

Quickest Change Detection In Multiple On-off Processes: Switching With Memory

Jia Ye, Qing Zhao

Department of Electrical and Computer Engineering, University of California, Davis, CA 95616
jiaye@ucdavis.edu, qzhao@ece.ucdavis.edu

Abstract—We consider the quickest detection of idle periods in multiple on-off processes. At each time, only one process can be observed, and the observations are random realizations drawn from two different distributions depending on the current state (on or off) of the chosen process. Switching back to a previously visited process is allowed, and measurements obtained during previous visits are taken into account in decision making. The objective is to catch an idle period in any of the on-off processes as quickly as possible subject to a constraint on the probability of mistaking a busy period for an idle one.

Assuming geometrically distributed busy and idle times, we establish a Bayesian formulation of the problem within a decision-theoretic framework. Basic structures of the optimal decision rules are established. Based on these basic structures, we propose a low-complexity threshold policy for switching among processes and declaring idle periods. The near optimal performance of this threshold policy is demonstrated by a comparison with a genie-aided system which defines an upper bound on the optimal performance. This problem finds applications in spectrum opportunity detection in cognitive radio networks where a secondary user searches for idle channels in the spectrum.

Index Terms—Quickest change detection, on-off process, spectrum opportunity detection, cognitive radio, genie-aided system

I. INTRODUCTION

The problem of quickest change detection was first studied in 1931 [1] for the application of on-line quality control of a manufacturing process. In the conventional setting, we have a single random process $X_1, \dots, X_{T_0-1}, X_{T_0}, \dots$. Before a random change point T_0 , the observations X_1, \dots, X_{T_0-1} are i.i.d according to a distribution f_0 ; after T_0 , the observations $X_{T_0}, X_{T_0+1}, \dots$, are i.i.d with a different distribution f_1 . The objective is to detect the change point T_0 as quickly as possible subject to a reliability constraint.

There are two standard mathematical formulations of the classic quickest change detection in a single stochastic process: Bayesian and minimax. The Bayesian formulation was developed by Shiriyayev in 1960's [2] [3], where the change point is assumed to have a geometric/exponential distribution and the objective is to minimize the expected detection delay subject to a constraint on the probability of false alarm. Generalizations of Shiriyayev's algorithm to arbitrary prior distributions of the change point and non-i.i.d observations have been studied (see, for example, [4], [5]). The minimax formulation was proposed by Lorden in 1971 [6], in which the priori

distribution of the change point is unknown and the objective is to minimize the worst-case conditional detection delay subject to a lower bound on the allowable mean time between false alarms. It was shown in [6] that the well-known cumulative sum (CUSUM) algorithm proposed by Page in 1954 [7] is asymptotically optimal under the minimax formulation.

An emerging application of quickest change detection is spectrum opportunity detection in cognitive radio networks where a secondary user searches for idle channels in the spectrum. [8]. The objective is to detect, as soon as possible, whether the sensed channel has become idle in order to maximize the transmission time before primary users reclaim the channel. The design constraint is on the interference to primary users, *i.e.*, the probability of declaring a busy channel as idle.

While the problem of spectrum opportunity detection appears to fit into the classic framework of quickest change detection, two distinct features of the cognitive radio systems call for a new formulation of and new solutions to the problem. Specifically, the busy/idle state of a channel constitutes an on-off process with *multiple* change points. Furthermore, there are multiple channels in the spectrum; the secondary user does not have to wait faithfully in a single channel for an idle period.

The above observations motivate the problem of quickest change detection in multiple on-off processes, which was first formulated and studied in our previous work [9] [10]. In [9] [10], we considered a large number of homogeneous independent channels, and the user always switches to a new channel should it decide to abandon the current channel. It was revealed in [9] [10] that the key to quickest detection in multiple processes lies in the tradeoff between avoiding long realizations of busy periods via channel switching and losing measurements obtained in the abandoned processes. The problem was formulated as a partially observable Markov decision process (POMDP), and the optimal policy was shown to have a simple threshold structure under the assumption of channel switching without memory [10].

