Optimal Search in a Multi-component Hypothesis Testing

Baha Alzalg

Department of Mathematics, The University of Jordan, Amman 11942, Jordan
Email: b.alzalg@ju.edu.jo

***Abstract*—We formulate and solve the component search problem under two different hypotheses: the exclusive hypothesis in which there is one and only one abnormal component, and the independent hypothesis in which there can be any number of abnormal components, but the abnormality is independently associated with the components. Under the exclusive hypothesis, we show that the optimal solution is given by a series of independent sequential probability ratio tests. Under the independent hypothesis, the threshold structure of the optimal decision rules is established.**

I. Introduction

We consider quickest localization of anomaly in a cyber system under noisy measurements. Consider that an intrusion to a subnet has been detected. The objective here is to locate the infected component(s) in the subnet as quickly and as reliably as possible. For example, a path may be detected as being in an abnormal state, the next step is to locate which links or routers in this path have been compromised. Under resource constraints, only a subset of components can be tested at each time, and the anomaly manifests itself not deterministically but rather through a different distribution of the measurements (for example, the delay over an infected link/router may exhibit a different distribution from that of a healthy one). As a consequence, reliable detection of the state of a component requires an accumulation of measurements to reveal the underlying distribution. The tradeoff here is thus between detection delay and detection accuracy: the state of a component can be more reliably detected by taking more measurements, but at the price of increasing the delay of localizing all infected components. The resource constraint adds another dimension to the problem: when to stop testing the current component and switch to a new set of components to locate the anomaly.

The following search problem is considered: K possible search components are given to be searched. With each component $k \in \{1, 2, \ldots, K\}$, we associate a hypothesis $H^{(k)}$ claiming that the component k is abnormal (or physically it contains an object of interest). The hypothesis $H^{(k)}$ is assumed to take one of two values: 0 or 1. $H^{(k)} = 0$ corresponds to the event that $H^{(k)}$ is false, and $H^{(k)} = 1$ corresponds to the event that $H^{(k)}$ is true.

In general, we can make simultaneous observations on $M \leq K$ components. The focus in this paper is on the single observation case (i.e., when M=1). When the k^{th} component is observed at a stage j and an action is taken, a measurement value $y_j^{(k)}$ is generated independently at the corresponding stage as a random variable with the following probability density:

$$\text{(pdf)} \begin{cases} f_0(y_j^{(k)}), & \text{if } \ H^{(k)} = 0; \\ f_1(y_j^{(k)}), & \text{if } \ H^{(k)} = 1. \end{cases}$$

Our goal is to determine the status of all components when we terminate the sensing by indicating whether each component is healthy or abnormal. This determination needs to be made by making one of the following two possible assumptions [2]:

Assumption 1 (The independent hypothesis): The events that the hypothesis are true are independent across hypothesis.

Assumption 2 (The exclusive hypothesis): One and only one hypothesis can be true.

We devote §2 to solve the problem under Assumption 1. By using tools from the optimal stopping theory, we minimize the average delay of termination time subject to the constraint that the error probabilities are less than preset thresholds. Our procedure in §2 closely follows that of Caromi, et al. [1, Subsection III(B)] for solving the non delay-limited scenario for multiband cognitive radio systems. §3 is devoted to study the problem under Assumption 2, and the basic structure of the optimal policy are established for the cases when we have two and three components.

II. The component search problem for independent hypothesis

In this section we solve our problem for independent hypothesis. We start by defining the following decision rules:

- $\tau \triangleq$ the *termination rule* that we use to decide whether or not to terminate the sensing,

- $\phi_j \triangleq$ the *component selection rule* that we use to select another component to make another observation if we decide to continue sensing at a stage j,
- $\boldsymbol{\delta}_j = (\delta_j^{(1)}, \delta_j^{(2)}, \ldots, \delta_j^{(K)}) \triangleq$ the *terminal decision rule* that we use to determine the status of all components if we decide to terminate at a stage j, in such a way that each $\delta_j^{(k)}$ takes values in $\{0,1\}$ with 0 indicating that component k is healthy and 1 indicating that component k is abnormal.

We also define $\boldsymbol{\delta} \triangleq \{\boldsymbol{\delta}_j\}_{j=1}^{\infty}$ to be the sequence of decision rules used, and $\boldsymbol{\phi} \triangleq \{\phi_j\}_{j=1}^{\infty}$ to be the sequence component selection functions. When we terminate the sensing, we have the following two error probabilities for component k: the *false-alarm probability* $P_{\text{FA}}^{(k)} \triangleq \Pr[\delta_\tau^{(k)} = 1 | H^{(k)} = 0]$, and the *missdetection probability* $P_{\text{MD}}^{(k)} \triangleq \Pr[\delta_\tau^{(k)} = 0 | H^{(k)} = 1]$.

