Dynamic sequential team multi-hypothesis testing under uniformly distributed nonstationary observations

Ji-Woong Lee, Geir E. Dullerud

Department of Electrical Engineering, Pennsylvania State University, University Park, PA 16802, USA
Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

A R T I C L E   I N F O

Article history:
Received 1 June 2005
Received in revised form 5 November 2007
Accepted 25 June 2008
Available online 6 August 2008

Keywords:
Information structure
Member-by-member optimality
Minimum principle
Team theory

A B S T R A C T

We consider a dynamic team of communicating decision makers performing Bayesian sequential multi-hypothesis testing under nonstationary observations. A simple decision rule is shown to be member-by-member optimal if the observations are uniformly distributed, if the observation costs are defined to reflect the quality of observations, and if “vague” terminal decisions are allowed without additional penalty. The “vague” terminal decisions can be made “clear” by performing a finite number of member-by-member optimal sequential decisions successively.

1. Introduction

In distributed sensor networks, it is often required that multiple decision makers be allowed due to decentralized information, and the problem of sequential hypothesis testing [1] becomes that of sequential teams where member-by-member optimal sequential decision rules are being sought [2,3]. However, even in the simplest problem with two hypotheses, static information (i.e., no communication among decision makers) and stationary observations, the coupling between the local decisions through the common cost causes considerable complexity in the computation of a member-by-member optimal solution [4]. Moreover, if the decision makers communicate without sharing their information, the information is dynamic (i.e., the separation between estimation and decision making no longer holds in general), and the structure of member-by-member optimal rules becomes unclear [5–7].

We are concerned with the case where there are more than two hypotheses, the information is dynamic, and the observations are nonstationary. Moreover, we allow the decisions at each time instant to be nonsequential, and so the decision problem does not revert to an equivalent problem with static information [8]. The decomposition technique developed in [4], therefore, does not generalize to this case. Our approach is to choose a candidate sequential decision rule, which admits a nonsequential representation, and check its optimality under the following simplifying assumptions: the observations are uniformly distributed, as well as independent, conditioned on each hypothesis; decision errors are uniformly penalized regardless of the number of errors; and the observation costs are defined in a way that reflects the quality of observations. We use Varaiya and Walrand’s minimum principle [9] to check the optimality of the candidate decision rule.

The M-ary sequential hypothesis testing problem, with $M > 2$, is considerably more difficult than the binary counterpart; even in the centralized case, the solution is very complex in general [10, 11]. Although our member-by-member optimal decision rule is valid for all $M$, it is weak in the sense that the decision makers are allowed to make “vague” (i.e., set-valued) terminal decisions. However, these “vague” decisions in effect reduce the $M$-ary problem to an $(M - 1)$-ary problem in an optimal way, so one can obtain “clear” decisions by performing at most $M - 1$ separate sequential decisions successively.

This work is motivated by the need to develop a systematic methodology for distributed networks of communicating mobile sensors that perform the tasks such as surveillance and reconnaissance that involve sequential hypothesis testing. The main contribution of this work is that an explicit solution is obtained for a tractable special case with multiple hypotheses, dynamic information structure, and nonstationary observations. The obtained optimal decision rule has an important property in the context of...
mobile sensor networks; namely, at each time the decisions can be made without knowing the future positions of the sensors. Another point to note is that prior to this work, whose preliminary version appears in a conference proceedings [12], the only application of Varaiya and Walrand’s minimum principle has been to packet switched satellite communication [13].

A simple motivating example is given in Section 2 and revisited in Section 6. Varaiya and Walrand’s minimum principle is summarized in Section 3. Section 4 formulates the problem and makes aforementioned assumptions. Then in Section 5 the optimality of a candidate decision rule is shown. Concluding remarks are made in Section 7.

2. An example

This section, together with Section 6, is independent of other sections and provides an example that illustrates how the optimal sequential decision rule works. Suppose that there are N agents in a bounded region with a partition \(D = \{D_1, \ldots, D_N\}\), where \(N < L\). The agents are to detect \(N\) unknown hot spots \(D_1, \ldots, D_N\), \(1 \leq i_1 < \cdots < i_N \leq L\), and to position themselves among \(D\) so that agent \(j\) is located in \(D_j\) for \(j = 1, \ldots, N\). Let \(H = \{D_1, \ldots, D_N\}\). The agents are mobile, and associated with each agent is a sensor that makes imperfect local observations. Let us assume the simple case where the observation \(y_i(t)\) of agent \(i\) at time \(t\) is as follows: if agent \(i\) is in \(D_j\) at time \(t\), then with probability \(p_j(t, j)\), the local observation \(y_i(t)\) contains full information on whether \(D_j\) belongs to \(H\) or not; with probability \(1 - p_j(t, j)\), the observation \(y_i(t)\) has no information on whether \(D_j\) is in \(H\) or not.