Differing from our previous work [9] [10], this paper addresses quickest detection with memory: switching back to a previously visited process is allowed, and measurements obtained during previous visits are taken into account in decision making. We show that this freedom of switching with memory significantly complicates the problem. The resulting POMDP changes from a one-dimensional problem to an N -dimensional problem, where N is the number of channels. The objective of this paper is to establish basic structures of the optimal policy

⁰This work was supported by the Army Research Laboratory under Grant DAAD19-01-C-0062, by the Army Research Office under Grant W911NF-08-1-0467, and by the National Science Foundation under Grant CCF-0830685.

of this N -dimensional POMDP and develop low-complexity policies with strong performance. In particular, we show that the optimal action of declaring always occurs in the channel with the largest posterior probability of being in the idle state. The monotonicity of the detection threshold is also established. Based on these basic structures, we propose a low-complexity threshold policy for switching among processes and declaring idle periods. Specifically, under the proposed policy, the user always observe the process with the largest posterior probability of being idle and declare if the largest posterior probability exceeds the declaring threshold. The near optimal performance of this threshold policy is demonstrated by a comparison with a genie-aided system which defines an upper bound on the optimal performance. Furthermore, we show that this low-complexity policy converges to the optimal policy for the infinite-channel case developed in [8] as the number N of channels increases.

II. THE PROBLEM STATEMENT AND A POMDP FORMULATION

In this section, we present the POMDP formulation of quickest detection in multiple on-off processes. We use spectrum opportunity detection in cognitive radio systems as an example application of this problem.

A. Problem Statement

Consider N homogeneous and independent on-off processes (the extension of the POMDP formulation to cases with heterogeneous channels is straightforward). Let $\{B_i\}_{i=-\infty}^{\infty}$ and $\{I_i\}_{i=-\infty}^{\infty}$ denote, respectively, the lengths of each busy and idle periods in the i th process. We assume that the busy periods $\{B_i\}_{i=-\infty}^{\infty}$ have a geometric distribution with parameter p_B , and the idle periods $\{I_i\}_{i=-\infty}^{\infty}$ have a geometric distribution with parameter p_I . The average busy and idle times are thus given by $m_B = 1/p_B$ and $m_I = 1/p_I$, respectively. Let λ_0 denote the fraction of channel idle time. It is given by

$$\lambda_0 \triangleq \frac{m_I}{m_B + m_I}. \quad (1)$$

As shown in Figure 1 (where $N = 3$), suppose that a secondary user starts to sense a channel at $t = 0$. The time unit here is the secondary user's sampling period (the time for taking one channel measurement). The objective is to catch an idle channel and start transmitting as quickly as possible subject to an interference constraint that caps the probability of transmitting over a busy channel below ζ . At each time instant, the user may choose any of the N channels to sense. Switching back to a channel that has been visited before is allowed, and the measurements obtained during previous visits to this channel will be taken into account in decision making.

The problem is a sequential decision making problem: at each time instant, based on all the observations obtained so far, the user decides whether to declare or to continue sensing and in which channel such an action (declare or continue) should be taken. The objective is to minimize the total detection time T_d subject to the constraint on the probability of false alarm.

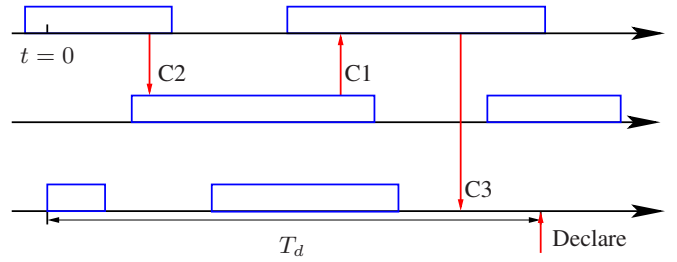


Fig. 1. Quickest detection of spectrum opportunities.

B. POMDP Formulation

We now formulate the problem as an N -dimensional POMDP over a random horizon.

