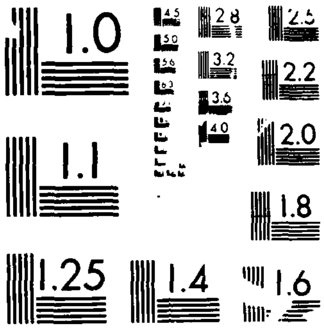
AD-A195 910 ON HYPOTHESIS TESTING IN DISTRIBUTED SENSOR NETWORKS (U) SYRACUSE UNIV NY DEPT OF ELECTRICAL AND COMPUTER ENGINEERING Z CHAIR ET AL. NOV 87 RADC-TR-87-180 1/2

UNCLASSIFIED F30602-81-C-0169 F/G 17/8 NL



MICROCOPY RESOLUTION TEST CHART
DS-1963-A

AD-A195 910

(4)

RADC-TR-87-180
Final Technical Report
November 1987

ON HYPOTHESIS TESTING IN DISTRIBUTED SENSOR NETWORKS

Syracuse University

Zeineddine Chair and Pramod K. Varshney

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

ROME AIR DEVELOPMENT CENTER
Air Force Systems Command
Griffiss Air Force Base, NY 13441-5700

This report has been reviewed by the RADC Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

RADC-TR-87-180 has been reviewed and is approved for publication.

APPROVED:

VINCENT C. VANNICOLA
Project Engineer

APPROVED:

FRANK J. REHM
Technical Director
Directorate of Surveillance

FOR THE COMMANDER:

JOHN A. RITZ
Directorate of Plans & Programs

If your address has changed or if you wish to be removed from the RADC mailing list, or if the addressee is no longer employed by your organization, please notify RADC (OCTS) Griffiss AFB NY 13441-5700. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS: N/A
2a. SECURITY CLASSIFICATION AUTHORITY: N/A	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE: N/A	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): N/A	5. MONITORING ORGANIZATION REPORT NUMBER(S): RADC-TR-87-180
6a. NAME OF PERFORMING ORGANIZATION: Syracuse University; 6b. OFFICE SYMBOL (If applicable):	7a. NAME OF MONITORING ORGANIZATION: Rome Air Development Center (OCTS)
6c. ADDRESS (City, State, and ZIP Code): Department of Electrical and Computer Engineering, Link Hall, Syracuse NY 13244-1240	7b. ADDRESS (City, State, and ZIP Code): Griffiss AFB NY 13441-5700
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: Rome Air Development Center; 8b. OFFICE SYMBOL (If applicable): OCTS	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: F30602-81-C-0169
8c. ADDRESS (City, State, and ZIP Code): Griffiss AFB NY 13441-5700	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO 61102F; PROJECT NO 2305; TASK NO J8; WORK UNIT ACCESSION NO PC

11. TITLE (Include Security Classification)
ON HYPOTHESIS TESTING IN DISTRIBUTED SENSOR NETWORKS

12. PERSONAL AUTHOR(S)
Zeineddine Chair, Pramod K. Varshney

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM Jun 84 TO Dec 86	November 1987	130

16. SUPPLEMENTARY NOTATION
N/A

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP
09	03	
17	04	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Fusion, Detection, Surveillance, Estimation, Remote Receivers

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

In this report, some hypothesis testing problems in distributed sensor networks are considered. Optimum data fusion rules are obtained when the decision rules at the detectors are known. The distributed hypothesis testing problem with a distributed data fusion is solved using the Bayesian as well as the Neyman-Pearson approach. The decentralized Neyman-Pearson hypothesis testing problem and the sequential hypothesis testing problem for a tandem topology network are investigated. The distributed sequential probability ratio test problem is also studied. In all these problems, optimal strategies at each site and at each time stage are obtained.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL: Vincent C. Vannicola	22b. TELEPHONE (Include Area Code): (315) 330-4437; 22c. OFFICE SYMBOL: RADC (OCTS)

TABLE OF CONTENTS

Page

CHAPTER ONE

INTRODUCTION

1.1. Overview

Traditional surveillance and communication systems use a single sensor such as a radar or a sonar for the detection, identification and tracking of targets. In these systems, complete sensor observations are available at a central location and classical hypothesis testing and estimation procedures are employed for signal processing [1]. There is an increasing interest in simultaneously employing several sensors and sensing techniques, such as, sonics, microwave, infra-red and X-ray sensors. The basic goal of such multiple sensor systems is to improve system performance, e.g., reliability or speed. This can be achieved by properly combining the information obtained from the various sensors. Other factors which have necessitated the use of multiple sensors are the increase in the number of targets under consideration and the increase in required coverage.

The decomposition of processing is essential for controlling the complexity of computation. Central computation is just too costly in both memory and time. Moreover, the distribution of computation may enable us to use parallel processing. This will reduce the computation time which often grows exponentially. Furthermore, for large-scale complex systems, a distributed implementation is almost essential. Also, as the external environment changes, adaptation to change is easier in distributed processing systems. In a complex information system comprising of a very large number of sensors and a large volume of information, a central processor will require a very large bandwidth, and therefore a distributed implementation will be much more attractive.

Distributed processing of signals is a natural way to treat the problem of hypothesis testing when there are many sensors located at different geographical sites. In general, classical statistical decision theory can not be used to solve problems that fall into a distributed decision-making framework. The mathematical tools that we have, usually assume a centralized configuration and can be employed to handle the hypothesis testing problems for systems shown in Figures 1.1 and 1.2. Note that in the system of Figure 1.2, raw data from the multiple sensors is transmitted to the data fusion center for centralized processing. Organizing a community of decision makers to perform a global task is a challenging problem. Even problems that are trivial in a centralized setting become very difficult in a distributed environment. To illustrate the difficulties, consider the binary Bayesian hypothesis

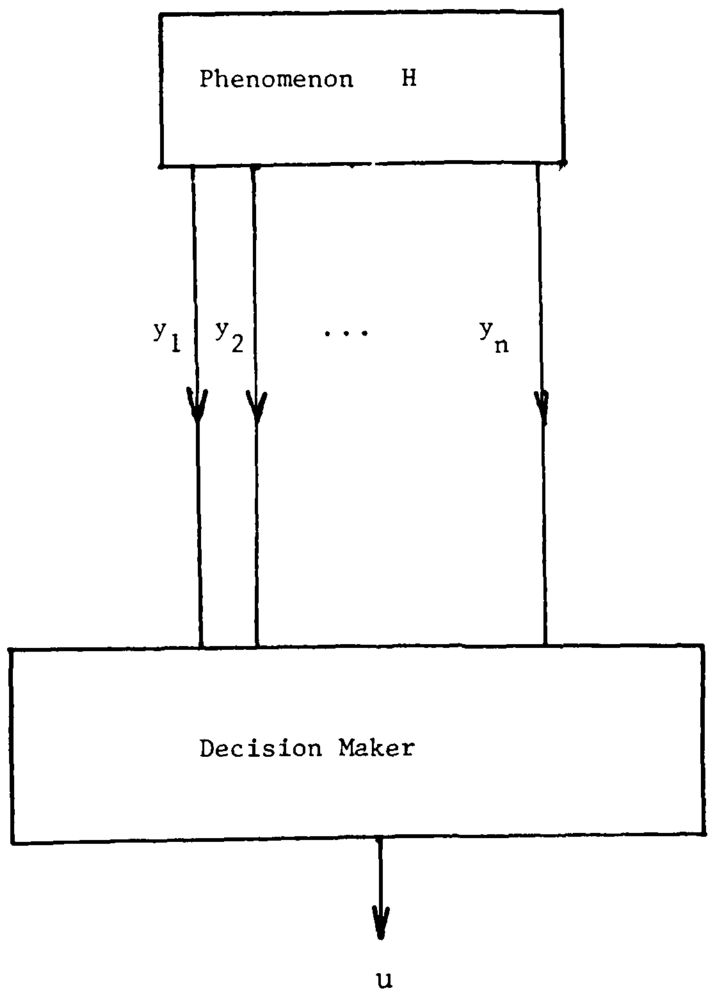


Figure 1.1 Single Sensor Detection System with Central Computation.