Our objective is to minimize the average delay $\mathbb{E}[\tau]$ subject to the constraint that the error probabilities over component k are less than preset thresholds, say $\alpha^{(k)}$ and $\beta^{(k)}$ for $k = 1, 2, \ldots, K$. More specifically, we are interested in the problem

$$\begin{array}{ll} \inf_{\tau, \boldsymbol{\delta}, \boldsymbol{\phi}} & \mathbb{E}[\tau] \\ \text{s.t.} & P_{\text{FA}}^{(k)} \le \alpha^{(k)}, k = 1, 2, \ldots, K, \\ & P_{\text{MD}}^{(k)} \le \beta^{(k)}, k = 1, 2, \ldots, K. \end{array} \quad (1)$$

By following the same argument as in [6, §4.3], the solution of (1) can be obtained once we can find the solution of the problem

$$\inf_{\tau, \boldsymbol{\delta}, \boldsymbol{\phi}} \quad \left[\mathbb{E}[\tau] + \sum_{k=1}^{K} \left(c_0 (1 - \pi_0^{(k)}) P_{\text{FA}}^{(k)} + c_1 \pi_0^{(k)} P_{\text{MD}}^{(k)} \right) \right], \quad (2)$$

where $\pi_0^{(k)} \triangleq \Pr[H^{(k)} = 1]$ is the priori probability that component k is abnormal, $\mathbb{E}$ is the expectation under the probability measure $\Big((1-\pi_0^{(1)}) f_0 + \pi_0^{(1)} f_1, \ldots, (1-\pi_0^{(K)}) f_0 + \pi_0^{(K)} f_1\Big)$, c_0 is the cost of a false-alarm event happening over each component, and c_1 is the cost of a missdetection event happening over each component.

We now define $\boldsymbol{\pi}_j \triangleq (\pi_j^{(1)}, \pi_j^{(2)}, \ldots, \pi_j^{(K)})$, where $\pi_j^{(k)}$ is the posterior probability that component k is abnormal after collecting observations up to stage j.

Under the independent hypothesis, if we select component k to sense at stage j (i.e., $\phi_j = k$), then from Bayes' rule, the posterior probability of component k being abnormal after collecting an observation $y_j^{(k)}$ can be updated as follows:

$$\pi_j^{(k)} = \frac{\pi_{j-1}^{(k)} f_1(y_j^{(k)})}{\pi_{j-1}^{(k)} f_1(y_j^{(k)}) + (1 - \pi_{j-1}^{(k)}) f_0(y_j^{(k)})}. \quad (3)$$

It is also clear $\pi_j^{(k)} = \pi_{j-1}^{(k)}$ for component k that is not selected at stage j. Now we will show that $\boldsymbol{\pi}_j$ is a sufficient statistic, that is, at stage j, we can make our decision rules τ, $\boldsymbol{\delta}$, and $\boldsymbol{\phi}$ solely based on $\boldsymbol{\pi}_j$.

For any given termination rule $\boldsymbol{\tau}$ and component selection rules $\boldsymbol{\phi}$, by following argument as in [5], one can show the optimality of the following simple terminal decision rule:

$$\delta_\tau^{(k)} = \begin{cases} 1, & \text{if } c_1 \pi_\tau^{(k)} \ge c_0 (1 - \pi_\tau^{(k)}); \\ 0, & \text{if } c_1 \pi_\tau^{(k)} < c_0 (1 - \pi_\tau^{(k)}). \end{cases}$$

That is, we declare that component k is abnormal if the cost of a missdetection event is larger than that of a false alarm and vice visa. This shows that the terminal decision rules $\boldsymbol{\delta}$ can be made only based on $\boldsymbol{\pi}_j$, and therefore Problem (2) can be converted into the problem

$$\inf_{\tau, \boldsymbol{\phi}} \quad \mathbb{E} \left[\tau + \sum_{k=1}^{K} \min \left\{ c_0 (1 - \pi_\tau^{(k)}), c_1 \pi_\tau^{(k)} \right\} \right]. \quad (4)$$

We will now take the advantage of the theory of optimal stopping rules [4] to solve the resulting problem. For any stopping time τ, let $\tau^{(k)}$ be the amount of time we spend on detecting component k. So, Problem (4) can be reformulated as the problem

$$\inf_{\{\tau^{(k)}\}_{k=1}^K} \mathbb{E} \left[\sum_{k=1}^{K} \tau^{(k)} + \sum_{k=1}^{K} \min \left\{ c_0 (1 - \pi_{\tau^{(k)}}^{(k)}), c_1 \pi_{\tau^{(k)}}^{(k)} \right\} \right],$$

or equivalently,

$$\inf_{\{\tau^{(k)}\}_{k=1}^K} \sum_{k=1}^{K} \mathbb{E} \left[\tau^{(k)} + \min \left\{ c_0 (1 - \pi_{\tau^{(k)}}^{(k)}), c_1 \pi_{\tau^{(k)}}^{(k)} \right\} \right]. \quad (5)$$

It is immediately follows that the objective function in (5) is regardless of sensing ordering $\boldsymbol{\phi}$, and, in particular, it is related to only the total amount of detection time. That is, once the quantity

$$\mathbb{E} \left[\tau^{(k)} + \min \left\{ c_0 (1 - \pi_{\tau^{(k)}}^{(k)}), c_1 \pi_{\tau^{(k)}}^{(k)} \right\} \right] \quad (6)$$

is minimized for each component $k = 1, 2, \ldots, K$, the sum is also minimized. The key observation about this situation is that these K optimization problems are independent of each other, and, as a result, we can *independently* minimize each term of the sum.