State Space: The system state at time t is given by $\{Z_1(t), \dots, Z_N(t)\}$, where $Z_i(t) \in \{0(\text{busy}), 1(\text{idle})\}$ denotes the state of channel i at time t . We then augment the state space by an absorbing state Δ which indicates the end of the decision horizon. Whenever the action of declaring is taken, the state of the system transits to Δ .

Action Space: The action space is $\{C_i, D_i, i = 1, \dots, N\}$ where C_i denotes the action of continuing taking measurements in channel i , D_i denotes the action of declaring that an idle state has been reached in the i th channel.

State Transition: The transition probabilities of channel i under all possible actions are given in Figure 2. Note that the busy/idle state of channel i evolves independently of the action as long as the user decides to continue sensing. Whenever the action of declaring is taken, the system state transits to Δ .

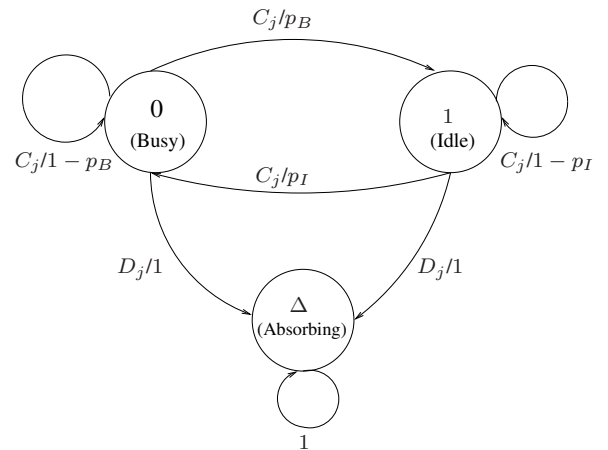


Fig. 2. The state transition diagram.

Observation Model: The observation at time t is X_t under actions C_i . The distribution of X_t is given by either $f_0(x)$ or $f_1(x)$ depending on the current state $Z_i(t)$. Under action D_i , no observation is available.

Cost: The action C_i has a unit cost that measures the delay in catching an idle period. Declaring a busy channel as idle incurs a cost of γ that models the tradeoff between detection delay and detection reliability. It is set to satisfy the interference constraint ζ . Note that it is not necessary to specify the value

of γ based on ζ . As shown in Sec. III, the optimal detection rule is specified by a detection threshold chosen to satisfy the interference constraint ζ .

The objective is to choose actions sequentially in time to minimize the expected total cost over an infinite horizon, or equivalently, over a random horizon defined by the hitting time of the absorbing state Δ . It is clear from the cost structure that the expected total cost is the expected delay in catching an idle channel.

A Sufficient Statistic: Since the full system state $[Z_1(t), \dots, Z_N(t)] \in \{0, 1\}^N$ at time t is not observable, what we have here is a POMDP. From the fundamental theory of stochastic control, we know that a sufficient statistic for choosing the optimal action at each time is the information state or the belief vector: $\Lambda(t) \triangleq [\lambda_1(t), \dots, \lambda_N(t)]$, where $\lambda_i(t)$ is the posterior probability that channel i is in idle state given all the past measurements taken on channel i .

Given the action a_t and the observation X_t at time t (if observation is available), the belief vector can be updated as follows.

$$\lambda_i(t) = \begin{cases} \mathcal{T}(\lambda_i(t-1)|x) & a_{t-1} = C_i, X_t = x \\ \mathcal{T}(\lambda_i(t-1)) & a_{t-1} = C_j, j \neq i \end{cases}, \quad (2)$$

where $\mathcal{T}(\lambda|x)$ denotes the updated information state based on a new measurement x , and $\mathcal{T}(\lambda)$ denotes the updated information state based purely on the underlying Markov chain defined by the geometric distributions of the busy and idle periods. Let $\bar{p} \triangleq 1 - p$ for $p \in [0, 1]$. We obtain $\mathcal{T}(\lambda|x)$ and $\mathcal{T}(\lambda)$ as follows.