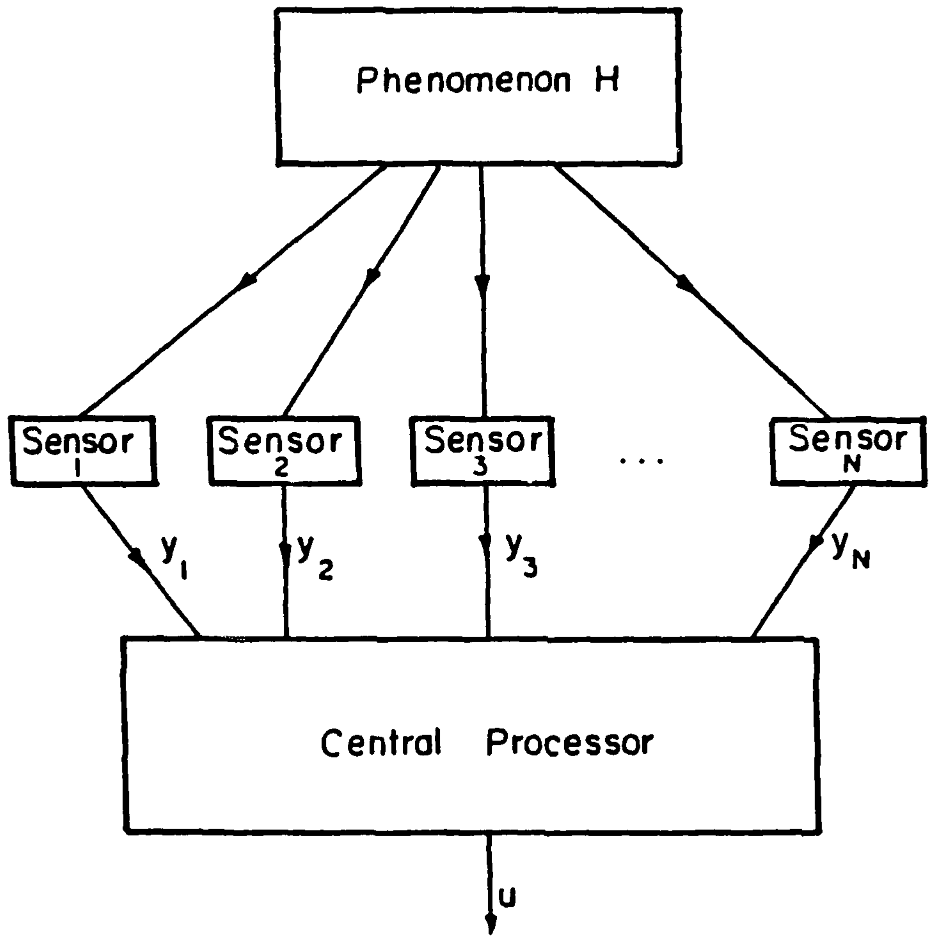


Figure 1.2 Distributed Sensor System with Central Computation.

testing problem for the system shown in Figure 1.1. The decision rule at the data fusion center is obtained using the well known result in classical detection theory [1]. The decision rule is a likelihood ratio test (LRT), where the value of the threshold is determined by the a priori probabilities and costs. In the distributed configuration as shown in Figure 1.3, the objective of the network consisting of all the decision makers is to minimize a joint cost function. Tenney and Sandell [2] have solved the problem and have shown that each decision maker implements a local likelihood ratio test, However, the thresholds are coupled and are obtained by solving a set of coupled nonlinear equations. In most cases, the solution of these nonlinear equations is a difficult matter.

The goal in this Report , is to consider some distributed signal detection problems where the signal processing is not centralized. The detection network consists of a multitude of geographically distributed sensors which collect observations from the environment. The problem is to design a distributed detection system that processes the noisy observations from these geographically distributed sensors and fuses the partially processed data to perform a global task. The objective of the work, reported in this Report , is to solve problems that will extend the

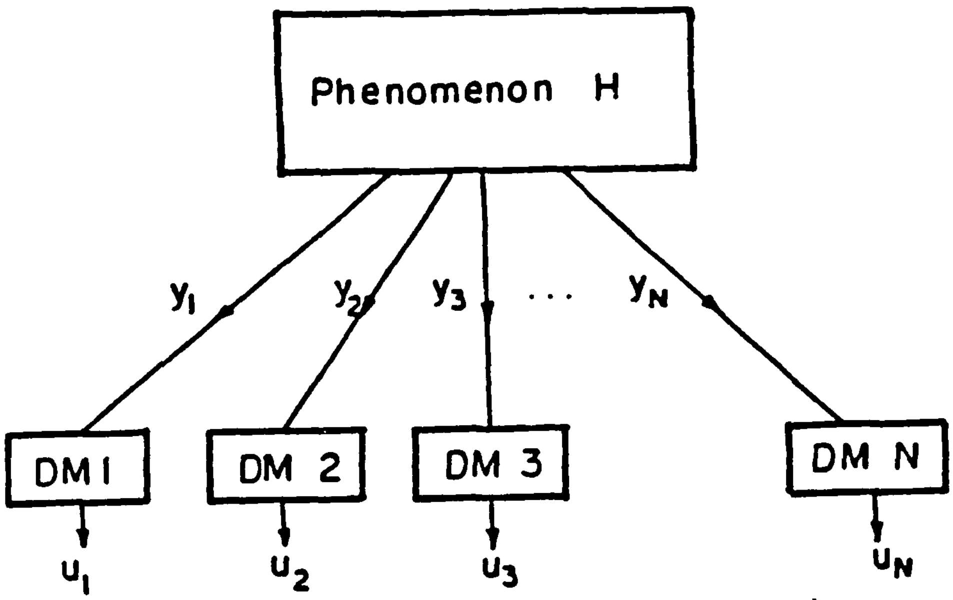


Figure 1.3 Distributed Detection System with Local Decisions.

boundaries of classical detection theory to a distributed framework. In the next section, we discuss the previous work on the subject that has been reported in the literature.

1.2. Previous Work

Some recent work on the detection problem with multiple sensors has been reported in the literature. Tenney and Sandell [2], in their important effort, have extended the classical binary Bayesian decision theory to the case of two distributed sensors. This extension does not yield results which are straightforward extensions of the classical results. The computation of detector thresholds at individual detectors is usually coupled. Further work along these lines has been performed in [2-9]. Sadjadi [3], extended the results of [2] to the case of N decentralized sensors and M hypotheses. Lauer and Sandell [4], considered the Bayesian detection of signal waveforms in the presence of noise. Ekchian and Tenney [5] formulated the Bayesian detection problem for various distributed sensor network topologies. Each member of a team of decision agents receives a conditionally independent observation about some discrete hypothesis. The decision makers seek to optimize a team cost function by making discrete decisions which are then transmitted to other decision makers depending upon the topology. They derive the optimal decision rules for each decision maker which turn out to be LRT's. Ekchian and Tenney in [6], presented a general recursive methodology for computing the optimal decision rules for each decision maker for a tandem topology. This is achieved by reformulating the stochastic optimal control problem as an equivalent deterministic one. The problem is then decomposed into subproblems which can be recursively solved using dynamic programming techniques. Kushner and Pacut [8] performed a simulation study for a distributed detection problem using exponential distributions. Tsitsiklis and Athans [9] demonstrated that, in general, distributed hypothesis testing problems are N-P complete. Their research provided theoretical evidence regarding the inherent complexity of solving optimal distributed decision problems as compared to their centralized counterparts.

The decentralized sequential detection problem has been investigated in [10-12]. In [10], Teneketzis formulated and solved a decentralized version of the Wald's problem [13]. In his model, each detector is given the flexibility of either stopping and making a decision or continuing to take more observations. He showed that the person to person optimal policies of the detectors are described by thresholds which are coupled. More specifically, the thresholds of detector i at any instant of time depend on the thresholds of detector j, $j=1,2,\ldots,N$, $j \neq i$, at all past, present and future times. For a two-detector N-stage detection system, he showed that the thresholds are determined by solving a set of 4N-2 equations in 4N-2 unknowns. Hashemipour and Rhodes [11] solved an N-stage, two-detector decentralized sequential hypothesis testing problem. In their model, each detector takes an observation and makes a binary decision which is sent to the data fusion center. The fusion center is given the option of either stopping and making a decision as to which hypothesis is true or continuing to the next time stage. It is shown that at each time instant, the optimal local strategies are LRT's. Furthermore, it is shown that the local decisions depend not only on the present and past observations, but on the past local decisions as well. In [12], the decentralized quickest detection problem has been considered. Tsitsiklis [14], considered a decentralized detection problem in which a number of identical sensors transmit a binary function of their observations to a fusion center which then decides which one of the two alternative hypothesis is true. He showed that, when the number of sensors grows to infinity, optimality is not lost if we constrain the sensors to use the same decision rule in deciding what to transmit. Recently Hoballah [15] solved various problems in distributed hypothesis testing. He solved the distributed detection problem with data fusion using both the Bayesian and Neyman Pearson approaches. Other related work appeared in [16, 21].