For each k, the procedure that minimizes (6) is the well-known SPRT algorithm [4]. More explicitly, this solution procedure is parametrized by two parameters $U^{(k)}$ and $L^{(k)}$ and is performed as follows: after taking each sample from component k we update the posterior probability $\pi_j^{(k)}$, then, as a rule, we stay on component k and take more samples if $\pi_j^{(k)} \in (L^{(k)}, U^{(k)})$, we stop sampling on component k, and claim that component k is abnormal if $\pi_j^{(k)} \geq U^{(k)}$, and we stop sampling on component k, and claim that component k is healthy if $\pi_j^{(k)} \leq L^{(k)}$. In fact, as Problem (6) does not depend on the component selection rules $\boldsymbol{\phi}$, we can start sensing from component 1, and then we switch to component 2 once we finish sensing component 1, etc. We terminate the whole sensing process once we finish sensing component K. The complete search procedure is summarized in Algorithm 1.

Algorithm 1: THE SPRT ALGORITHM FOR SOLVING THE COMPONENT SEARCH PROBLEM (5)

Require: f_0, f_1, c_0, c_1, and $\boldsymbol{\pi}_0$
  for $k = 1, 2, \ldots, K$ **do**
    compute parameters $L^{(k)}$ and $U^{(k)}$,
    take each sample from comp. k and update the posterior probability $\pi_j^{(k)}$ using (3)
    while $\pi_j^{(k)} \in (L^{(k)}, U^{(k)})$ **do**
      take more samples from component k and update the posterior probability $\pi_j^{(k)}$ using (3)
    end while
    if $\pi_j^{(k)} \geq U^{(k)}$ **then**
      claim that component k is abnormal
    end if
    if $\pi_j^{(k)} \leq L^{(k)}$ **then**
      claim that component k is healthy
    end if
  end for

III. THE COMPONENT SEARCH PROBLEM FOR EXCLUSIVE HYPOTHESIS

In this section we give basic structure of the optimal policy of our problem for exclusive hypothesis. Under the exclusive hypothesis, if we select component k to sense at stage j (i.e., $\phi_j = k$), then from Bayes' rule, the posterior probability of component k being abnormal after collecting an observation $y_j^{(k)}$ can be updated as follows:

$$\mathcal{T}(\pi_{j-1}^{(k)}|y_j^{(k)}) \triangleq \pi_j^{(k)} = \frac{\pi_{j-1}^{(k)} f_1^{(k)}(y_j^{(k)})}{\pi_{j-1}^{(k)} f_1^{(k)}(y_j^{(k)}) + (1-\pi_{j-1}^{(k)}) f_0^{(k)}(y_j^{(k)})}.$$

For component k that is not selected at time j, the posterior probability $\pi_j^{(k)}$ is also updated. In fact, if we select component $l \neq k$, then the measurement $y_j^{(l)}$ affects the posterior probability $\pi_{j-1}^{(k)}$ using the following equation

$$\mathcal{T}(\pi_{j-1}^{(k)}|y_j^{(l)}) \triangleq \pi_j^{(k)} = \frac{\pi_{j-1}^{(k)} f_0^{(k)}(y_j^{(l)})}{\pi_{j-1}^{(l)} f_1^{(l)}(y_j^{(l)}) + (1-\pi_{j-1}^{(l)}) f_0^{(l)}(y_j^{(l)})}.$$

We define $\boldsymbol{\pi}_j \triangleq (\pi_j^{(1)}, \pi_j^{(2)}, \ldots, \pi_j^{(K)})$. Note that under Assumption 2 we have that $\sum_{k=1}^{K} \pi_j^{(k)} = 1$. Let $V(\boldsymbol{\pi}_j)$ denote the minimal expected total remaining cost when the current information state is $\boldsymbol{\pi}_j$. Note that $V(\boldsymbol{\pi}_j)$ specifies the performance of the optimal policy starting from the information state $\boldsymbol{\pi}_j$. Let $V_{C_k}(\boldsymbol{\pi}_j)$ denote the minimal expected total remaining cost when we take action C_k at time j and follow the optimal policy. Let $V_{S_k}(\boldsymbol{\pi}_j)$ be similarly defined. We thus have

$$V(\bar{\boldsymbol{\pi}}_j) = \min\{V_{C_k}(\bar{\boldsymbol{\pi}}_j), V_{S_k}(\bar{\boldsymbol{\pi}}_j) : k = 1, 2, \ldots, K\}.$$

In general, want to choose an action from the following set

$$\{C_1, C_2, \ldots, C_K, S_1, S_2, \ldots, S_K\},$$

where C_k means that we continue taking samples from component k, and S_k means that we stop taking any more samples and declare that the k^{th} component is abnormal. In the following two subsections we consider the cases of $K = 2$ and $K = 3$. Extensions to an arbitrary $K > 3$ is straightforward.

A. The case K=2

Let us consider the case when $K = 2$. In this case we choose an action from the set $\{C_1, C_2, S_1, S_2\}$, and we have $\pi_j^{(1)} + \pi_j^{(2)} = 1$.

We first make the following additional assumption concerning the probability densities f_0 and f_1.

Assumption 3: There exists a constant $a \in \mathbb{R}$ such that $f_0(x) = f_1(a - x)$.

Observe that when the distributions are symmetric about constant means (such as Gaussian distributions), Assumption 3 is satisfied.

Note that in this case we have

$$\begin{aligned} V_{S_1}(\pi_j^{(1)}) &= (c_0 + c_1)(1 - \pi_j^{(1)}), \\ V_{S_2}(\pi_j^{(1)}) &= (c_0 + c_1)\pi_j^{(1)}. \end{aligned} \tag{7}$$

We then have the following lemma.