$$\mathcal{T}(\lambda|x) = \frac{(\lambda\bar{p}_I + \bar{\lambda}p_B)f_1(x)}{(\lambda\bar{p}_I + \bar{\lambda}p_B)f_1(x) + (\lambda p_I + \bar{\lambda}\bar{p}_B)f_0(x)} \quad (3)$$

$$\mathcal{T}(\lambda) = \lambda\bar{p}_I + \bar{\lambda}p_B. \quad (4)$$

A channel selecting and change detection policy π specifies a function that maps a belief vector $\Lambda(t)$ to an action $a_t = \pi(\Lambda(t))$ for each time t . Quickest change detection in multiple on-off processes switching with memory can thus be formulated as the following stochastic control problem:

$$\pi^* = \arg \min_{\pi} \mathbb{E}_{\pi} \left[\sum_{t=0}^{\infty} R_{\pi(\Lambda(t))} | \Lambda(0) \right], \quad (5)$$

where $\pi(\Lambda(t))$ is the action specified by policy π under belief vector $\Lambda(t)$, $R_{\pi(\Lambda(t))}$ is the cost incurred under this action, and $\Lambda(0) = \{\lambda_0, \dots, \lambda_0\}$ where $\lambda_0 = \frac{m_I}{m_I + m_B}$.

III. BASIC STRUCTURE OF THE OPTIMAL POLICY

In this section, we establish the basic structure of the optimal policy. For simplicity of the presentation, we consider the case of $N = 2$. Extensions to $N > 2$ is straightforward.

Let $V(\lambda_1, \lambda_2)$ denote the minimum expected total remaining cost when the current belief vector is (λ_1, λ_2) . It specifies the performance of the optimal policy π^* starting from the belief vector (λ_1, λ_2) . Let $V_{C_1}(\lambda_1, \lambda_2)$ denote the expected total remaining cost when we take action C_1 at the current

time and then follow the optimal policy π^* in the future. Let $V_{C_2}(\lambda_1, \lambda_2)$, $V_{D_1}(\lambda_1, \lambda_2)$ and $V_{D_2}(\lambda_1, \lambda_2)$ be similarly defined. We thus have

$$V(\lambda_1, \lambda_2) = \min \{ V_{C_1}(\lambda_1, \lambda_2), V_{C_2}(\lambda_1, \lambda_2), V_{D_1}(\lambda_1, \lambda_2), V_{D_2}(\lambda_1, \lambda_2) \}. \quad (6)$$

From the cost structure, we obtain the following:

$$\begin{aligned} V_{C_1}(\lambda_1, \lambda_2) &= 1 + \int_x P(x; \lambda_1) V(\mathcal{T}(\lambda_1|x), \mathcal{T}(\lambda_2)) dx, \\ V_{C_2}(\lambda_1, \lambda_2) &= 1 + \int_x P(x; \lambda_2) V(\mathcal{T}(\lambda_1), \mathcal{T}(\lambda_2|x)) dx, \\ V_{D_1}(\lambda_1, \lambda_2) &= (1 - \lambda_1)\gamma, \\ V_{D_2}(\lambda_1, \lambda_2) &= (1 - \lambda_2)\gamma, \end{aligned} \quad (7)$$

where

$$P(x; \lambda) \triangleq (\lambda\bar{p}_I + \bar{\lambda}p_B)f_1(x) + (\lambda p_I + \bar{\lambda}\bar{p}_B)f_0(x)$$

is the probability of observing x when the process has probability λ to be idle. It is easy to see that $V_{D_i}(\lambda_i)$ is linearly decreasing with λ_i for all i .