1.3. Report Organization

In this Report, we consider some distributed detection problems. We will derive the optimal decision rules for the various distributed sensor network topologies using a variety of optimality criteria.

In Chapter Two, we present the derivation of the optimal decision rule at the data fusion center for the distributed detection problem. We use two configurations for the data fusion center, one centralized and the other distributed. We assume that the local decisions and the thresholds of all the detectors are known a priori.

In Chapter Three, we solve the problem of distributed Bayesian hypothesis testing with distributed data fusion. We derive the decision rules both at the detectors and at the data fusion centers. An example is presented to illustrate the results of this chapter.

In Chapter Four, we solve the problem of distributed Neyman-Pearson hypothesis testing for various network topologies. First, we formulate and solve the distributed detection problem for a two-detector tandem topology. Then we extend the results in two directions. First, we solve the problem of distributed Neyman-Pearson hypothesis testing with distributed data fusion. Then, we consider the problem of distributed Neyman-Pearson hypothesis testing for a N-detector tandem topology. An example is presented to illustrate the results of this chapter.

In Chapter Five, we present the derivation of the optimal decision rule for a decentralized sequential detection problem using the Neyman-Pearson approach. Furthermore, we solve the Bayesian sequential hypothesis testing problem for a tandem topology configuration.

In Chapter Six, a summary of results, conclusions, and suggestions for future research are presented.

CHAPTER TWO

OPTIMAL DATA FUSION IN MULTIPLE SENSOR DETECTION SYSTEMS

2.1 Introduction

Tenney and Sandell [2], have treated the Bayesian detection problem with distributed sensors. They only considered the design of decision rules at the individual sensors and did not consider the design of data fusion algorithms. In this chapter, we investigate the latter problem, i.e., the design of data fusion algorithms when the decision rules at the individual sensors are known. We consider two major options for signal processing in multi-sensor systems with data fusion. In the first option, all sensors report to a centralized computing facility as shown in Figure 2.1. This is the conventional configuration for a multi-sensor system with data fusion. Some signal processing is done at the local sensor, and partial results are transmitted to the data fusion center for further processing. Global results can then be obtained at the data fusion center. This option is attractive for many applications due to the communication bandwidth constraints found in practice. The principal weakness of this configuration is its vulnerability to the loss of the data fusion center. The second option is to have a distributed data fusion con-

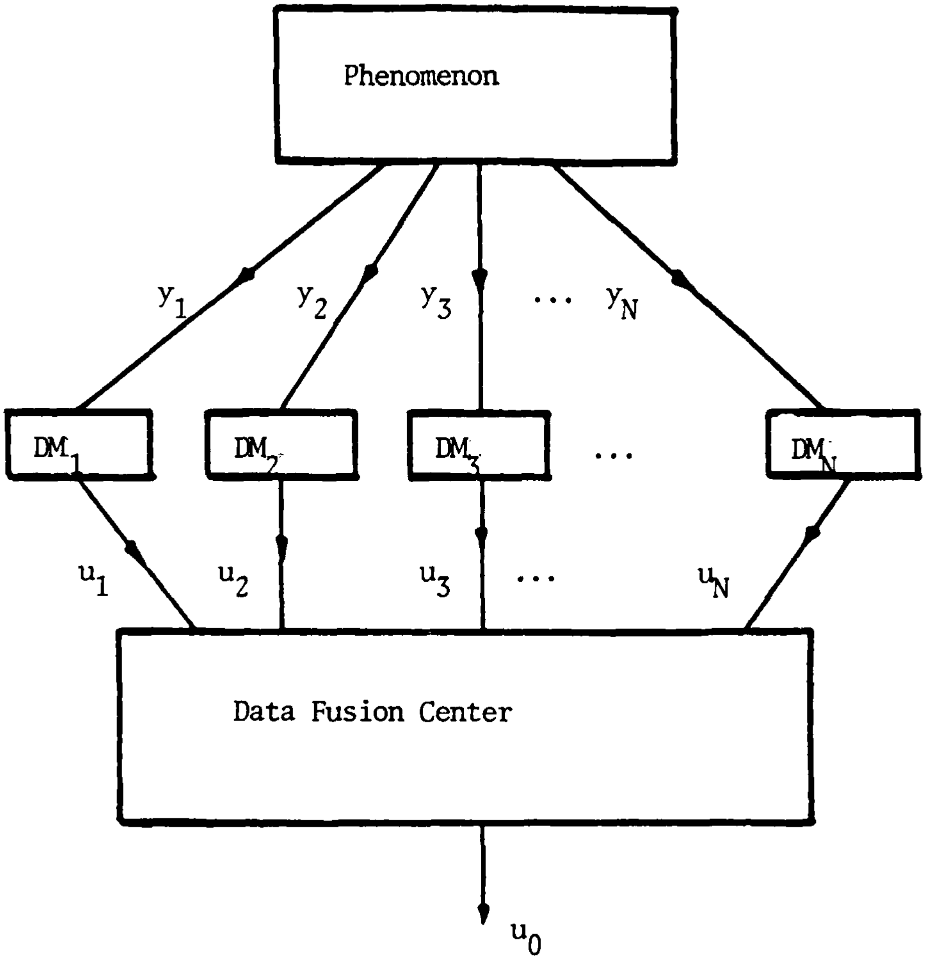


Figure 2.1 Distributed Detection System with Centralized Data Fusion Center.

figuration as shown in Figure 2.2 (only three sensors are shown for the sake of clarity). In this option, some signal processing is done locally at the sensor and partial results are transmitted to all of the local data fusion centers for further processing. Global results can then be obtained at all of the data fusion centers. This option is attractive for many applications because it is more robust than the centralized configuration in terms of its vulnerability to the loss of the data fusion center. Furthermore, each decision maker in the distributed configuration has access to the actual raw data at that site which may result in a performance enhancement.

In this chapter, we consider the problem of optimal decision combining in multiple sensor detection systems. In Section 2.2, we derive the optimum fusion rule for the centralized configuration when the individual detector decision rules are known. The combining rule turns out to be a function of the probability of false alarm and probability of miss of individual detectors, i.e., the reliability of individual detectors. An example is presented to illustrate the result. In Section 2.3, we derive the optimum fusion rules for the distributed configuration when the individual detector decision rules are known. The combining rule at a fusion site turns out to be a function of the probability of false alarm and the probability of miss of the incoming decisions from the other decision makers, and its