Lemma 1: $V_{S_1}(\pi_j^{(1)})$ is a linearly decreasing function of $\pi_j^{(1)}$, and $V_{S_2}(\pi_j^{(1)})$ is a linearly increasing function of $\pi_j^{(1)}$.

Proof. Obvious from (7). □

Note also that

$$V_{C_1}(\pi_j^{(1)}) = 1 + \int_y [\pi_j^{(1)} f_1(y) + (1 - \pi_j^{(1)}) f_0(y)] V(\mathcal{T}_1(\pi_j^{(1)}|y)) dy, \tag{8}$$

where

$$\mathcal{T}_1(\pi_j^{(1)}|y) = \frac{\pi_j^{(1)} f_1(y)}{\pi_j^{(1)} f_1(y) + (1 - \pi_j^{(1)}) f_0(y)},$$

and that

$$V_{C_2}(\pi_j^{(1)}) = 1 + \int_y [(1-\pi_j^{(1)})f_1(y) + \pi_j^{(1)} f_0(y)] V(\mathcal{T}_2(\pi_j^{(1)}|y))dy, \tag{9}$$

where

$$\mathcal{T}_2(\pi_j^{(1)}|y) = \frac{\pi_j^{(1)} f_0(y)}{(1-\pi_j^{(1)})f_1(y) + \pi_j^{(1)} f_0(y)}.$$

We also have the following lemma.

Lemma 2: $V_{C_1}(\pi_j^{(1)})$ and $V_{C_2}(\pi_j^{(1)})$ are concave and identical to each other.

Proof. We proof the lemma by considering the finite horizon problem of length T, i.e., we need to declare within K units of time. Let $V^T(\cdot), V_{C_1}^T(\cdot)$ and $V_{C_2}^T(\cdot)$ denote the corresponding value functions. We have that

$$V^0(\pi_j) = \min\{V_{S_1}(\pi_j^{(1)}), V_{S_2}(\pi_j^{(1)})\},$$

and

$$V^T(\pi_j) = \min\{V_{C_1}^T(\pi_j^{(1)}), V_{C_2}^T(\pi_j^{(1)}), V_{S_1}(\pi_j^{(1)}), V_{S_2}(\pi_j^{(1)})\},$$

for $T > 0$.

We prove that $V_{C_1}^T(\pi_j^{(1)})$ is concave and that it is equal to $V_{C_2}^T(\pi_j^{(1)})$ by induction. For the initial condition, we have

$$\begin{aligned} V_{C_1}^1(\pi_j^{(1)}) &= 1 + \int_y [\pi_j^{(1)} f_1(y) + (1-\pi_j^{(1)})f_0(y)] \\ &\qquad V^0(\mathcal{T}_1(\pi_j^{(1)}|y))dy \\ &= 1 + \int_y [\pi_j^{(1)} f_1(y) + (1-\pi_j^{(1)})f_0(y)] \\ &\qquad \min\{V_{S_1}(\mathcal{T}_1(\pi_j^{(1)}|y)), V_{S_2}(\mathcal{T}_1(\pi_j^{(1)}|y))\}dy \\ &= 1 + (c_0+c_1)\int_y [\pi_j^{(1)} f_1(y) + (1-\pi_j^{(1)}) \\ &\qquad f_0(y)] \min\{(1-\mathcal{T}_1(\pi_j^{(1)}|y)), \mathcal{T}_1(\pi_j^{(1)}|y)\}dy \\ &= 1 + (c_0+c_1) \\ &\qquad \int_y \min\{(1-\pi_j^{(1)})f_0(y), \pi_j^{(1)} f_1(y)\}dy. \end{aligned} \tag{10}$$

Thus, for any given y, the integrand (10) is the minimum of two linear functions of $\pi_j^{(1)}$. The concavity of $V_{C_1}^1(\pi_j^{(1)})$ thus follows. As a consequence, $V^1(\pi_j^{(1)})$ is the minimum of six linear functions. It is thus also concave. In addition, using Assumption 3, we also obtain

$$\begin{aligned} V_{C_2}^1(\pi_j^{(1)}) &= 1 + \int_y [(1-\pi_j^{(1)})f_1(y) + \pi_j^{(1)} f_0(y)] \\ &\qquad V^0(\mathcal{T}_2(\pi_j^{(1)}|y))dy \\ &= 1 + \int_y [(1-\pi_j^{(1)})f_1(y) + \pi_j^{(1)} f_0(y)] \\ &\qquad \min\{V_{S_1}(\mathcal{T}_2(\pi_j^{(1)}|y)), V_{S_2}(\mathcal{T}_2(\pi_j^{(1)}|y))\}dy \\ &= 1 + (c_0+c_1)\int_y [(1-\pi_j^{(1)})f_1(y) + \pi_j^{(1)} \\ &\qquad f_0(y)] \min\{(1-\mathcal{T}_2(\pi_j^{(1)}|y)), \mathcal{T}_2(\pi_j^{(1)}|y)\}dy \\ &= 1 + (c_0+c_1)\int_y \min\{(1-\pi_j^{(1)})f_1(y), \\ &\qquad \pi_j^{(1)} f_0(y)\}dy \\ &= 1 + (c_0+c_1)\int_y \min\{(1-\pi_j^{(1)})f_1(a-y), \\ &\qquad \pi_j^{(1)} f_0(a-y)\}dy \\ &= 1 + (c_0+c_1)\int_y \min\{(1-\pi_j^{(1)})f_0(y), \\ &\qquad \pi_j^{(1)} f_1(y)\}dy \\ &= V_{C_1}^1(\pi_j^{(1)}). \end{aligned}$$