Let \(I\) be the set of all \(N\)-tuples \((j_1, \ldots, j_N)\) of positive integers such that \(1 \leq j_1 < \cdots < j_N \leq L\); let \(H_j = \{D_{j_1}, \ldots, D_{j_N}\}\) for each \(j = (j_1, \ldots, j_N) \in I\). Then \(H_j \subseteq D\) for all \(j \in I\), so each \(H_j\) can be considered to be a hypothesis on the hot spots, and the number of hypotheses is equal to \(M = \binom{L}{N}\), which is the cardinality of \(I\). The true hypothesis is \(H\), and \(H_j = H\) only when \(J = (i_1, \ldots, i_N)\).

As a team, the \(N\) agents perform sequential M-ary hypothesis testing. When agent \(i\) is located in \(D_j\) at time \(t\), agent \(i\) makes its (set-valued) local decision \(u_i(t)\) and broadcasts it to other agents as follows: if \(y_i(t)\) reveals that \(D_j \in H\), then \(u_i(t) = \{j_1, \ldots, j_N\} \in I: j_k = j\) for some \(k\); if \(y_i(t)\) tells that \(D_j \notin H\), then \(u_i(t) = \{j_1, \ldots, j_N\} \in I: j_k \neq j\) for each \(k\); if \(y_i(t)\) contains no information, then agent \(i\) is undecided and \(u_i(t)\) is empty. If agent \(i\) is undecided but another agent’s decision is nonempty, then agent \(i\) decides to stop observations and follows the other agent’s decision. Consequently, at time \(t + 1\), the agents start a new sequential hypothesis testing as a team, where the hypotheses are now those in the intersection of nonempty decisions made at time \(t\). Proceeding in this manner, the \(N\) agents sequentially reduce the number of hypotheses until only the true hypothesis remains and all the hot spots are identified.

3. Varaiya and Walrand’s minimum principle

A finite-dimensional discrete-time decentralized stochastic control system with \(N\) controllers has the representation

\[
x(t + 1) = f(t, x(t), u(t), w(t))
\]

for \(t = 0, 1, \ldots, T\). Here \(x(t)\) is a state-like vector that defines how information is updated at time \(t\); \(x(t)\) is allowed to have time-varying dimension. The vector \(u(t) = (u_1(t), \ldots, u_N(t))\) is the tuple of local decisions, where \(u_i(t)\) is the local decision of the \(i\)-th controller at time \(t\). The initial state \(x(0)\) and the disturbances \(w(0), w(1), \ldots, w(T)\), are mutually independent random variables. The controllers have perfect recall; if \(z_i(t)\) denotes the new information that becomes available to controller \(i\) at time \(t\), then the information available to controller \(i\) up to time \(t\) is given by

\[
z_i(t) = (z_i(0), \ldots, z_i(t)).
\]

The only requirement is that \(z_i(t)\) is a subvector of \(x(t)\) for each \(i\) and \(t\).

The decisions \(u_i(t)\) are images of information under functions \(\gamma\):

\[
u_i(t) = \gamma_i(t, z_i^1).
\]

Here \(\gamma_i\) are such that the decision rule (or control law) \(\gamma = (\gamma_1, \ldots, \gamma_N)\) belongs to a class \(\Gamma\) of admissible decision rules. The performance of the system (1) under the decision rule \(\gamma \in \Gamma\) is evaluated by the cost

\[
J(\gamma) = E^\gamma \sum_{i=0}^{T} c(t, x^t, u(t))
\]

for some real-valued function \(c\) and a given terminal time \(T\), where \(E^\gamma\) denotes the expectation that depends on the decision rule \(\gamma\), and \(x^t = (x(0), \ldots, x(t))\). For \(\gamma, \eta \in \Gamma\), denoted by \(\gamma|\eta\) and \(\gamma|\eta|\) are the decision rules

\[
(\gamma|\eta)_i(s, \cdot) = \begin{cases} 
\gamma_i(s, \cdot) & \text{if } f \neq i; \\
\eta_i(s, \cdot) & \text{if } f = i.
\end{cases}
\]

(2a)

\[
(\gamma|\eta)^{-1}_i(s, \cdot) = \begin{cases} 
\gamma_i(s, \cdot) & \text{if } f \neq i \text{ or } s > t; \\
\eta_i(s, \cdot) & \text{if } f = i \text{ and } s \leq t.
\end{cases}
\]

(2b)

We assume that \(\gamma|\eta, \gamma|\eta| \in \Gamma\) for all \(i\) and \(t\) whenever \(\gamma, \eta \in \Gamma\).

Definition 1. A decision rule \(\gamma \in \Gamma\) is said to be member-by-member optimal if \(J(\gamma|\eta) \leq J(\gamma|\eta|)\) for all \(\eta = (\eta_1, \ldots, \eta_N) \in \Gamma\) and for all \(i = 1, \ldots, N\).