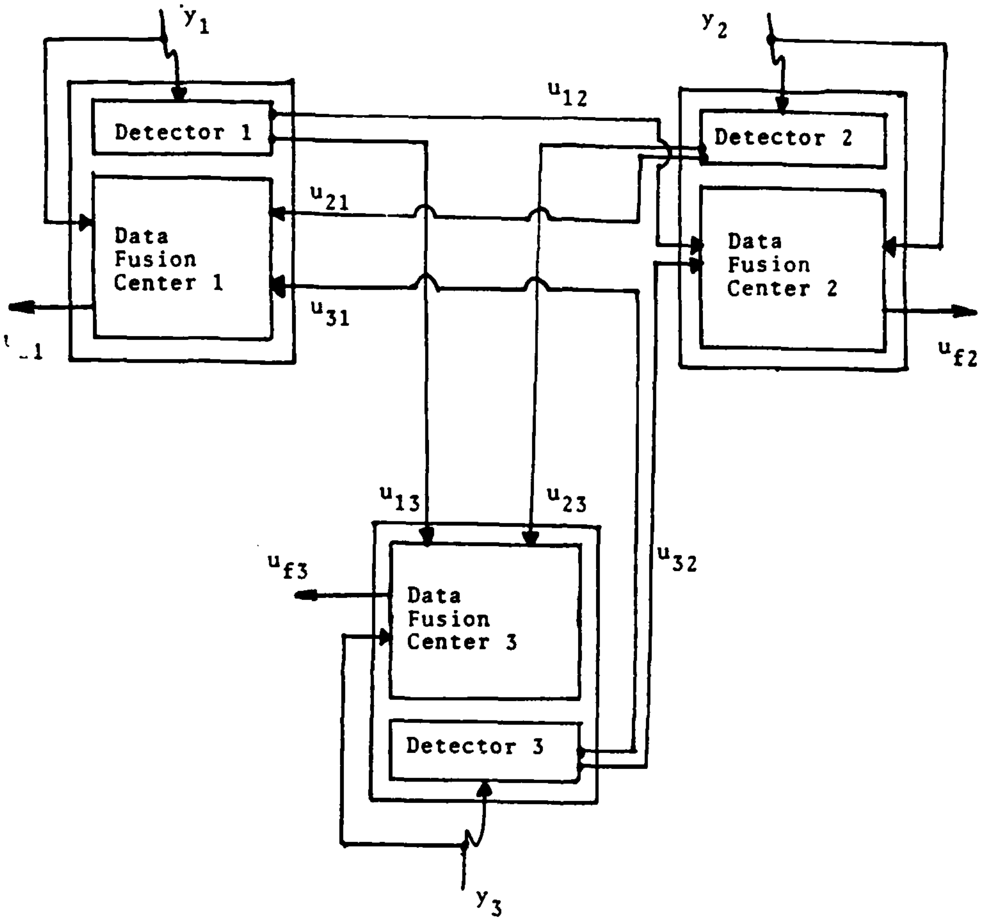


Figure 2.2 Distributed Detection with Distributed Data Fusion Configuration for Three Decision Makers.

own observation. An example is presented to illustrate the result. Finally the results of this chapter are discussed in Section 2.4.

2.2. Optimal Centralized Data Fusion

2.2.1. Problem Statement and Solution

Let us consider a binary hypothesis testing problem with the following two hypotheses:

$H_o \triangleq$ signal is absent, and

$H_1 \triangleq$ signal is present.

The a priori probabilities of the two hypotheses are denoted by $P(H_o) = P_o$ and $P(H_1) = P_1$. As shown in Figure 2.1, we assume that there are N detectors, D_i, i=1,...,N, and the observations at each D_i are denoted by y_i, i=1,...,N. We will further assume that the observations at the individual detectors are statistically independent and that the conditional density function of the observations is denoted by $p(y_i|H_j)$; i=1,...,N; j=0,1. The independence assumption implies that

$$p(y_i|y_j, H_k) = p(y_i|H_k) \quad \text{for} \quad i \neq j \ . \tag{2-1}$$

Each detector, D_i, employs a decision rule $g_i(y_i)$ to make a decision u_i, i=1,2,...,N, where

$$u_i = \begin{cases} 0 & \text{if } H_o \text{ is declared} \\ 1 & \text{if } H_1 \text{ is declared} \ . \end{cases} \tag{2-2}$$

We denote the probabilities of false alarm and miss of each detector by P_{Fi} and P_{Mi} respectively. These error probabilities are defined as

$$P_{Fi} \triangleq \text{Prob. } [u_i = 1|H_o] \quad \text{and}$$
$$P_{Mi} \triangleq \text{Prob. } [u_i = 0|H_1] = 1-P_{Di}. \tag{2-3}$$

After processing the observations locally, the decisions u_i are transmitted to the data fusion center as shown in Figure 2.1. The centralized data fusion center determines the overall decision u_o for the system, based on the individual decisions, i.e.,

$$u_o = \gamma(u_1, u_2, \ldots, u_N) . \tag{2-4}$$

Our objective is to find the optimal data fusion rule when the individual detectors have already been designed. Data fusion rules are often implemented as "k out of N" logical functions. This means that if k or more detectors decide hypothesis H_1, then the global decision is H_1, otherwise it is H_o, i.e.,

$$u_o = \begin{cases} 1 & \text{if } u_1 + u_2 + \ldots + u_N \geq k \\ 0 & \text{otherwise ,} \end{cases} \tag{2-5}$$

where $u_i \in \{0,1\}$ for $i = 1,2,\ldots,N$. Common logical functions such as AND, OR, and majority gate are special cases of the 'k out of N" rule. In this section, we will consider a more general formulation of the data fusion problem. The data fusion problem can be viewed as a two-hypothesis detection problem with individual detector decisions being the observations. The optimum decision rule is obtained by minimizing the overall risk function R defined as follows:

$$R = E\{J(u_o, H)\} \tag{2-6}$$

where

$u_o \triangleq$ global decision,

$H \triangleq$ hypothesis .

Our aim is to minimize R with respect to all the uncertainties present ($u_1, \ldots, u_N$ and H) over all possible strategies γ, i.e.,

$$\min_{\gamma} \; E_{\underline{u},H} \, J[\gamma(\underline{u}), H] \tag{2-7}$$

where

$\underline{u} \triangleq (u_1, u_2, \ldots, u_N)$,

$u_o = \gamma(\underline{u})$, $u_o \in \{0,1\}$,

and

$J(u_o = i, H_j) \triangleq$ cost of deciding $u_o = i$ when H_j is present.

Consequently, the problem faced by the data fusion center is centralized in nature and can be solved using the classical approach. The result is presented in Theorem 2.1.

<u>Theorem 2.1</u>

Given N detectors D_i, $i=1,2,\ldots,N$, along with their associated decision rules operating at the point (P_{Di}, P_{Fi})

of the Receiver Operating Characteristic (ROC); the optimal centralized data fusion rule is given by

$$\ln \frac{P_1}{P_o} + \sum_{S_+} \ln \frac{1 - P_{Mi}}{P_{Fi}} + \sum_{S_-} \ln \frac{P_{Mi}}{1 - P_{Fi}} \underset{u_o=0}{\overset{u_o=1}{\gtrless}} \ln \frac{J_{10}-J_{00}}{J_{01}-J_{11}} \tag{2-8}$$

where

$S_+ \triangleq$ the set of all i such that $u_i = 1$,

$S_- \triangleq$ the set of all i such that $u_i = 0$

<u>Proof</u>:

The optimum data fusion rule is given by the following likelihood ratio test [1]

$$\Lambda(\underline{u}) \triangleq \frac{p(\underline{u}|H_1)}{p(\underline{u}|H_o)} \underset{u_o=0}{\overset{u_o=1}{\gtrless}} \frac{P_o(J_{10} - J_{00})}{P_1(J_{01} - J_{11})} \tag{2-9}$$

where

$J_{ij} \triangleq J(u_o = i|H = H_j)$.

Since u_i depends only on y_i,

$$p(\underline{u}|H_1) = \prod_{i=1}^{N} p(u_i|H_1) = \prod_{S_+} p(u_i{=}1|H_1) \prod_{S_-} p(u_i{=}0|H_1)$$

$$= \prod_{S_+} (1 - P_{Mi}) \prod_{S_-} P_{Mi} . \tag{2-10}$$

In a similar manner,

$$p(\underline{u}|H_o) = \prod_{i=1}^{N} p(u_i|H_o) = \prod_{S_+} p(u_i=1|H_o) \prod_{S_-} p(u_i=0|H_o)$$

$$= \prod_{S_+} P_{Fi} \prod_{S_-} (1 - P_{Fi}) \tag{2-11}$$

Substituting (2-10) and (2-11) into (2-9) and taking the logarithm, we get

$$\ln \frac{p(\underline{u}|H_1)}{p(\underline{u}|H_o)} = \sum_{S_+} \ln \frac{1 - P_{Mi}}{P_{Fi}} + \sum_{S_-} \ln \frac{P_{Mi}}{1-P_{Fi}} \underset{u_o=0}{\overset{u_o=1}{\gtrless}} \ln \frac{P_o(J_{10}-J_{00})}{P_1(J_{01}-J_{11})} \tag{2-12}$$

A manipulation of (2-12) yields the desired result which completes the proof. Q.E.D.

