{10}^2}{p_{01}H - E} + 1 \leq U(\Omega_o) \leq \frac{p_{10}^2}{p_{01}G - E} + 1 \tag{12}$$

where $E = p_{10}^2(1 + p_{01}) + p_{01}(1 - F)$,

$$F = (1 - p_{01})(1 - \omega_o)(\frac{1}{2 - p_{01}} - \frac{p_{01}(p_{11} - p_{01})^4}{1 - (p_{11} - p_{01})^2(1 - p_{01})^2}),$$

$$G = (1 - \omega_o)(\frac{1}{2 - p_{01}} - \frac{p_{01}(p_{11} - p_{01})^6}{1 - (p_{11} - p_{01})^2(1 - p_{01})^2})$$

$$H = (1 - \omega_o)(\frac{1}{2 - p_{01}} - \frac{p_{01}(p_{11} - p_{01})^{2N-1}}{1 - (p_{11} - p_{01})^2(1 - p_{01})^2})$$

- Monotonicity: for both cases, the difference between the upper and lower bounds monotonically decreases with N; when $p_{11} \geq p_{01}$, the lower bound converges to the upper bound as $N \to \infty$.

Proof: See [9]. ∎

The monotonicity of the difference between the upper and lower bounds with respect to N shows that the performance of the multi-channel opportunistic system improves with the number N of channels, as suggested by intuition. Note that the upper bounds on U for both cases are independent of N. For $p_{11} \geq p_{01}$, the upper bound gives the limiting performance of the opportunistic system when $N \to \infty$.

V. NUMERICAL EXAMPLES

In this section, we demonstrate the tightness of the bounds on U given in Sec. IV-B by examining the relative difference $d(N)$ between the upper and the lower bound. In Fig. 4, we plot $d(N = 5)$ with respect to the upper bound for $p_{11} \geq p_{01}$. From Fig. 4 we observe that for most values of p_{11} and p_{01}, $d(N = 5)$ is below 6%, demonstrating the tightness of the bounds even for a small number of channels. Furthermore, Fig. 4 shows that the bounds are tighter for larger p_{01}. Similarly observations can be drawn from Fig. 5 for the case of $p_{11} < p_{01}$.
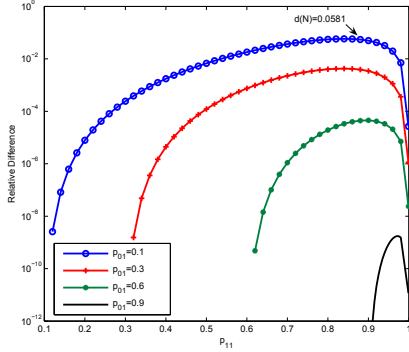

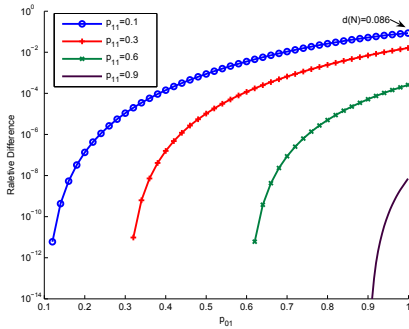Fig. 4. The relative difference $d(N = 5)$ between the upper and the lower bound for $p_{11} \geq p_{01}$.

Fig. 5. The relative difference $d(N = 5)$ between the upper and the lower bound for $p_{11} < p_{01}$.

In Fig. 6 and 7 we examine the rate at which the lower bound approaches to the upper bound as N increases. Specifically, we plot the ratio of $d(N = 10)$ to $d(N = 3)$. We observe that in both cases, the lower bound approaches to the upper bound quickly. While demonstrating the usefulness of the bounds for small N, this observation conveys a pessimistic message: the optimal link throughput of a multi-channel opportunistic system with limited sensing quickly saturates as N increases.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have analyzed the performance of myopic sensing in multi-channel opportunistic access under an i.i.d. Gilbert-Elliot channel model. Based on the conjectured optimality of myopic sensing policy, the obtained analytical results allow us to systematically examine the impact of the number of channels and channel dynamics (transition probabilities) on the system performance. Future work
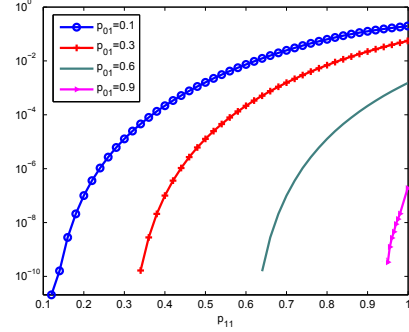

Fig. 6. The rate at which the lower bound approaches to the upper bound as N increases ($p_{11} \geq p_{01}$).
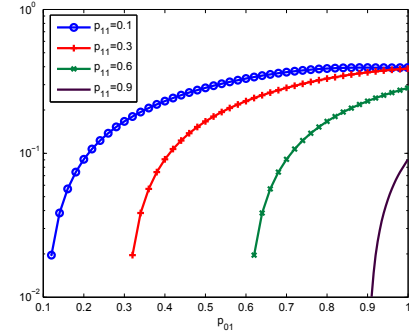

Fig. 7. The rate at which the lower bound approaches to the upper bound as N increases ($p_{11} < p_{01}$).

includes the generalization to cases with sensing errors and non-identical channels. The former can again be addressed by exploiting the structure and optimality of the myopic policy in the presence of sensing errors as established in [10].

REFERENCES

[1] R. Knopp and P. Humblet, "Information capacity and power control in single cell multi-user communications," in *Proc. Intl Conf. Comm.*, (Seattle, WA), pp. 331–335, June 1995.

[2] Q. Zhao and B. Sadler, "A Survey of Dynamic Spectrum Access," *IEEE Signal Processing magazine*, vol. 24, pp. 79–89, May 2007.

[3] E.N. Gilbert, "Capacity of burst-noise channels," Bell Syst. Tech. J., vol. 39, pp. 1253-1265, Sept. 1960.

[4] R. Smallwood and E. Sondik, "The optimal control of partially ovservable Markov processes over a finite horizon," *Operations Research*, pp. 1071–1088, 1971.

[5] P. Whittle, "Restless bandits: Activity allocation in a changing world", in *Journal of Applied Probability*, Volume 25, 1988.

[6] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework," *IEEE JSAC*, April 2007.

[7] Q. Zhao and B. Krishnamachari, "Structure and optimality of myopic sensing for opportunistic spectrum access," in *Proc. of IEEE CogNet*, pp. 6476–6481, June, 2007.

[8] T. Javidi, B. Krishnamachari, Q. Zhao, and M. Liu, "Optimality of Myopic Sensing in Multi-Channel Opportunistic Access," to appear in Proc. of ICC, May, 2008.

[9] K. Liu and Q. Zhao, "Link Throughput of Multi-Channel Opportunistic Access with Limited Sensing," Technical Report, Univ. of California, Davis, July, 2007, http://www.ece.ucdavis.edu/~qzhao/Report.html.

[10] Q. Zhao, B. Krishnamachari, and K. Liu, "Low-Complexity Approaches to Spectrum Opportunity Tracking," in *Proc. of CrownCom*, August 2007.