Theorem 2.\] A decision rule \(\gamma \in \Gamma\) is member-by-member optimal if there exist functions \(W(t, x^t, u(t), t = 0, \ldots, T+1\), with

\[
E^\eta[W(t, x^t, u(t)) | z_i^1] \leq E^\eta[W(t, x^t, u(t)) + W(t + 1, x^{t+1})] | z_i^1],
\]

\[
E^{\gamma|\eta}_{t+1}[W(t, x^t, u(t)) | z_i^1] = E^{\gamma|\eta^1}_{t+1}[W(t, x^t, u(t))]
\]

for all \(\eta = (\eta_1, \ldots, \eta_N) \in \Gamma\), \(i = 1, \ldots, N\), and \(t = 0, \ldots, T\).

According to Definition 1, a decision rule \(\gamma\) is member-by-member optimal if any unilateral deviation from it cannot reduce the cost. Theorem 2 gives a sufficient condition for \(\gamma\) being member-by-member optimal. The condition states that the functions \(W(t, \cdot, u(t), \cdot)\) make the decision rule \(\gamma|\eta\), which is a unilateral deviation from \(\gamma\), “appear” to be no better than \(\gamma\) for any \(\eta\) and \(i\). The condition, however, is highly restrictive because it implies that \(\gamma\) also appears to give the best present and future performance even if \(\gamma|\eta\) was used in the past. The proof of Theorem 2 provided in [9] shows that one may take \(W(t, x^t) = E^\Gamma[\sum_{s=t}^{T} c(s, x^s, u(s)) | x^t]\).

4. Problem formulation

There are \(M\) hypotheses \(H_0, H_1, \ldots, H_{M-1}\) on a static parameter \(h\), so that \(h = k\) under \(H_k\). The a priori probabilities of these hypotheses are \(P(H_k) = \lambda_k\) with \(\lambda_k \in (0, 1)\) for \(k = 0, \ldots, M - 1\), and \(\sum_{k=0}^{M-1} \lambda_k = 1\). (We abuse the notation by using the same symbol \(P\) to denote the probability of any event; we write \(P^\Omega\) when...
$P$ depends on a decision rule $\gamma$.) We write $\lambda = (\lambda_0, \ldots, \lambda_{M-1})$. Let $S = \{0, \ldots, M-1\}$ be the set of all hypotheses, and define

$$k^* = \arg \min \{\lambda_k; k \in S\}; \quad L(\lambda) = \min \{\lambda_k; k \in S\}$$

Let $N$ be the number of decision makers (or detectors). Let $y_i(t) \in \mathbb{R}^{m_i(t)}$ be the observation made by detector $i$ at time $t$, given by

$$y_i(t) = g_i(t, h) + u_i(t)$$

for $i = 1, \ldots, N$, and $t = 0, 1, \ldots$, where $g_i(t, \cdot)$ are known functions and $m_i(t)$ the time-varying dimension of detector $i$'s observation. It is assumed that $h$, $w_1(0), \ldots, w_N(0)$, $w_1(1), \ldots, w_N(1)$, are independent random variables, so that $y_i(t)$ are independent conditioned on each hypothesis.

We allow $2^M - 1$ possible decisions for each detector at each time instant, and represent each decision as a proper subset of $S$. At time $t$, detector $i$ samples $y_i(t)$ and makes a decision $u_i(t) \in \mathcal{P}(S) \setminus \{\emptyset\}$, where $\mathcal{P}(S)$ is the power set of $S$. If detector $i$ decides at time $t$ to continue to sample $y_i(t+1)$, then $u_i(t) = \emptyset$; in this case, no signal is transmitted from detector $i$ to the others. If detector $i$ decides at time $t$ to stop observation and declare its terminal decision, then $u_i(t)$ is nonempty; that is, if detector $i$ declares that the true hypothesis is among a nonempty proper subset $K$ of $S$, then $u_i(t) = K$. A terminal decision (i.e., a nonempty proper subset $K$ of $S$) is said to be “clear” if it is a singleton; otherwise, it is said to be “vague”. Whenever $K$ is a terminal decision, we write $H_K = \{H_k; k \in K\}$.

**Fig. 1** shows the block diagram of the decision model, where the decisions $u_i(t)$ are sent to every detector at each time $t$. This diagram signifies that each detector in effect conveys information to the others even when $u_i(t)$ are empty, and that the decisions at each time are nonsequential (i.e., the order of decisions depends on the observations) [3,14]. A nonsequential system can potentially perform better than sequential ones as long as it is deadlock-free [15,16]. The evolution of $x$ and $z_i$ will be determined in Section 5.

An appropriate class $\Gamma$ of admissible decision rules will also be defined in Section 5. The cost of a decision rule $\gamma \in \Gamma$ is

$$J^T(\gamma) = E^\gamma \sum_{t=0}^T c^T(t, h, u(t)).$$

The function $c^T$ is given by

$$c^T(t, h, u(T)) = \sum_{j=1}^N \mu_j(T) 1_{[u_j(T) = u]} + d^T(h, u(T)),$$

$$c^T(t, h, u(t)) = \sum_{j=1}^N \mu_j(t) 1_{[u_j(t) = u]}$$

for $t = 0, \ldots, T - 1$, where $1_{[\cdot]}$ denotes the indicator function, and $\mu_j(t)$ the cost of sampling $y_j(t+1)$. The initial observations $y_j(0)$ are free. The function $d^T$ represents the cost of terminal decisions. The objective is to find a decision rule $\gamma \in \Gamma$ such that $J^T(\gamma) \leq J^T(\gamma|\eta_i)$ for $\eta_i \in \Gamma$ and $i = 1, \ldots, N$.