Note that, we may also express the data fusion rule as

$$u_o = \gamma(u_1,u_2,\ldots,u_N) = \begin{cases} 1 & \text{if } a_o + \sum_{i=1}^{N} a_i(2u_i-1) > \ln \dfrac{J_{10}-J_{00}}{J_{01}-J_{11}} \\ 0 & \text{otherwise} \end{cases} \tag{2-13}$$

where

$$a_o = \ln \frac{P_1}{P_o}$$

$$a_i = \ln \frac{1 - P_{Mi}}{P_{Fi}} \quad \text{if} \quad u_i = 1 \tag{2-14}$$

and

$$a_i = \ln \frac{1 - P_{Fi}}{P_{Mi}} \quad \text{if} \quad u_i = 0 \ .$$

The optimum data fusion rule can be implemented as shown in Figure 2.3. As we can observe, individual detector decisions are weighted according to their reliability, i.e., the weights

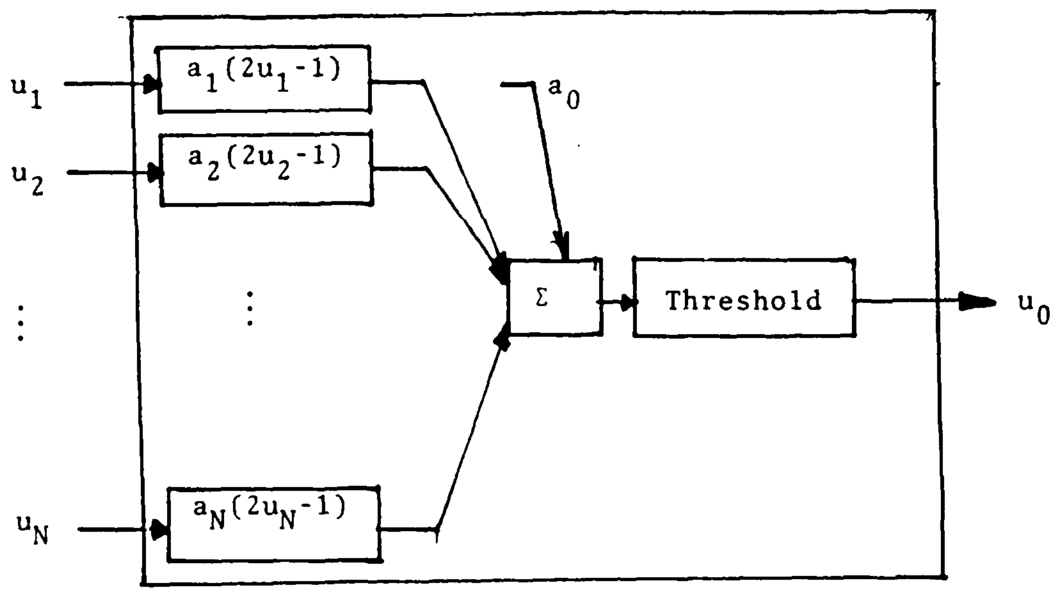


Figure 2.3 Optimum Centralized Data Fusion Center Structure.

are functions of probability of false alarm and probability of miss. The data fusion center structure obtained here attempts to optimally use the individual detector decisions by forming a weighted sum and then comparing it to a threshold.

2.2.2. Example

We consider a simple binary hypothesis problem with two detectors D_1 and D_2. The probabilities of false alarm are given to be $P_{F1} = P_{F2} = \frac{1}{2}$, and the probabilities of miss are given to be $P_{M1} = P_{M2} = \frac{1}{4}$. We assume the hypotheses to be equally likely, i.e., $P_0 = P_1 = \frac{1}{2}$ and choose the minimum probability of error cost function, i.e., $J_{00} = J_{11} = 1$, $J_{01} = J_{10} = 0$. Now, the optimum data fusion rule is obtained using Theorem 2.1. We consider the case when $u_1 = u_2 = 1$. The log-likelihood ratio is given by

$$\ln \frac{p(H_1|\underline{u})}{p(H_o|\underline{u})} = 2 \ln \frac{3}{2} \; . \tag{2-15}$$

Since it is greater than the threshold zero, the decision is H_1. The results for the other combinations of u_1 and u_2 are summarized below:

u_1	u_2	Global Decision u_o
1	1	H_1
0	1	H_0
1	0	H_0
0	0	H_0

Note that the fusion rule turns out to be the AND rule. Depending on the values of P_0, P_1, P_{Mi}, P_{Fi} and the costs, the fusion rule can be some other logical function. The ROC curve for this example is shown in Figure 2.6.

2.3. Optimum Distributed Data Fusion

2.3.1 Problem Statement and Solution

In this section, we consider the same binary hypothesis testing problem as stated in Section 2.2.1 for the distributed data fusion configuration. In this system, after processing the observations locally, i.e., at the level of the detectors, the decisions u_i are transmitted to all the local data fusion centers for further processing as shown in Figure 2.4. Now, we describe the sequence of operations at the ith sensor, i.e., at the detector D_i, and the data fusion center i, DF_i. First, D_i takes an observation y_i and based on y_i makes a decision u_i. Then, the decision $u_i \epsilon \{0,1\}$ is transmitted to all the data fusion centers DF_k, k=1,...,N, $k \neq i$. After receiving the decisions u_k, k=1,...,N, $k \neq i$, from the other detectors, DF_i makes a final decision u_{fi} based on $\{u_k, k=1,\ldots,N, k \neq i\}$ and y_i. In the distributed data fusion problem, our goal is to derive the optimum decision rules to obtain the final decisions u_{fi}, i=1,...,N. This problem can again be viewed as a two-hypothesis detection problem at each site. The individual detector decisions u_i's and the measurement y_i at each site are the set of observations, i.e., the data fusion center DF_i receives the set of observations

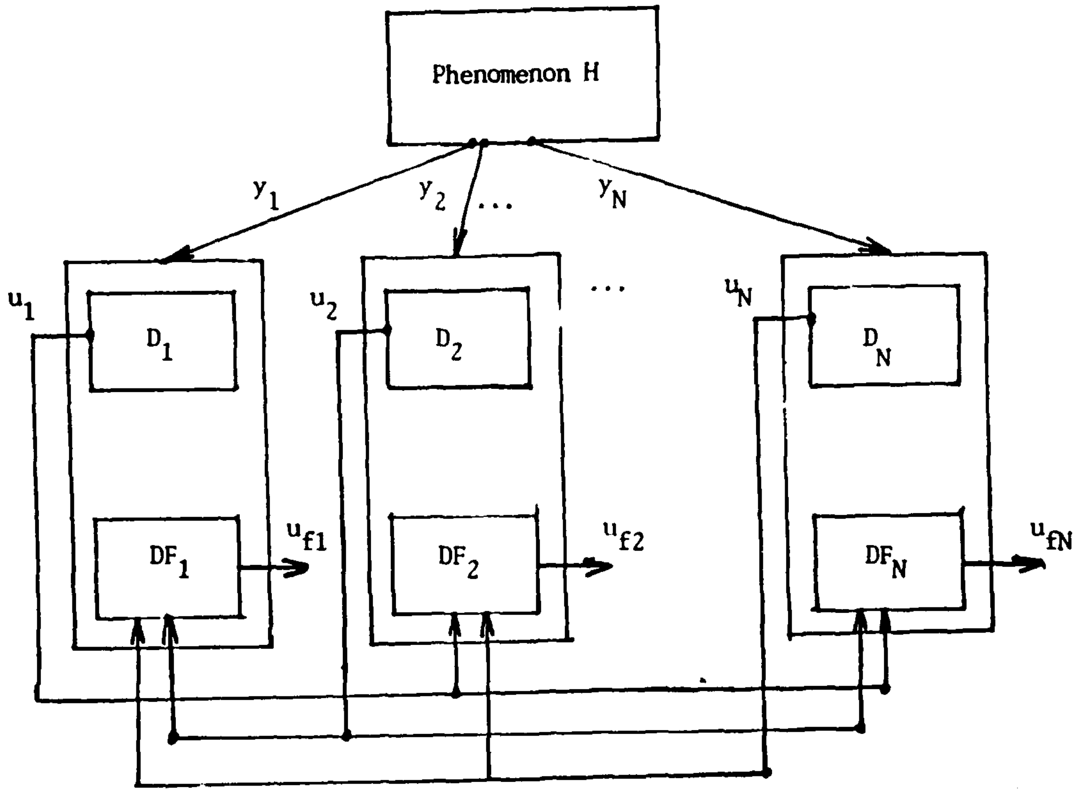


Figure 2.4 Distributed Detection System with Distributed Data Fusion Centers.