Now assume that $V_{C_1}^T(\pi_j^{(1)}) = V_{C_2}^T(\pi_j^{(1)})$ and $V^T(\pi_j^{(1)})$ are concave. Then $V^T(\pi_j^{(1)})$ can be written as the minimum of finitely many (potentially uncountable) linear functions of $\pi_j^{(1)}$. Let us, by abuse of notation, index these linear functions by $i \in \mathbb{R}$, i.e., there exist $a_i, b_i \in \mathbb{R}$ such that

$$V^T(\pi_j^{(1)}) = \min_{i\in\mathbb{R}}\{a_i + b_i\pi_j^{(1)}\}.$$

We thus have

$$\begin{aligned} V_{C_1}^{T+1}(\pi_j^{(1)}) &= 1 + \int_y [\pi_j^{(1)} f_1(y) + (1-\pi_j^{(1)})f_0(y)] \\ &\qquad V^T(\mathcal{T}_1(\pi_j^{(1)}|y))dy \\ &= 1 + \int_y [\pi_j^{(1)} f_1(y) + (1-\pi_j^{(1)})f_0(y)] \\ &\qquad \min_{i\in\mathbb{R}}\{a_i + b_i\mathcal{T}_1(\pi_j^{(1)}|y)\}dy \\ &= 1 + \int_y \min_{i\in\mathbb{R}}\{a_i[\pi_j^{(1)} f_1(y) + (1-\pi_j^{(1)}) \\ &\qquad f_0(y)] + b_i\pi_j^{(1)} f_1(y)\}dy \\ &= 1 + \int_y \min_{i\in\mathbb{R}}\{[a_i f_0(y)] + [(a_i+b_i) \\ &\qquad f_1(y) - a_i f_0(y)]\pi_j^{(1)}\}dy. \end{aligned} \tag{11}$$

We can see that for any given y, the integrand in (11) is the minimum of linear functions of $\pi_j^{(1)}$, thus concave in $\pi_j^{(1)}$. It follows that $V_{C_1}^{T+1}(\pi_j^{(1)})$ is concave. In addition, using Assumption 3, we also obtain

$$
\begin{aligned}
V_{C_2}^{T+1}(\pi_j^{(1)}) &= 1+\int_y [(1-\pi_j^{(1)})f_1(y)+\pi_j^{(1)}f_0(y)] \\
&\qquad V^T(\mathcal{T}_2(\pi_j^{(1)}|y))dy \\
&= 1+\int_y [(1-\pi_j^{(1)})f_1(y)+\pi_j^{(1)}f_0(y)] \\
&\qquad \min_{i\in\mathbb{R}}\{a_i+b_i\mathcal{T}_2(\pi_j^{(1)}|y)\}dy \\
&= 1+\int_y \min_{i\in\mathbb{R}}\{a_i[(1-\pi_j^{(1)})f_1(y) \\
&\qquad +\pi_j^{(1)}f_0(y)]+b_i\pi_j^{(1)}f_0(y)\}dy \\
&= 1+\int_y \min_{i\in\mathbb{R}}\{a_i[(1-\pi_j^{(1)})f_1(a-y)+ \\
&\qquad \pi_j^{(1)}f_0(a-y)]+b_i\pi_j^{(1)}f_0(a-y)\}dy \\
&= 1+\int_y \min_{i\in\mathbb{R}}\{a_i[(1-\pi_j^{(1)})f_0(y) \\
&\qquad +\pi_j^{(1)}f_1(y)]+b_i\pi_j^{(1)}f_1(y)\}dy \\
&= 1+\int_y [\pi_j^{(1)}f_1(y)+(1-\pi_j^{(1)})f_0(y)] \\
&\qquad V^T(\mathcal{T}_1(\pi_j^{(1)}|y))dy \\
&= V_{C_1}^{T+1}(\pi_j^{(1)}).
\end{aligned}
$$

The proof is complete. □

One can easily obtain from (7), (8) and (9) the following observations:

- $V_{S_1}(1)=V_{S_1}(0)=0$, and $V_{S_1}(\pi_j^{(1)})=V_{S_2}(\pi_j^{(1)})$ if and only if $\pi_j^{(1)}=1/2$.
- $V_{C_k}(0)=V_{C_k}(1)=1$ for $k=1,2$.

Based on the above observations and Lemmas 1 and 2, we obtain the following basic structure of the optimal policy.

Theorem 1: The quickest detection of the two component search problem for exclusive hypothesis is given by two thresholds $\eta\in(0,1/2]$ and $1-\eta$: stop taking samples and declare that the 2nd component is abnormal if $\pi_j^{(1)}<\eta$, continue taking samples on either the 1st or the 2nd component if $\eta\le\pi_j^{(1)}\le 1-\eta$, and stop taking samples and declare that the 1st component is abnormal if $\pi_j^{(1)}>1-\eta$. That is, the optimal action $a(\pi_j^{(1)})$ under the belief value $\pi_j^{(1)}$ is in the following form:

$$
a(\pi_j^{(1)})=\begin{cases} S_2, & \text{if } \pi_j^{(1)}<\eta; \\ C_1 \text{ or } C_2, & \text{if } \eta\le\pi_j^{(1)}\le 1-\eta; \\ S_1, & \text{if } \pi_j^{(1)}>1-\eta. \end{cases}
$$

Fig. 1 illustrates the basic structure of the optimal policy given in Theorem 1.