We make a few assumptions. First, we assume that each $u_i(t)$ is uniformly distributed on a bounded subset of $\mathbb{R}^{m_i(t)}$. Then, for each $i$ and $t$, we can define disjoint measurable sets $Y_i(t, K) \subset \mathbb{R}^{m_i(t)}$, $K \in \mathcal{P}(S) \setminus \{\emptyset\}$, as the set difference of the intersection of the sets of all possible values of $y_i(t)$ under hypothesis $H_K$ over all $k \in K$ and the union of the sets of all possible values of $y_i(t)$ under hypothesis $H_S$ over all $k \not\in K$; write $Y_i(t, \emptyset) = Y_i(t, S)$. Then associated with $Y_i(t, K)$ are numbers $\rho_i(t)$ and $\rho_i^K(t)$ such that

$$P(y_i(t) \in Y_i(t, \emptyset) \mid H_K) = \rho_i^K(t) = \rho_i(t)$$

whenever $k \in S$, and such that

$$P(y_i(t) \in Y_i(t, K) \mid H_K) = \rho_i^K(t)$$

whenever $K \neq \emptyset$ and $k \in K$, where

$$\rho_i(t) = \sum_{K \neq \emptyset \cap k \in K} \rho_i^K(t) = 1$$

for $k \in S$. **Fig. 2** shows an example of these sets. Next, we assume a uniform cost $\epsilon > 0$ of terminal decision errors where no additional penalty is imposed on multiple terminal decision errors and “vague” terminal decisions:

$$d^T(h, u(T)) = \epsilon 1_{[u_j(T) \neq \emptyset \text{ and } u_j(T) \neq u_j(T) \text{ for some } j]}$$

Finally, it is assumed that the observation costs reflect the quality of observations and that the cost of terminal decision errors is sufficiently large:

$$\mu_j(t) < \epsilon(1 - \rho(t+1))L(\lambda)$$

where

$$\rho(t) = \prod_{j=1}^N \rho_j(t); \quad \mu(t) = \sum_{j=1}^N \mu_j(t)$$

for $t = 0, 1, \ldots$. Inequality (4) implies that the non-free observations $y_j(t+1), t = 0, 1, \ldots$, are such that the ratio between their aggregate cost $\mu(t)$ and aggregate quality $1 - \rho(t+1)$ is bounded. We assume $\rho(t) \in (0, 1)$ for all $t$.

**Notation.**

We shall write $y_i(t) = (y_i(0), \ldots, y_i(t))$ as well as $u(t) = (u(0), \ldots, u(t))$ for each $t$. Likewise, $u^d = (u(0), \ldots, u(t))$, $y^d = (y(0), \ldots, y(t))$, and $y^d_i = (y_i(0), \ldots, y_i(t))$. We use the convention that $u_i(-1) = \emptyset$ and $u_i^d = \emptyset$ for all $i$; likewise, $u(-1) = \emptyset$ and $u^{-1} = \emptyset$. Define $U(K) \subset \mathcal{P}(S)^N, K \in \mathcal{P}(S)$, as follows:

$$U(\emptyset) = \{(\emptyset, \ldots, \emptyset)\},$$

$$U(K) = \{u_1, \ldots, u_K \in \mathcal{P}(S)^N; \bigcap_{j \not\in K} u_j = K\}$$
for nonempty proper subsets $K$ of $S$, and

$$U(S) = \mathcal{P}(S)^N \setminus \bigcup_{K \in \mathcal{P}(S) \setminus \{S\}} U(K).$$

Then, for each $t$, $u(t)$ belongs to one of the disjoint sets $U(K)$, $K \in \mathcal{P}(S)$. The observation sequence $y^t_i$ belongs to one of the disjoint sets $Y^t_i(K)$, $K \in \mathcal{P}(S) \setminus \{S\}$, defined as follows: $y^t_i \in Y^t_i(\emptyset)$ if $y_i(s) = Y_i(s, \emptyset)$ for all $s \leq t$; and $y^t_i \in Y^t_i(K)$ for some nonempty proper subset $K$ of $S$ if $y^t_i \notin Y^t_i(\emptyset)$ and $y_i(s) \in Y_i(s, K)$ for $s \leq t$ with