$I_i = \{ u_k, k=1,\ldots,N, k \neq i$ and $y_i\}$. Based on I_i, DF_i makes a final decision u_{fi}, i.e.,

$$u_{fi} = \gamma_i(I_i) \quad . \tag{2-16}$$

We wish to find the optimal data fusion rule at DF_i, $i = 1,\ldots,N$. The risk function R_i, $i=1,\ldots,N$, at DF_i is defined as

$$R_i \triangleq E[J(u_{fi}, H)] \ . \tag{2-17}$$

Our aim is to minimize R_i with respect to all uncertainties present $\{I_i, H\}$ over all possible strategies γ_i, i.e.,

$$\min_{\gamma_i} E_{I_i,H} J(\gamma_i, H) \ .$$

The optimal data fusion strategy at DF_i for the distributed data fusion configuration is given in the following theorem.

<u>Theorem 2.2</u>

Given N detectors D_i along with their associated decision rules and operating points (P_{Di}, P_{Fi}), and N data fusion centers DF_i, $i=1,\ldots,N$; the optimal distributed data fusion rule at the site i is given by

$$\ln \frac{P_1}{P_0} + \sum_{S_+^i} \ln \frac{1-P_{Mj}}{P_{Fj}} + \sum_{S_-^i} \ln \frac{P_{Mj}}{1-P_{Fj}} + \ln \frac{p(y_i|H_1)}{p(y_i|H_0)} \underset{u_{fi}=0}{\overset{u_{fi}=1}{\gtrless}} \ln \frac{J_{10} - J_{00}}{J_{01}-J_{11}} \tag{2-18}$$

where

$$S_+^i \triangleq \text{ the set of all j such that } u_j = 1, \; j \neq i.$$

$$S_-^i \triangleq \text{ the set of all j such that } u_j = 0, \; j \neq i. \tag{2-19}$$

Proof:

The objective is to find the strategy γ_i so as to minimize E $\{J(u_{fi}, H)\}$, i=1,...,N. This is a classical detection problem, the solution is given by the following likelihood ratio test

$$\Lambda(\underline{u}^i, y_i) \triangleq \frac{p(\underline{u}^i, y_i | H_1)}{p(\underline{u}^i, y_i | H_o)} \underset{u_{fi}=0}{\overset{u_{fi}=1}{\gtrless}} \frac{P_0(J_{10}-J_{00})}{P_1(J_{01}-J_{11})} \tag{2-20}$$

where

$$\underline{u}^i \triangleq (u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_N) \; . \tag{2-21}$$

We have

$$p(\underline{u}^i, y_i | H) = p(u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_N, y_i | H) \tag{2-22}$$

We assume that u_k depends only on y_k, i.e.,

$$u_k = g_k(y_k) \tag{2-23}$$

Then, due to the independence assumption (2-22) becomes

$$p(\underline{u}^i, y_i | H) = \prod_{\substack{k=1 \\ k \neq i}}^{N} p(u_k | H) \cdot p(y_i | H) \tag{2-24}$$

Using (2-24) in (2-20), we have

$$\prod_{\substack{k=1\\k\neq i}}^{N} \frac{p(u_k|H_1)}{p(u_k|H_o)} \cdot \frac{p(y_i|H_1)}{p(y_i|H_o)} \underset{u_{fi}=0}{\overset{u_{fi}=1}{\gtrless}} \frac{J_{10}-J_{00}}{J_{01}-J_{11}} \frac{P_0}{P_1} \qquad (2\text{-}25)$$

Taking the logarithm of both sides of (2-25), and proceeding as in the proof of Theorem 2.1, we have

$$\ln \frac{P_1}{P_0} + \sum_{S_+^i} \ln \frac{1 - P_{Mj}}{P_{Fj}} + \sum_{S_-^i} \ln \frac{P_{Mj}}{1 - P_{Fj}}$$

$$+ \ln \frac{p(y_i|H_1)}{p(y_i|H_0)} \underset{u_{fi}=0}{\overset{u_{fi}=1}{\gtrless}} \ln \frac{J_{10}-J_{00}}{J_{01}-J_{11}} \qquad (2\text{-}26)$$

which completes the proof. Q.E.D.

The data fusion rule at the site i can be implemented as shown in Fig. 2.5, where

$$a_o = \ln \frac{P_1}{P_0}$$

$$a_i = \ln \frac{1 - P_{Mi}}{P_{Fi}}, \text{if} \quad u_i = 1,$$

and

$$a_i = \ln \frac{1 - P_{Fi}}{P_{Mi}}, \text{if} \quad u_i = 0, \qquad (2\text{-}27)$$

for $j \neq i$.

2.3.2. Example

We consider the same example as in Section 2.2.2,

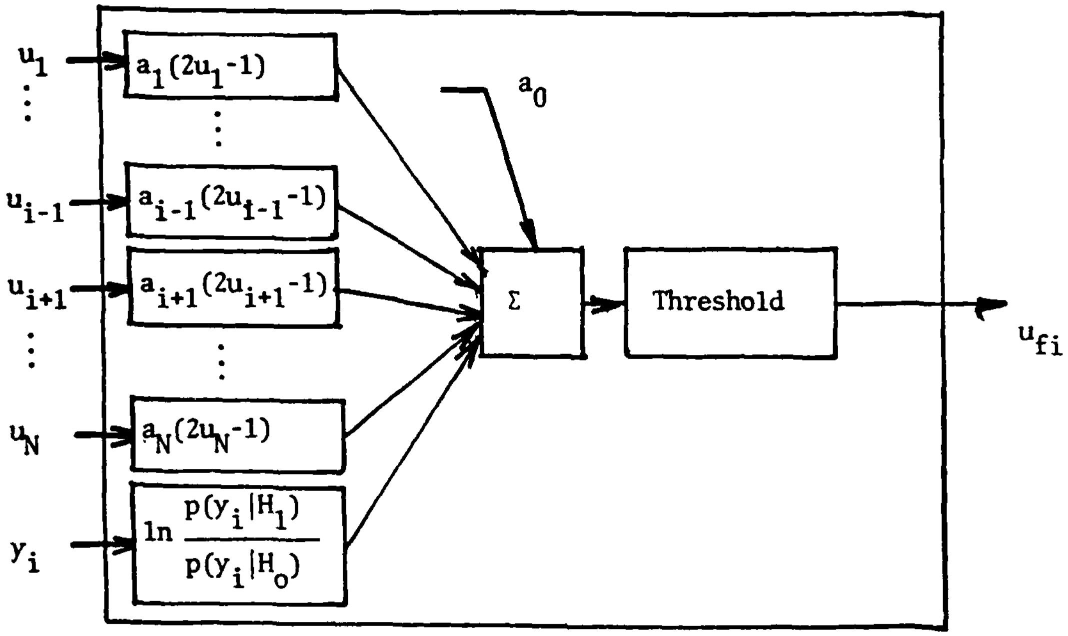


Figure 2.5 Optimum Distributed Data Fusion Center Structure at The Site i.

the given data is summarized below:

P_{F1}	P_{F2}	P_{M1}	P_{M2}	J_{10}	J_{01}	J_{00}	J_{11}	P_0	P_1
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	1	1	0	0	$\frac{1}{2}$	$\frac{1}{2}$

We further assume that the conditional densities $p(y_i|H_j)$, $i=1,2,;\ j=0,1$, are exponential, i.e.,

$$p(y_i|H_o) = \lambda \exp(-\lambda y_i)$$

$$p(y_i|H_1) = 2\lambda \exp(-2\lambda y_i) . \quad \lambda>0,\ i=1,2$$

for $\quad y_i \geq 0$,

and

$$p(y_i|H_j) = 0. \quad i=1,2 , \quad j=0,1 \quad \text{elsewhere.} \tag{2-28}$$

Using Theorem 2.2, we obtain the optimum data fusion rule for this problem. We first consider the case when $u_2=1$, and in this case we obtain the data fusion rule at the site 1. We have

$$\ln[2 \exp(-\lambda y_1)] + \ln \frac{1 - P_{M2}}{P_{F2}} + \ln \frac{P_1}{P_0}$$

$$\underset{u_{f1}=0}{\overset{u_{f1}=1}{\gtrless}} \ln \frac{J_{10} - J_{00}}{J_{01} - J_{11}} . \tag{2-29}$$

Substituting the numerical values in (2-29), we get

$$\ln[2\exp(-\lambda y_1)] + \ln\frac{3}{2} \underset{u_{f1}=0}{\overset{u_{f1}=1}{\gtrless}} 0 \qquad (2\text{-}30)$$