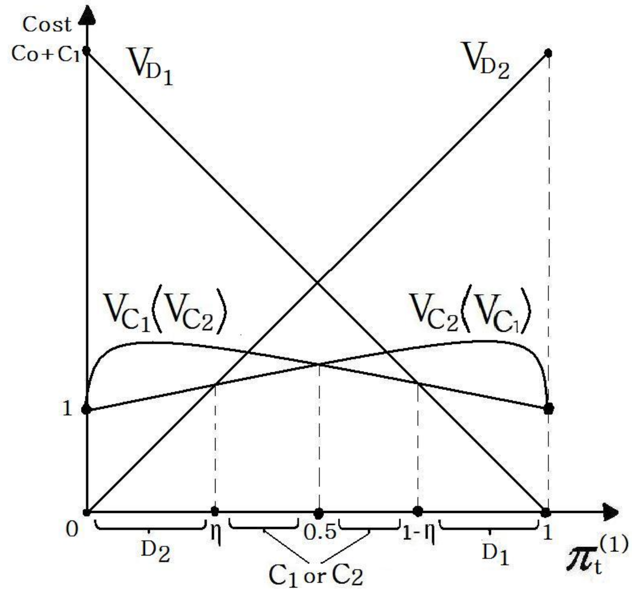

Fig. 1. The structure of the optimal policy under the exclusive model ($K=2$).

B. The case K=3

Let us consider the case when $K=3$. In this case we choose an action from the set $\{C_1,C_2,C_3,S_1,S_2,S_3\}$, and we have that $\pi_j^{(1)}+\pi_j^{(2)}+\pi_j^{(3)}=1$. Note that

$$
\begin{aligned}
V_{S_1}(\pi_j^{(1)},\pi_j^{(2)}) &= (c_0+c_1)(1-\pi_j^{(1)}), \\
V_{S_2}(\pi_j^{(1)},\pi_j^{(2)}) &= (c_0+c_1)(1-\pi_j^{(2)}), \\
V_{S_3}(\pi_j^{(1)},\pi_j^{(2)}) &= (c_0+c_1)(\pi_j^{(1)}+\pi_j^{(2)}).
\end{aligned} \tag{12}
$$

We have the following lemma.

Lemma 3: $V_{S_1}(\pi_j^{(1)},\pi_j^{(2)})$ and $V_{S_2}(\pi_j^{(1)},\pi_j^{(2)})$ are linearly decreasing functions of $(\pi_j^{(1)},\pi_j^{(1)})$, and $V_{S_3}(\pi_j^{(1)},\pi_j^{(2)})$ is a linearly increasing function of $(\pi_j^{(1)},\pi_j^{(2)})$.

Proof. Obvious from (12). □

Note also that

$$
\begin{aligned}
V_{C_1}(\pi_j^{(1)},\pi_j^{(2)}) &= 1+\textstyle\int_y [\pi_j^{(1)}f_1(y)+(1-\pi_j^{(1)})f_0(y)] \\
&\qquad V(\mathcal{T}_1(\pi_j^{(1)}|y),\mathcal{T}_1(\pi_j^{(2)}|y))dy,
\end{aligned} \tag{13}
$$

where

$$
\mathcal{T}_1(\pi_j^{(1)}|y)=\frac{\pi_j^{(1)}f_1(y)}{\pi_j^{(1)}f_1(y)+(1-\pi_j^{(1)})f_0(y)}
$$

and

$$
\mathcal{T}_1(\pi_j^{(2)}|y)=\frac{\pi_j^{(2)}f_0(y)}{\pi_j^{(1)}f_1(y)+(1-\pi_j^{(1)})f_0(y)};
$$

$$V_{C_2}(\pi_j^{(1)},\pi_j^{(2)}) = 1+\int_y[\pi_j^{(2)}f_1(y)+(1-\pi_j^{(2)})f_0(y)] V(\mathcal{T}_2(\pi_j^{(1)}|y),\mathcal{T}_2(\pi_j^{(2)}|y))dy, \quad (14)$$

where

$$\mathcal{T}_2(\pi_j^{(1)}|y)=\frac{\pi_j^{(1)}f_0(y)}{\pi_j^{(2)}f_1(y)+(1-\pi_j^{(2)})f_0(y)}$$

and

$$\mathcal{T}_2(\pi_j^{(2)}|y)=\frac{\pi_j^{(2)}f_1(y)}{\pi_j^{(2)}f_1(y)+(1-\pi_j^{(2)})f_0(y)};$$

and that

$$V_{C_3}(\pi_j^{(1)},\pi_j^{(2)}) = 1+\int_y[(1-\pi_j^{(1)}-\pi_j^{(2)})f_1(y) +(\pi_j^{(1)}+\pi_j^{(2)})f_0(y)]V(\mathcal{T}_3(\pi_j^{(1)}|y), \mathcal{T}_3(\pi_j^{(2)}|y))dy, \quad (15)$$

where

$$\mathcal{T}_3(\pi_j^{(1)}|y)=\frac{\pi_j^{(1)}f_0(y)}{(1-\pi_j^{(1)}-\pi_j^{(2)})f_1(y)+(\pi_j^{(1)}+\pi_j^{(2)})f_0(y)}$$

and

$$\mathcal{T}_3(\pi_j^{(2)}|y)=\frac{\pi_j^{(2)}f_0(y)}{(1-\pi_j^{(1)}-\pi_j^{(2)})f_1(y)+(\pi_j^{(1)}+\pi_j^{(2)})f_0(y)}.$$

We also have the following lemma.