Theorem 1: When $p_B + p_I \leq 1$, the basic structure of the optimal policy π^* for channel selection and change detection is given in Figure 3. Specifically, the optimal action $a^*(\lambda_1, \lambda_2)$ under the belief vector (λ_1, λ_2) is in the following form.

$$a^*(\lambda_1, \lambda_2) = \begin{cases} D_1, & \text{if } \lambda_1 \geq \lambda_2, \lambda_1 > \eta(\lambda_2) \\ D_2, & \text{if } \lambda_1 < \lambda_2, \lambda_2 > \eta(\lambda_1) \\ C_1 \text{ or } C_2, & \text{Otherwise} \end{cases}. \quad (8)$$

where $\eta(\lambda) : [0, 1] \rightarrow [0, 1]$ is the detection threshold which is monotonically increasing in λ . Furthermore, the action of continuing has a symmetric structure, *i.e.*,

$$a^*(\lambda_1, \lambda_2) = C_1 \Leftrightarrow a^*(\lambda_2, \lambda_1) = C_2. \quad (9)$$

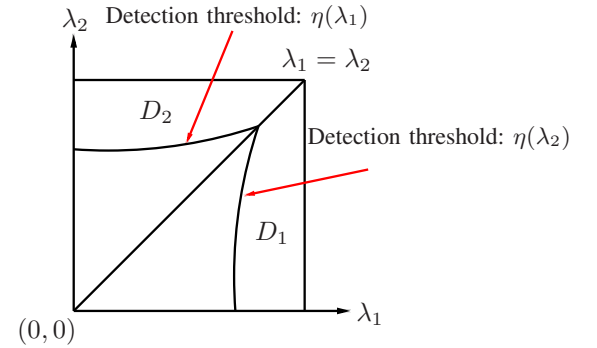


Fig. 3. The basic structure of the optimal policy.

Proof: This Theorem is proved based on the following lemmas. Details are omitted due to the space limit.

Lemma 1: $V_{C_1}(\lambda_1, \lambda_2)$ and $V_{C_2}(\lambda_1, \lambda_2)$ are concave w.r.t λ_1 and λ_2 respectively.

Lemma 2: When $p_B + p_I \leq 1$, $V_{C_1}(\lambda_1, \lambda_2)$ and $V_{C_2}(\lambda_1, \lambda_2)$ are monotonically decreasing w.r.t λ_1 and λ_2 respectively.

Lemma 3: $V_{C_1}(\lambda_1, \lambda_2)$ and $V_{C_2}(\lambda_1, \lambda_2)$ are symmetric w.r.t the plane $\lambda_1 = \lambda_2$, *i.e.*, $V_{C_1}(\lambda_1, \lambda_2) = V_{C_2}(\lambda_2, \lambda_1)$.

We point out that the assumption of $p_B + p_I < 1$ imposes a mild condition on the system which generally holds. For example, if the average busy and idle times are more than two sample periods, the condition is satisfied. The detection threshold $\eta(\lambda)$ is chosen to satisfy the interference constraint ζ . Setting $\eta(\lambda) \equiv 1 - \zeta$ always meets the constraint but potentially leads to suboptimal performance.

The basic structure of the optimal policy given in Theorem 1 holds for the general case of $N > 2$. Specifically, under the optimal policy, the user always declares on the channel with the largest belief value, and the declaring threshold is monotonically increasing with the belief values of the rest $N - 1$ channels.

IV. A LOW-COMPLEXITY THRESHOLD POLICY

In this section, we propose a low-complexity threshold policy based on the basic structure of the optimal policy established in III.

Again, we first focus on the two-channel case. Based on the basic structure of the optimal policy, one possibility of the relationship among $V_{C_1}(\lambda_1, \lambda_2), V_{C_2}(\lambda_1, \lambda_2), V_{D_1}(\lambda_1)$ and $V_{D_2}(\lambda_2)$ suggests the following simple threshold policy for channel selecting and change detection as illustrated in Figure 4.

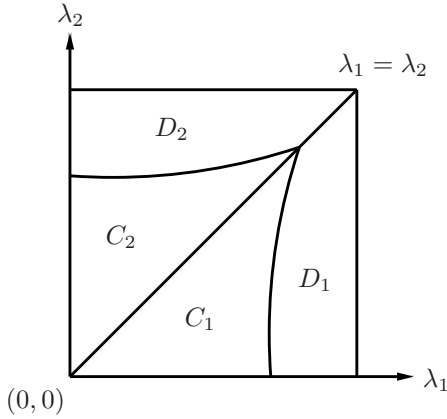


Fig. 4. A threshold structure of the resulting POMDP.