From this, we obtain

$$y_1 \underset{u_{f1}=1}{\overset{u_{f1}=0}{\gtrless}} \frac{\ln 3}{\lambda} \triangleq t_1(u_2 = 1) \qquad (2\text{-}31)$$

For the case when $u_2 = 0$, we obtain as before

$$y_1 \underset{u_{f1}=1}{\overset{u_{f1}=0}{\gtrless}} 0 \triangleq t_1(u_2 = 0) \qquad (2\text{-}32)$$

Similarly for detector 2, we obtain the following likelihood ratio test

$$y_2 \underset{u_{f2}=1}{\overset{u_{f2}=0}{\gtrless}} \frac{\ln 3}{\lambda} \triangleq t_2(u_1 = 1) \qquad (2\text{-}33)$$

and

$$y_2 \underset{u_{f2}=1}{\overset{u_{f2}=0}{\gtrless}} 0 \triangleq t_2(u_1 = 0) \qquad (2\text{-}34)$$

Note that in the case of distributed data fusion configuration, we need to have the actual value of the received data $\{y_i\}$ before we can make a final decision. Furthermore, the final decision depends on the value of the received local decisions $\{u_i\}$. Basically, D1 waits for the local decision u_2 from D2, then DF_1 makes a final decision based on the value of u_2 and the raw data y_1. The ROC curve for this example is shown in Figure 2.6. As expected, the performance of a distributed data fusion configuration is superior to its counterpart - the centralized data fusion configuration.

2.4. Discussion

In this chapter, we have considered the design of optimum data fusion algorithms for the signal detection problem when multiple sensors are used for surveillance and a global decision is desired. We have considered two approaches to the data fusion problem. In the first approach, a centralized optimum data fusion structure has been derived which combines the decisions from the individual decision makers while minimizing the overall risk function. Individual decisions are weighted according to their reliability, i.e., the weights are a function of the probability of miss and the probability of false alarm of the individual decision makers. In the second approach, an optimum distributed data fusion structure has been derived which combines the decisions from other decision makers and the raw observation at that site while minimizing the risk function at that site. This approach gives a better performance than the first approach since we are using the actual observation y_i instead of its quantized value u_i. If we use the hard

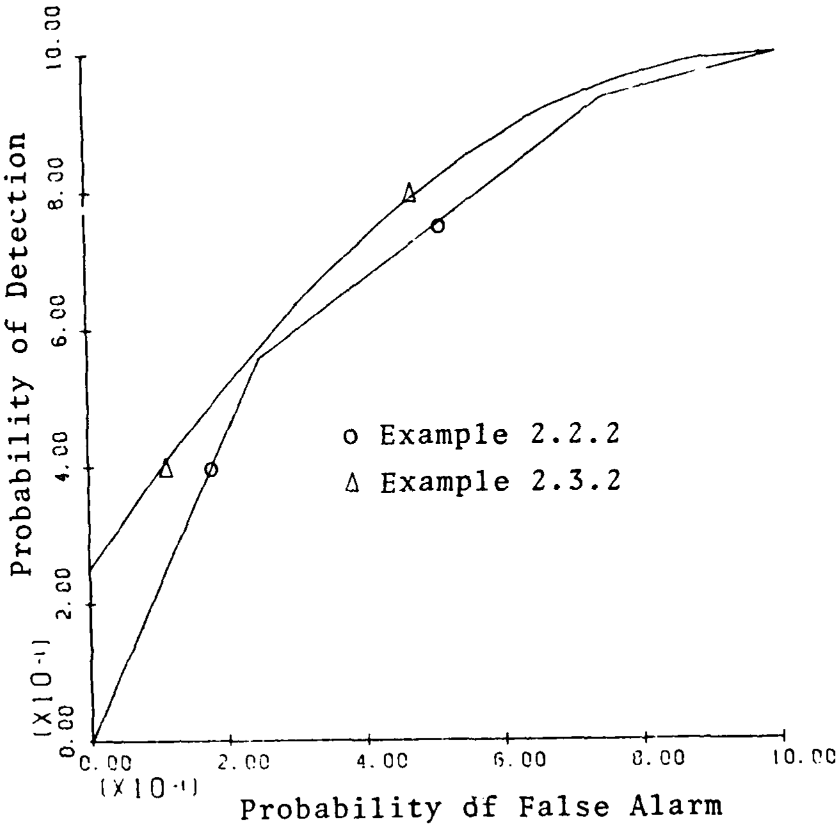


Figure 2.6. Receiver Operating Characteristic for Example 2.2.2 and 2.3.2 .

decision u_i at the fusion site i in the distributed configuration, we will obtain the same performance as in the centralized configuration. In the distributed configuration, we have a global decision (fused result) at each site which makes the system more survivable for military applications. This is achieved at the expense of added communication amongst the sensors. In the next two chapters, we will study the overall problem of distributed detection using a distributed data fusion structure.

CHAPTER THREE

DISTRIBUTED BAYESIAN HYPOTHESIS TESTING WITH DISTRIBUTED DATA FUSION

3.1. Introduction

Some recent work on the Bayesian detection problem with multiple sensors has been reported in the literature. Tenney and Sandell [2] extended the classical Bayesian decision theory to the case of distributed sensors without a data fusion center. Ekchian and Tenney [5] then formulated and solved the Bayesian hypothesis testing problem for various distributed sensor network topologies. Hoballah [15] considered the problem of distributed Bayesian hypothesis testing using a centralized data fusion center. In this chapter, we solve the problem of distributed Bayesian hypothesis testing with distributed data fusion. In Section 3.2, we formulate and solve the distributed Bayesian hypothesis testing problem with distributed data fusion for the case of N decision makers. In Section 3.3, we present the results for the case of two decision makers, and illustrate with an example. Finally, in Section 3.4, we discuss the results obtained in this chapter.

3.2. Distributed Bayesian Hypothesis Testing with Distributed Data Fusion

In this section, we consider the system as shown in Figure 3.1. The Binary hypothesis testing problem is considered and the two possible hypotheses are denoted by H_o and H_1, with given a priori probabilities $P(H_j) = P_j$, $j=0,1$. The ith decision maker, DM_i, consists of two elements - a detector D_i and a data fusion center DF_i, as shown in Figure 3.2. The decision maker i, DM_i, takes an observation y_i, $i=1,\ldots,N$, based on which the detector i, D_i, makes a set of decisions $(u_{i1}, u_{i2},\ldots, u_{ii-1}, u_{ii+1},\ldots, u_{iN})$ where u_{ij} represents the decision made by D_i and intended for DF_j. We will denote, by $\underline{u}_{it}$, this decision vector generated at D_i. As we will see subsequently, D_i employs a different decision rule while determining the different elements of $\underline{u}_{it}$. Each data fusion center j, DF_j, receives a decision vector $(u_{1j}, u_{2j},\ldots,u_{j-1j}, u_{j+1j},\ldots,u_{Nj})$ denoted by $\underline{u}_{jr}$. The decision vectors $\underline{u}_{it}$ and $\underline{u}_{ir}$ are transported over a communication network which is responsible for all the processing necessary for transmission to the appropriate DM's. For illustration, we show an explicit signalling configuration for a three decision maker network in Figure 3.3. In the above network each DM receives a decision vector of length two and transmits a decision vector of length two. In this case, the vector $\underline{u}_{kr}$ consists of two elements, e.g., $\underline{u}_{2r} \triangleq \{u_{12}, u_{32}\}$, and the vector $\underline{u}_{kt}$