Lemma 4: $V_{C_k}(\pi_j^{(1)},\pi_j^{(2)})$ is concave for $k=1,2,3$.

Proof. The proof is a trivial extension of the proof of Lemma 2 and therefore omitted. □

From (12), (13), (14) and (15), one can easily obtain the following observations:

- $V_{S_1}(1,0) = V_{S_2}(0,1) = V_{S_3}(0,0) = 0$, and $V_{S_1}(\pi_j^{(1)},\pi_j^{(2)}) = V_{S_2}(\pi_j^{(1)},\pi_j^{(2)}) = V_{S_3}(\pi_j^{(1)},\pi_j^{(2)})$ if and only if $(\pi_j^{(1)},\pi_j^{(2)})=(1/3,1/3)$.
- The functions $V_{S_1}(\cdot)$ and $V_{S_2}(\cdot)$ are symmetric with respect to the plane $\pi_j^{(1)}=\pi_j^{(2)}$. That is,
$$V_{S_1}(\pi_j^{(1)},\pi_j^{(2)})=V_{S_2}(\pi_j^{(2)},\pi_j^{(1)}).$$
- The functions $V_{S_1}(\cdot)$ and $V_{S_3}(\cdot)$ are symmetric with respect to the plane $2\pi_j^{(1)}+\pi_j^{(2)}=1$. That is,
$$V_{S_3}(\pi_j^{(1)},\pi_j^{(2)})=V_{S_1}(1-\pi_j^{(1)}-\pi_j^{(2)},\pi_j^{(2)}).$$
- The functions $V_{S_2}(\cdot)$ and $V_{S_3}(\cdot)$ are symmetric with respect to the plane $\pi_j^{(1)}+2\pi_j^{(2)}=1$. That is,
$$V_{S_3}(\pi_j^{(1)},\pi_j^{(2)})=V_{S_2}(\pi_j^{(1)},1-\pi_j^{(1)}-\pi_j^{(2)}).$$
- $V_{C_k}(0,0) = V_{C_k}(0,1) = V_{C_k}(1,0) = 1$ for all $k=1,2,3$, and $V_{C_1}(\pi_j^{(1)},\pi_j^{(2)}) = V_{C_2}(\pi_j^{(1)},\pi_j^{(2)}) = V_{C_3}(\pi_j^{(1)},\pi_j^{(2)})$ if and only if $(\pi_j^{(1)},\pi_j^{(2)}) = (1/3,1/3)$.
- The functions $V_{C_1}(\cdot)$ and $V_{C_2}(\cdot)$ are symmetric with respect to the plane $\pi_j^{(1)}=\pi_j^{(2)}$. That is,
$$V_{C_1}(\pi_j^{(1)},\pi_j^{(2)})=V_{C_2}(\pi_j^{(2)},\pi_j^{(1)}).$$
- The functions $V_{C_1}(\cdot)$ and $V_{C_3}(\cdot)$ are symmetric with respect to the plane $2\pi_j^{(1)}+\pi_j^{(2)}=1$. That is,
$$V_{C_3}(\pi_j^{(1)},\pi_j^{(2)})=V_{C_1}(1-\pi_j^{(1)}-\pi_j^{(2)},\pi_j^{(2)}).$$
- The functions $V_{C_2}(\cdot)$ and $V_{C_3}(\cdot)$ are symmetric with respect to the plane $\pi_j^{(1)}+2\pi_j^{(2)}=1$. That is,
$$V_{C_3}(\pi_j^{(1)},\pi_j^{(2)})=V_{C_2}(\pi_j^{(1)},1-\pi_j^{(1)}-\pi_j^{(2)}).$$

Based on the above observations and Lemmas 3 and 4, we can easily obtain the following basic structure of the optimal policy.

Theorem 2: The quickest detection of the three component search problem for exclusive hypothesis is given by the following four rules: stop taking samples and declare that the 1st component is abnormal if $\pi_j^{(1)}+2\pi_j^{(2)}\le 1, 2\pi_j^{(1)}+\pi_j^{(2)}\le 1$, and $\pi_j^{(1)}\ge\eta(\pi_j^{(2)})$, stop taking samples and declare that the 2nd component is abnormal if $\pi_j^{(1)}+2\pi_j^{(2)}>1, \pi_j^{(1)}\le\pi_j^{(2)}$, and $\pi_j^{(2)}\ge\eta(\pi_j^{(1)})$, stop taking samples and declare that the 3rd component is abnormal if $2\pi_j^{(1)}+\pi_j^{(2)}>1, \pi_j^{(1)}>\pi_j^{(2)}$, and $\pi_j^{(1)},\pi_j^{(2)}\le\eta(1-\pi_j^{(1)}-\pi_j^{(2)})$, and continue taking samples on either the 1st, the 2nd or the 3rd component if otherwise. That is, the optimal action $a(\pi_j^{(1)},\pi_j^{(2)})$ under the belief vector $(\pi_j^{(1)},\pi_j^{(2)})$ is given in (16), where $\eta(\pi_j):[0,1]\longrightarrow[0,1]$ is the detection threshold which is concave in π_j. Furthermore, the action $a(\pi_j^{(1)},\pi_j^{(2)})$ is symmetric with respect to the lines $\pi_j^{(1)}=\pi_j^{(2)}, \pi_j^{(1)}+2\pi_j^{(2)}=1$, and $2\pi_j^{(1)}+\pi_j^{(2)}=1$, as follows:

$$\begin{aligned} a(\pi_j^{(1)},\pi_j^{(2)})=C_1 &\iff a(\pi_j^{(2)},\pi_j^{(1)})=C_2,\\ a(\pi_j^{(1)},\pi_j^{(2)})=C_3 &\iff a(1-\pi_j^{(1)}-\pi_j^{(2)},\pi_j^{(1)})=C_1,\\ a(\pi_j^{(1)},\pi_j^{(2)})=C_3 &\iff a(\pi_j^{(1)},1-\pi_j^{(1)}-\pi_j^{(2)})=C_2.\end{aligned}$$

Fig. 2 illustrates the basic structure of the optimal policy given in Theorem 2.

IV. Conclusion

In this paper, we have formulated and solved the component search problem under two different assumptions: When there is one and only one abnormal component (the exclusive hypothesis), and when there can be any number of abnormal components, but the abnormality

$$a(\pi_j^{(1)},\pi_j^{(2)}) = \begin{cases} S_1, & \text{if } \pi_j^{(1)}+2\pi_j^{(2)} \leq 1, 2\pi_j^{(1)}+\pi_j^{(2)} \leq 1, \text{ and } \pi_j^{(1)} \geq \eta(\pi_j^{(2)}); \\ S_2, & \text{if } \pi_j^{(1)}+2\pi_j^{(2)} > 1, \pi_j^{(1)} \leq \pi_j^{(2)}, \text{ and } \pi_j^{(2)} \geq \eta(\pi_j^{(1)}); \\ S_3, & \text{if } 2\pi_j^{(1)}+\pi_j^{(2)} > 1, \pi_j^{(1)} > \pi_j^{(2)}, \text{ and } \pi_j^{(1)},\pi_j^{(2)} \leq \eta(1-\pi_j^{(1)}-\pi_j^{(2)}); \\ C_1, C_2, \text{ or } C_3, & \text{otherwise.} \end{cases} \quad (16)$$

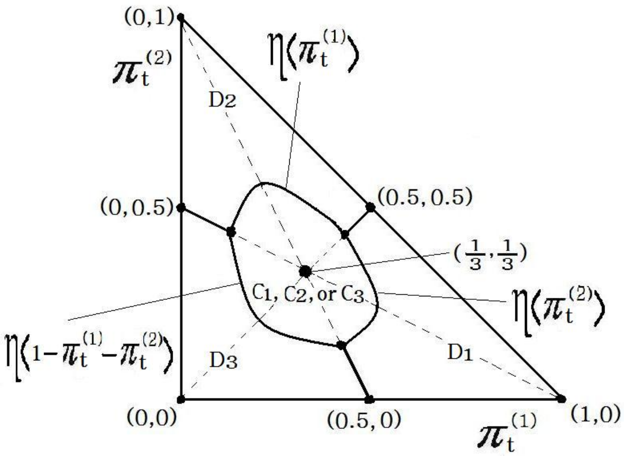

Fig. 2. The structure of the optimal policy under the exclusive model ($K = 3$).

is independently associated with the components (the independent hypothesis). For the exclusive hypothesis, using tools from the optimal stopping theory, we have developed an optimal sensing algorithm that minimizes the average delay of termination time subject to the constraint that the error probabilities (the false-alarm probability and the missdetection probability) over are less than preset thresholds. this development is analogue to the results of Caromi, et al. [1, Subsection III(B)] for solving the non delay-limited scenario for multiband cognitive radio systems. For the independent hypothesis, we established the simple threshold structure of the optimal decision rule for the case when we have two components and the basic structure of the optimal decision rule for the case when we have three components.

REFERENCES

[1] R. Caromi, Y. Xin and L. Lai. Fast multiband spectrum scanning for cognitive radio systems. *IEEE Trans. Communications*. To appear. The paper is available online at: users.wpi.edu/[tilde]llai/spectrumsurveyv6.pdf

[2] D. Castañon. Optimal search strategies in dynamic hypothesis testing. *IEEE Trans. Systems, Man. and Cybernetics*, 25(7):1130–1138, Jul. 1995.

[3] L. Lai, H. V. Poor, Y. Xin and G. Georgiadis. Quickest search over multiple sequences.*IEEE Trans. Inform. Theory*, 57(8):5375–5386, Aug. 2011.

[4] H. V. Poor. *An introduction to Signal Detection and Estimation*. Springer-Verlag, 1994.

[5] H. V. Poor and O. Hadjiliadis. *Quickest Detection*. Cambridge, UK: Cambridge University Press, 2008.

[6] A. N. Shiryaev. *Optimal Stopping Rules*. New York: Springer, 1978.