$$\bigcap_{1 \leq j \leq t, K_j \neq \emptyset} K_j = K.$$

(Here, it is not possible to have $y^t_i \notin Y^t_i(\emptyset)$ and $K = \emptyset$.) Similarly, the observation sequence $\tilde{y}^t$ belongs to one of the disjoint sets $Y^t(\emptyset), K \in \mathcal{P}(S) \setminus \{S\}$: $\tilde{y}^t_i \in Y^t(\emptyset)$ if $y_i(s) \in Y_i(s, \emptyset)$ for all $j$ and for all $s \leq t$; and $\tilde{y}^t_i \in Y^t(K)$ for some nonempty proper subset $K$ of $S$ if $\tilde{y}^t_i \notin Y^t(\emptyset)$ and $y_i(s) \in Y_i(s, K)$, for all $j$ and $s \leq t$ with

$$\bigcap_{(j,s): s \leq t, K_{j,s} \neq \emptyset} K_{j,s} = K.$$

5. Member-by-member optimal solution

To obtain a decision model that conforms with that of Section 3, let us introduce intermediate decisions $\tilde{u}(t) = (\tilde{u}_1(t), \ldots, \tilde{u}_N(t))$, with associated running costs $\tilde{c}^t(t, h, \tilde{u}(t)) = 0$, in addition to the actual decisions $u(t)$. Let $x$ and $z_i$ evolve according to

$$x(2t) = (h, y(t), u(t - 1)); \quad x(2t + 1) = (h, \tilde{u}(t))$$

and

$$z_i(2t) = (y_i(t), u(t - 1)); \quad z_i(2t + 1) = \tilde{u}(t)$$

for $i = 1, \ldots, N$ and $t = 0, 1, \ldots$. Then the information available to detector $i$ up to time $t$ is given by $z_i^{2t+1}$ where

$$z_i^{2t} = (y_i^{2t}, \tilde{u}_i^{2t - 1}, u^{2t - 1}); \quad z_i^{2t + 1} = (y_i^{2t}, \tilde{u}_i, u^{2t - 1})$$

for each $i$ and $t$.

Let $\Gamma$ be the set of measurable decision rules $\eta$ satisfying the following: $\tilde{u}_i(t) = \eta_i(2t, z_i^{2t}) \in \mathcal{P}(S) \setminus \{S\}$ and $u_i(t) = \eta_i(2t + 1, z_i^{2t+1}) \in \mathcal{P}(S) \setminus \{S\}$ for all $i$ and $t$; whenever $\tilde{u}_i(t) \neq \emptyset$ for some $(i, t)$, we have $\tilde{u}_i(s) = \tilde{u}_i(t)$ for $s > t$, and $u_i(s) = \tilde{u}(s)$ for $s \geq t$; whenever $u_i(t) \neq \emptyset$ for some $(i, t)$, we have $u_i(s) = u_i(s) = u_i(t)$ for $s > t$.

Let $\gamma$ be a candidate decision rule such that $\gamma \in \Gamma$, and such that, for $i = 1, \ldots, N, t = 0, 1, \ldots$, and $K \in \mathcal{P}(S) \setminus \{\emptyset, S\}$,

$$\gamma_i(2t, z_i^{2t}) = \begin{cases} \emptyset & \text{if } u(t - 1) \in U(\emptyset) \text{ and } y_i(t) \in Y_i(t, \emptyset); \\ \emptyset & \text{if } u(t - 1) \in U(\emptyset) \text{ and } y_i(t) \in Y_i(t, K); \\ S \setminus \{k^*\} & \text{if } u(t - 1) \in U(S), \end{cases}$$

whenever $z_i^{2t}$ has $u_i(t - 1) = \emptyset$, and

$$\gamma_i(2t + 1, z_i^{2t+1}) = \begin{cases} \emptyset & \text{if } \tilde{u}(t) \in U(\emptyset) \text{ and } y_i(t) \in Y_i(t, \emptyset); \\ K & \text{if } \tilde{u}(t) \in U(\emptyset) \text{ and } y_i(t) \in Y_i(t, K); \\ S \setminus \{k^*\} & \text{if } \tilde{u}(t) \in U(S) \end{cases}$$

whenever $z_i^{2t+1}$ has $u_i(t) = \emptyset$. Note that, under $\gamma$, the detectors are allowed not to make terminal decisions by time $T$; that is, if a detector is not able to make a reliable decision by time $T$, it may remain undecided with the expense of additional observation cost. The fourth case of (7a) and the second and fourth cases of (7b) never occur under $\gamma$. However, these cases are useful in dealing with the decision rules $\gamma'_{\eta_i}, \gamma'_{\eta_i, 2t+1}$, and $\gamma'_{\eta_i}$ defined as in (2).