Specifically, under the proposed policy $\hat{\pi}$, when both λ_1 and λ_2 are below their corresponding detection thresholds $\eta(\lambda_2)$ and $\eta(\lambda_1)$, the user continues taking observations on the channel with a larger belief value; Otherwise, the user declares that an idle state has been reached on the channel with a larger belief value.

Similarly, we can extend this simple threshold policy to the N -channel case.

Define that $\Lambda^{-i} = \{\lambda_1, \dots, \lambda_{i-1}, \lambda_{i+1}, \dots, \lambda_N\}$, then

$$\hat{\pi}_N(\Lambda) = \begin{cases} C_i, & i = \arg \max_{1 \leq j \leq N} \{\lambda_j\} \text{ \& } \lambda_i < \eta(\Lambda^{-i}) \\ D_i, & i = \arg \max_{1 \leq j \leq N} \{\lambda_j\} \text{ \& } \lambda_i \geq \eta(\Lambda^{-i}) \end{cases} \quad (10)$$

This low-complexity threshold policy $\hat{\pi}_N$ agrees with our intuition: continue on the channel where the prospect of

catching an opportunity on that channel is the best. Below, we show that this proposed policy converges to the optimal policy for the infinite-channel case developed in [10].

Let π_∞^* denotes the optimal switching and change detection policy presented in [10] where we considered an infinite number of homogeneous independent channels. The threshold structure is with respect to the posterior probability λ_t : the user should switch to a new channel when $\lambda_t \in [0, \eta_s)$, should continue observing the current channel when $\lambda_t \in [\eta_s, \eta_d)$, and should declare that the current channel is idle and start transmitting when $\lambda_t \in [\eta_d, 1]$, where η_s and η_d are, respectively, the switching and detection thresholds, and $\eta_s = \lambda_0$.

Theorem 2: $\hat{\pi}_N$ converges to the optimal switching and change detection policy π_∞^* for the infinite-channel case as the number N of channels increases.

V. A LOWER BOUND ON THE AVERAGE DETECTION TIME: A GENIE-AIDED SYSTEM

In this section, we provide a performance benchmark by considering the full sensing case where the user can obtain observations from all N channels simultaneously.

The value functions under full sensing are as follows:

$$\begin{aligned} V(\lambda_1, \dots, \lambda_N) &= \min\{V_C(\lambda_1, \dots, \lambda_N), V_{D_1}(\lambda_1), \dots, V_{D_N}(\lambda_N)\}, \\ V_C(\lambda_1, \dots, \lambda_N) &= 1 + \int_{x_1} \dots \int_{x_N} P(x_1, \lambda_1) \dots P(x_N, \lambda_N) \\ &\quad V(\mathcal{T}(\lambda_1|x_1), \dots, \mathcal{T}(\lambda_N|x_N)) dx_1 \dots dx_N, \\ V_{D_1}(\lambda_1) &= \gamma(1 - \lambda_1), \\ &\vdots \\ V_{D_N}(\lambda_N) &= \gamma(1 - \lambda_N). \end{aligned}$$

Theorem 3: When $p_B + p_I \leq 1$, the optimal policy π_{FS}^* is given in Figure 5. Specifically,

$$\pi_{FS}^*(\lambda_1, \lambda_2) = \begin{cases} D_1, & \text{if } \lambda_1 \geq \lambda_2, \lambda_1 > \eta(\lambda_2) \\ D_2, & \text{if } \lambda_1 < \lambda_2, \lambda_2 > \eta(\lambda_1) \\ C, & \text{Otherwise} \end{cases} \quad (11)$$

where $\eta(\lambda) : [0, 1] \rightarrow [0, 1]$ is the detection threshold which is monotonically increasing in λ .