Given $\mu(t), \rho(t), \epsilon, \rho(t)$, and $\mu(t)$ as in (3) and (4), let $V^T(t)$ be a value function-like quantity defined by

$$V^T(T + 1) = 0, \quad V^T(t) = \mu(t) + \rho(t + 1) V^T(t + 1)$$

for $t = 0, \ldots, T$. Then it follows from (4) that

$$V^T(t) < \epsilon L(\lambda)$$

for all $t$. Define $W$, which will be used to test the optimality of $\gamma$, by

$$W(2(T + 1), x^{2(T + 1)}) = 0;$$

and

$$W(2t, x^{2t + 1}) = \begin{cases} V^T(t) & \text{if } u(t - 1) \in U(\emptyset) \text{ and } y^t_i \in Y^t_i(\emptyset); \\ \sum_{\eta \in \mathcal{P}(S) \setminus \{\emptyset, S\}} \eta_i(1)_{\tilde{y}^t_i} t + 1 & \text{if } u(t - 1) \in U(\emptyset) \text{ and } y^t_i \in Y^t_i(K); \\ \epsilon & \text{if } u(t - 1) \in U(S); \end{cases}$$

for $i = 1, \ldots, N, t = 0, 1, \ldots, T$, and $K \in \mathcal{P}(S) \setminus \{\emptyset, S\}$. Then it is not difficult to verify that

$$E^{\gamma_{\eta_{i}}}[W(2T + x^{2T + 1})] z_i^{2T + 1}$$

$$= E^{\gamma_{\eta_{i}}}[E^T(h, u(t)) + W(2T + 1, x^{2T + 1})] z_i^{2T + 1}$$

and

$$E^{\gamma_{\eta_{i}}}[W(2T + 1, x^{2T + 1})] z_i^{2T + 1}$$

for $i = 1, \ldots, N, t = 0, 1, \ldots, T$, and $K \in \mathcal{P}(S) \setminus \{\emptyset, S\}$.

The following lemma establishes that the conditional expectations on the left-hand sides of equalities (10a) and (10c), with $\gamma'_{\eta_{i}}$ replaced by $\gamma'_{\eta_{i}}$, are lower bounds on $E^{\gamma_{\eta_{i}}}[E^T(t, u(t)) + W(2T + 1, x^{2T + 1})] z_i^{2T + 1}$ and $E^{\gamma_{\eta_{i}}}[E^T(t, u(t)) + W(2T + 1, x^{2T + 1})] z_i^{2T + 1}$, respectively, over all $\eta \in \Gamma$.

**Lemma 3.** Let $x$ and $z_i$, $i = 1, \ldots, N$, be as in (5) and (6); let $\gamma' \in \Gamma$ be such that (7a) holds whenever $z_i^{2t}$ has $u_i(t - 1) = \emptyset$, and such
Under $\gamma$, the intermediate decisions $\hat{u}_i(t)$ manifest themselves only if $\hat{u}_i(t) \neq \varnothing$ for some $i$, in which case $u_i(t) = \hat{u}_i(t)$ and $u_j(t) \neq \varnothing$ for all $j$; otherwise, $u_i(t) = \varnothing$ for all $j$. Therefore, $\hat{u}(t)$ are completely “hidden”, and $\gamma$ admits a deadlock-free nonsequential representation $\gamma^*$ defined as follows: $u_i(t) = \gamma_i^*(t, \zeta^*_i)$ where $\zeta^*_i = (y_i, u_i, u_i', j \neq i)$ and

$$
\gamma^*_i(t, \zeta^*_i) = \bigcap_{\varnothing \neq j, y_i(t) \in Y_j(t, \varnothing)} u_i(t)
$$

for $i = 1, \ldots, N$, $t = 0, 1, \ldots$, and $K \in \mathcal{P}(S \setminus \varnothing, S)$. Underlying assumptions here are that all members $i$ having $y_i(t) \in Y_i(t, K)$ for some nonempty proper subset $K$ of $S$ are required to make their terminal decisions simultaneously, and that these terminal decisions are conveyed to all the other members instantaneously.

In practice, the decision $\gamma_i(y_i(t), \gamma_i(t)) \in \mathcal{P}(Y_i(t, \varnothing), \varnothing)$ is accompanied by a set of decisions $\gamma_j(t)$ of the members having $y_j(t) \neq \varnothing$ for some $j \neq i$, and $y_i(t) \in Y_i(t, K)$ can be replaced with the first $u_i(t) \neq \varnothing$ that is conveyed to member $i$ during the sampling period $[t, t + 1]$;

this does not affect the cost of $\gamma^*$.

The decision rules $\gamma$ and $\gamma^*$ are equivalent in the sense that they share the same $u_i(t)$ for all $i$ and $t$. The cost of $\gamma$, and hence of $\gamma^*$, is

$$
J^I(\gamma) = \rho(0) V^I(0),
$$

which is the same as the optimal cost under the assumption (4) for the centralized problem with observation costs $\mu_k(t), t = 0, 1, \ldots$, and terminal decision cost $\epsilon$. Therefore, $\gamma$ is optimal over “all” decision rules. The optimal cost $J^I(\gamma)$ does not involve the cost $\epsilon$ of decision errors because, under $\gamma$, the terminal decisions are either $\varnothing$ (i.e., undecided) or a proper subset $K$ of $S$ such that $i \in K$.

Suppose that $M > 2$. Then there are $M$ possible “vague” terminal decisions of cardinality $M - 1$, and the assumed cost structure is such that there is no advantage in making terminal decisions of cardinality less than $M - 1$. Therefore, the decision rule $\gamma$ in effect gives an optimal way to reduce the $M$-ary hypothesis testing problem to an $(M - 1)$-ary hypothesis testing problem. A “clear” terminal decision, if desired, can be obtained by performing at most $M - 1$ separate sequential decisions, where each sequential decision reduces a “vague” decision to another decision which is strictly less “vague” (i.e., of strictly less cardinality) in an optimal way.

A nice property of the decision rule $\gamma$ in the context of mobile sensor networks is that the decision of one detector at the present time does not depend on the future distributions of its and the other detectors’ observations. That is, decisions at time $t$ can be made without knowledge of the quantities $\mu_i(t), \rho_i(t + 1), \mu_i(t + 1), \rho_i(t + 2), \ldots, j = 1, \ldots, N$, that are directly related to the future sensor positions.

Inequality (4) has a simple interpretation in the case of the infinite-horizon problem with stationary observations: if $\mu(t) = \mu$ and $\rho(t) = \rho$ for all $t$, then inequality (4) becomes

$$
\frac{\mu}{1 - \rho} < L(\lambda),
$$

and the optimal cost reduces to

$$
\lim_{t \to \infty} J^I(\gamma) = \lim_{t \to \infty} \left(1 - \frac{\rho^{t+1}}{1 - \rho}\right) \rho \mu = \rho \mu \frac{\mu}{1 - \rho}.
$$

Inequality (13) guarantees that the probability that all the detectors make their terminal decisions at time $0$ is strictly less than 1. That is, if (13) does not hold, and if $y_0^i \in Y^0(\varnothing)$, then it is optimal for every detector to stop at $t = 0$ and declare $H_{\emptyset}(k_i)$.

that (7b) holds whenever $z_{i+1}^T$ has $u_i(t) = \varnothing$. Then $W$ defined in (9) satisfies

$$
E^V Y^{10}(W(2t + 1, z_{i+1}^T) \mid z_{i+1}^T) \leq E^V \left(\mathbf{u}_i(t) + W(2(t + 1), x_{i+1}^2) \mid z_{i+1}^T\right)
$$

and

$$
E^V \left(Y^{10}(W(2t + 1, z_{i+1}^T) \mid z_{i+1}^T) \leq W(2t + 1, z_{i+1}^T) + W(2(t + 1), x_{i+1}^2) \mid z_{i+1}^T\right)
$$

(11a)

for all $\eta \in \Gamma$, $i = 1, \ldots, N$, and for all $t = 0, 1, \ldots, T$.

Proof. What is given here is a much abbreviated version of the proof. It is readily seen that $E^V Y^{10}(W(2t, x_{i+1}^2) \mid z_{i+1}^T)$ and $E^V \left(Y^{10}(W(2t + 1, x_{i+1}^2) \mid z_{i+1}^T) \right)$ are given by the right-hand sides of equalities (10b) and (10d), respectively, with $P^{Y_{i+1}^2}$ and $P^{Y_{i+1}^4}$ replaced by $P^{Y_{i+1}^3}$. Furthermore, it is tedious but not difficult to see that

$$
E^V \left(Y^{10}(W(2(t + 1), x_{i+1}^2) \mid z_{i+1}^T) \right) \leq E^V \left(\mathbf{u}_i(t) + W(2(t + 1), x_{i+1}^2) \mid z_{i+1}^T\right)
$$

and

$$
E^V \left(Y^{10}(W(2(t + 1), x_{i+1}^2) \mid z_{i+1}^T) \right) \leq E^V \left(\mathbf{u}_i(t) + W(2(t + 1), x_{i+1}^2) \mid z_{i+1}^T\right)
$$

(11b)

for all $\eta \in \Gamma$, $i = 1, \ldots, N$, and for all $t = 0, 1, \ldots, T$.

Proof. The result follows from (10a), (10c), Lemma 3, and Theorem 2. □
On the other hand, for the finite-horizon problem with stationary observations, the proof of Lemma 3 suggests we can replace (4) with a less conservative condition. Under this condition, the decision rule $\gamma$ is not necessarily optimal over all decision rules but is still member-by-member optimal over the sequential decision rules in $I'$. This result is stated below as a corollary.

**Corollary 5.** Suppose that $p_i(t) = \rho_i$ and $\mu_i(t) = \mu_i$ for all $i = 1, \ldots, N$ and for all $t = 0, 1, \ldots$. If

$$
\epsilon L(\lambda) \geq \max \left\{ \mu_i + \frac{(1 - \rho)^i \rho_{\mu_i} (1 - \rho^{i+1}) \rho_{\mu_i}}{1 - \rho} \right\}
$$

for all $i = 1, \ldots, N$, then $\gamma$ is member-by-member optimal for terminal time $T$.