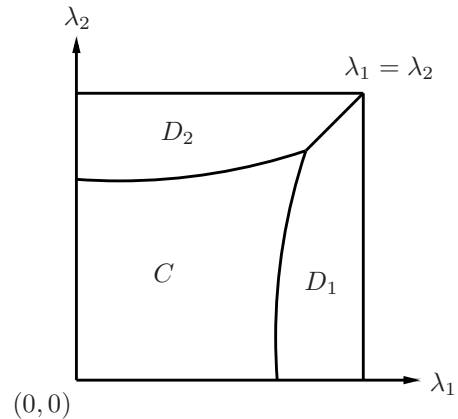


Fig. 5. The basic structure of the optimal policy.

Proof: Omitted due to the space limit. ■
 The extension to the N channel case is straightforward.

VI. SIMULATION EXAMPLES

In this section, we demonstrate the performance of the proposed low-complexity threshold policy. The primary signals are modeled as Gaussian signals in Gaussian noise, *i.e.*, $f_0(x)$ and $f_1(x)$ are both Gaussian distributions with zero mean and different variances. In all these examples, we set the detection threshold to $1 - \zeta$.

In Fig. 6, we compare the single-channel and multi-channel strategies for different on-off processes. Specifically, we increase both the average busy time m_B and the average idle time m_I while keeping the fraction λ_0 of idle time unchanged. We observe that the low-complexity threshold policy offers significant reduction in the detection delay over the single-channel strategy that employs the optimal change detection rule in a single channel. The performance improvement is dramatic when the average busy time is large. This is due to the channel selection strategy that avoids with large realizations of busy time. We also see that the detection performance improves with the number N of channels. When $N = 6$ the average detection time converges to that of the optimal policy π_∞^* for the infinite-channel case.

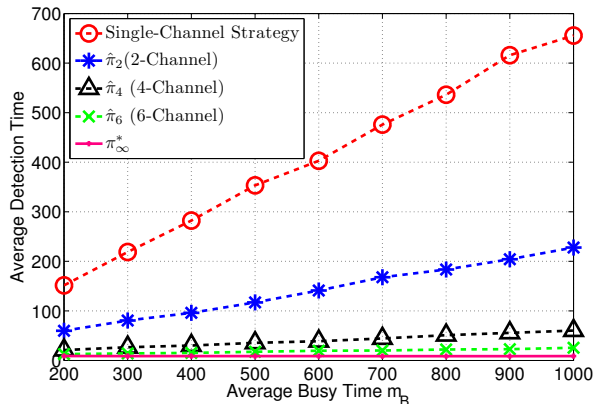


Fig. 6. Average detection time vs. the average busy time m_B . ($\lambda_0 = 0.3$, $m_B = 200 : 100 : 1000$, $p_B = 1/m_B$, $p_I = p_B \frac{1-\lambda_0}{\lambda_0}$, $\text{SNR} = 10$, $\zeta = 0.1$)

Shown in Fig. 7 is the expected time to catch an opportunity as a function of SNR, the signal to noise ratio of the observation model. Compared to the strategy that stays in a single channel, the proposed low-complexity threshold policy offers significant improvement for a large range of SNR. And the performance of the proposed policy $\hat{\pi}$ is close to the benchmark given by the full sensing case.

Fig. 8 illustrates the average detection time as a function of the number of channels, where we observe a significant reduction in the average detection time offered by the proposed policy over the single-channel strategy, especially when the

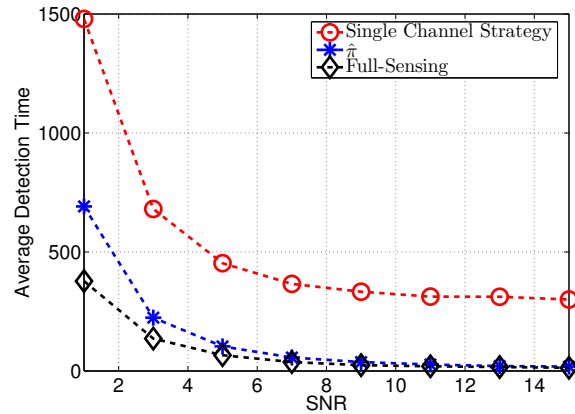


Fig. 7. Average detection time vs. SNR. ($p_I = 0.002$, $m_B = 600$, $p_B = 1/m_B$, $\zeta = 0.1$, $\text{SNR} = 1 : 2 : 15$, $N = 4$)