**Proof.** Inequality (14) follows from (12) with $\rho(t) = \rho_i, \mu(t) = \mu_i$ and $V^1(t) = (1 - \rho^{I+1}) / (1 - \rho)$. $\square$

6. An illustration

Let us continue to consider the example described in Section 2. If agent $i$ is positioned in $D_i$ at time $t$, write $pos_i(t) = j$. Then we have

$$p_i(t) = p_i^k(t) = 1 - p_i(t, pos_i(t)),$$

and that

$$p_i^k(t) = 1 - \rho_i(t)$$

for either $K = K_{pos_i(t)}$ or $K = I \setminus K_{pos_i(t)}$, where $K = \{1, \ldots, N\} \setminus \{j : \text{pos}_i(t) = j \text{ for some} k\}$.

we have $p_i^k(t) = 0$ for all other nonempty subsets $K$ of $I$, so

$$p_i(t) + \sum_{K:|i| \cup \{j:K_{pos_i(t)}\}} p_i^k(t) = 1$$

for all $\{j_1, \ldots, j_n\} \in I$. The sets $Y_i(t, K), K \subset I$, are defined as follows: if $D_{pos_i(t)} \in H$, then

$$Y_i(t, K) = \begin{cases} \{y_i(t) : (y_i(t) \text{ has no information}) \} & \text{if } K = \emptyset; \\ \{y_i(t) : (y_i(t) \text{ has full information}) \} & \text{if } K = K_{pos_i(t)}; \\ \emptyset & \text{otherwise}, \end{cases}$$

and, if $D_{pos_i(t)} \not\in H$, then

$$Y_i(t, K) = \begin{cases} \{y_i(t) : (y_i(t) \text{ has no information}) \} & \text{if } K = \emptyset; \\ \{y_i(t) : (y_i(t) \text{ has full information}) \} & \text{if } K = I \setminus K_{pos_i(t)}; \\ \emptyset & \text{otherwise}. \end{cases}$$

Now let us illustrate how the optimal decision rule $\gamma$ is implemented. Let $t$ be the first time where $u_i(t) \not\in \emptyset$ for some $i$ under $\gamma$; let

$I' = \bigcap_{t \in u_i(t)} u_i(t).$

Then the original $M$-ary sequential hypothesis testing problem is reduced to the $M'$-ary problem, where $M'$ is the cardinality of $I'$, to determine the true hypothesis among $H_j, j \in I'$. For instance, let $N = 3, I = 6$, and $H = \{D_1, D_2, D_3\}$. Suppose that $pos_i(t) = 3$ with $y_1(t)$ having full information, that $pos_i(t) = 6$ with $y_2(t)$ having no information, and that $pos_i(t) = 4$ with $y_3(t)$ having full information. Then, writing $j_1j_2j_3$ for $(j_1,j_2,j_3)$ to simplify notation, we have

$$u_1(t) = \{123, 134, 135, 136, 234, 235, 236, 345, 346, 356\},$$

$$u_2(t) = \{123, 125, 126, 135, 136, 235, 236, 256, 356\},$$

and so agents 1 and 3 stop observations and broadcast their local decisions. Agent 2 is undecided at first, but upon receiving $u_1(t)$ and $u_2(t)$, agent 2 also decides to stop observations and makes its terminal decision

$$u_3(t) = u_1(t) \cap u_2(t) = \{123, 135, 136, 235, 236, 356\},$$

which is broadcasted. Although agents 1 and 3 have made their own terminal decisions that are “vaguer” than that of agent 2, they are not allowed to change their terminal decisions. Nevertheless, all the agents share a common knowledge about the possible locations of the hot spots by the end of time $t$. Thus, the original 20-ary hypothesis testing problem, in effect, reduces to a six-ary problem where the six hypotheses are $H_{123}, H_{135}, H_{136}, H_{235}, H_{236}, H_{356}$; if desired, this new sequential hypothesis testing problem starts at time $t + 1$ to reduce the number of hypotheses further.

The optimal sequential decision rule $\gamma$ does not depend on the positions of the agents (as long as (4) is satisfied). Hence, the problem of optimally positioning the agents can be considered separately. For instance, if $p(t_i, j)$ are bounded away from zero, and if $\rho_i(t)$ are such that $\mu(t_i)/(1 - \rho(t_i))$ is constant, then in the case of infinite horizon problem, it is optimal to place the agents at time zero so that $[1 - p_i(0, \text{pos}(0))]$ is minimized until an agent makes a nonempty decision at some time $t$. It is also possible to derive an optimal agent placement strategy that minimizes the total cost over successive execution of sequential hypothesis testing until a “clear” terminal decision is reached.

7. Concluding remarks

We focused on a problem of decentralized sequential multi-hypothesis testing with dynamic information and nonstationary observations. We chose a simple candidate decision rule, and showed that it is member-by-member optimal under the assumptions that the observations are uniformly distributed conditioned on each hypothesis, that decision errors are uniformly penalized, and that the observation costs depend on the observation quality in a certain way.

If the observations are nonuniform, there are no constructive methods to solve this problem, and our result does not apply. Possible future work includes theoretical and experimental verification of the performance of the obtained decision rule as a suboptimal one under bounded but nonuniform observations.

Acknowledgments

The authors thank the anonymous referees for their helpful comments and suggestions.

References