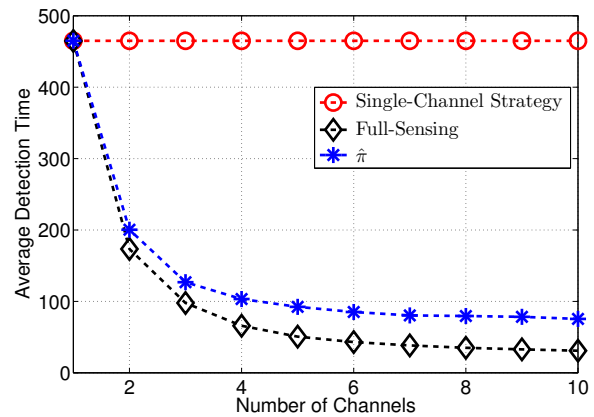


Fig. 8. Average detection time vs. the number of channels. ($p_I = 0.002$, $m_B = 600$, $p_B = 1/m_B$, $\zeta = 0.1$, $\text{SNR} = 5$, $N = 1 : 10$)

number of channels is large. The performance of the proposed policy is close to that of the full sensing case. The difference, however, increases with N since the advantage of full sensing becomes more significant when N is large.

VII. CONCLUSION

In this paper, we consider the quickest change detection problem in multiple on-off processes where switching back to a previously visited on-off process is allowed and measurements obtained during previous visits to this process are taken into consideration. A Bayesian formulation has been obtained within a decision-theoretic framework. The basic structure of the optimal decision rule is established. Based on these structures, we proposed a low-complexity threshold policy for the joint design of the channel selecting rule and the detection rule and demonstrated its near optimal performance.

REFERENCES

- [1] W. Shewhart, *Economic Control of Quality of Manufactured Product*. D. Van Nostrand Company, 1931.
- [2] A. N. Shiriyayev, "The problem of quickest detection of a violation of stationary behavior," *Dokl. Akad. Nauk SSSR*, vol. 138, pp. 1039–1042, 1961.
- [3] A. N. Shiriyayev, "On optimum methods in quickest detection problems," *Theory Prob. Applications*, vol. 8, pp. 22–46, June 1963.
- [4] A. G. Tartakovsky and V. V. Veeravalli, "General asymptotic bayesian theory of quickest change detection," *Theory Prob. Applications*, vol. 49, no. 3, pp. 458–497, 2005.
- [5] A. A. Borovkov, "Asymptotically optimal solutions in the change-point problem," *Theory Prob. Applications*, vol. 43, pp. 539–561, Oct. 1997.
- [6] G. Lorden, "Procedures for reacting to a change in distribution," *Annals Mathematical Statistics*, vol. 42, pp. 1897–1908, 1971.
- [7] E. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, pp. 100–115, 1954.
- [8] Q. Zhao and B. Sadler, "A Survey of Dynamic Spectrum Access," *IEEE Signal Processing magazine*, vol. 24, pp. 79–89, May 2007.
- [9] Q. Zhao and J. Ye, "When to Quit for A New Job: Quickest Detection of Spectrum Opportunities in Multiple Channels," in *Proc. of IEEE Military Communication Conference (MILCOM)*, Nov. 2008.
- [10] Q. Zhao and J. Ye, "Quickest change detection in multiple on-off processes," to appear in *ICASSP'09*, Apr. 2009.
- [11] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Applications*. Prentice Hall, Englewood Cliffs, 1993.
- [12] S. Kent, "On the trial of intrusions into information systems," *IEEE Spectrum*, pp. 52–56, 2000.