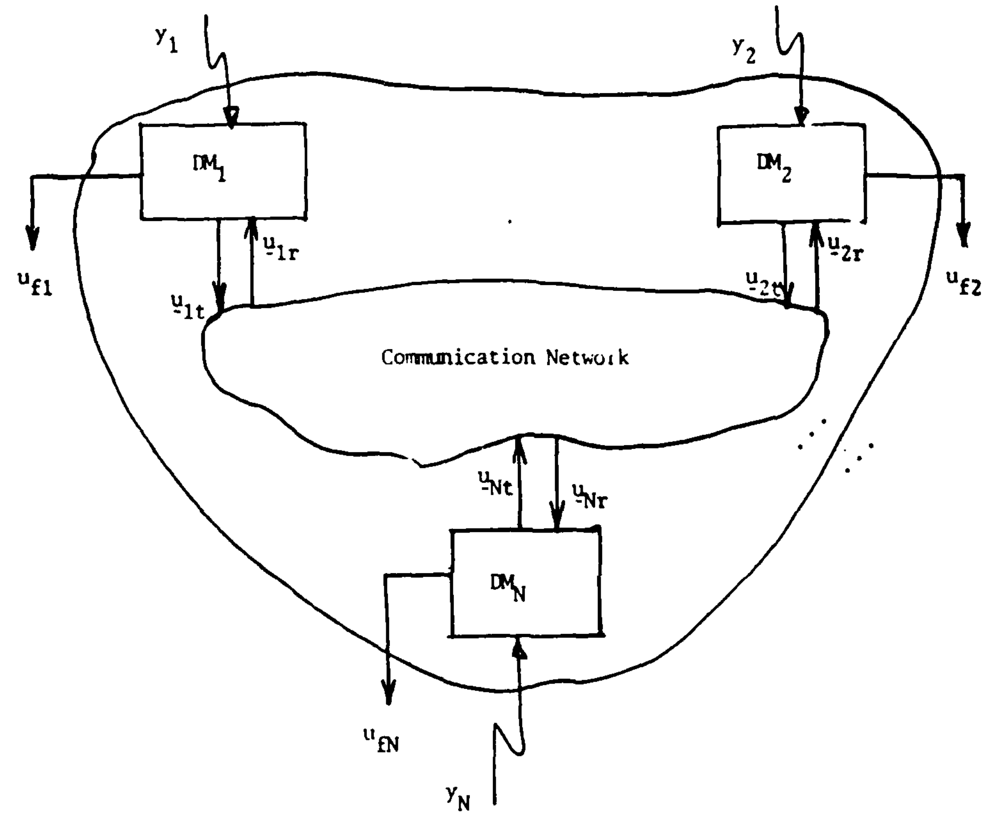


Figure 3.1 Distributed N-Sensor Detection System with Distributed Data Fusion.

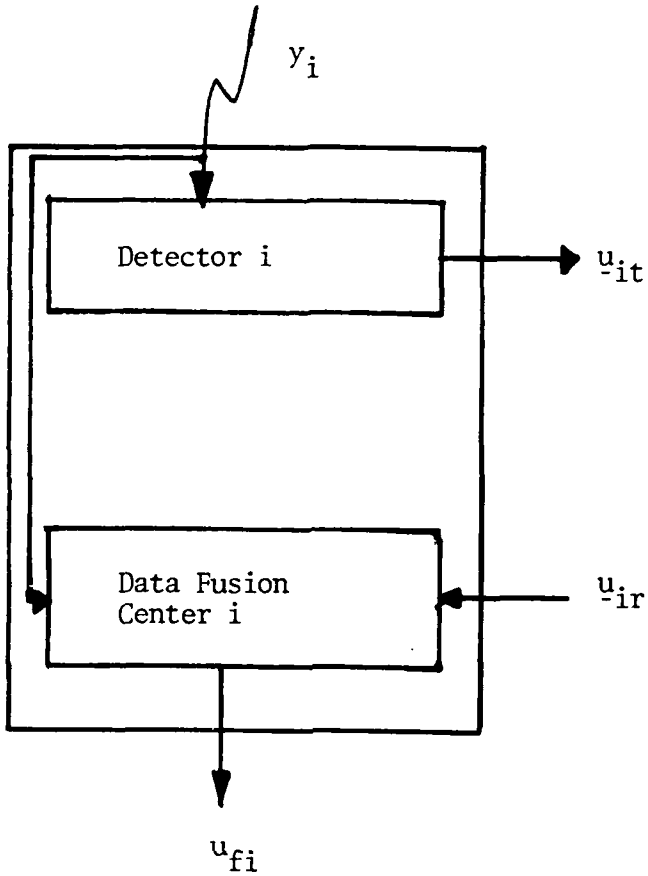


Figure 3.2 Configuration of the Decision Maker i.

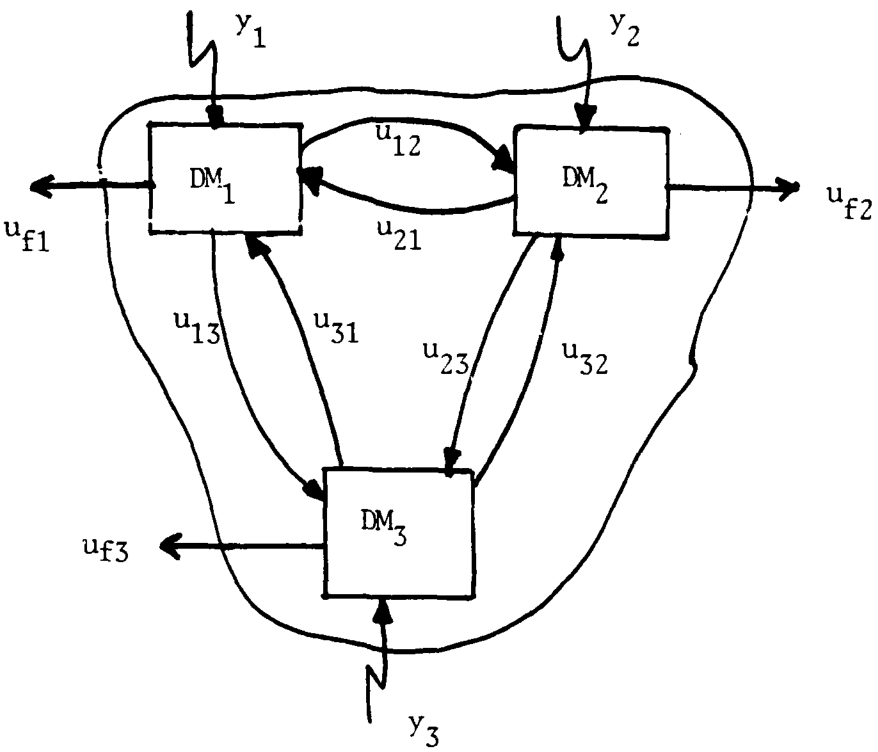


Figure 3.3 Signalling Configuration for a Three-Sensor Network with Distributed Data Fusion.

also consists of two elements, e.g., $\underline{u}_{2t} \triangleq \{u_{21}, u_{23}\}$. We further assume that all $u_{ij} \in \{0,1\}$, $i=1,\ldots,N$, $j=1,\ldots,N$. The final decision of DM_i at the site i, $i=1,\ldots,N$, is denoted by u_{fi}.

The aim in this chapter is to find the optimal decision and fusion rules at each site so as to minimize the expected value of the overall cost at each site. The problem can be formulated as follows. Consider the system configuration shown in Figure 3.1.

(A.1) DM_i takes an observation y_i, $i=1,\ldots,N$. We assume that $\{y_i, i=1,\ldots,N\}$ are conditionally independent random variables given the hypothesis H, that is,

$$p(y_1,\ldots,y_N|H) = \prod_{i=1}^{N} p(y_i|H) \tag{3-1}$$

(A.2) Based on its observation y_i, D_i makes N-1 local decisions u_{ij}, $j=1,\ldots,i-1,i+1,\ldots,N$, i.e.,

$$u_{ij} = \gamma_{ij}(y_i), \tag{3-2}$$

$i=1,\ldots,N$; $j=1,\ldots,N$; $i \neq j$.

The decision u_{ij} is an intermediate decision made at D_i whose purpose is to help DM_j in making its final decision u_{fj}.

(A.3) Based on y_i and $\underline{u}_{ir}$, DF_i makes a final decision u_{fi}, $i=1,\ldots,N$, i.e.,

$$u_{fi} = \gamma_{fi}(y_i, \underline{u}_{ir}) \; . \tag{3-3}$$

The decision u_{fi} is the final decision of DM_i at the site i.

(A.4) The cost incurred in making the decision u_{fj} at DM_j is denoted by $J(u_{fj}, H)$ where H is the true hypothesis. We assume that

$$J(u_{fj} = i, H_k) \geq J(u_{fj} = i \; , H_i), \quad k \neq i$$

$$j=1,\ldots,N \; ; \; i=0,1 \; ; \; k=0,1 \; . \tag{3-4}$$

All these inequalities are reasonable because they imply that an error is more costly than no error.

(A.5) Also, we assume that

$$p(u_{fi} = 0 | \underline{u}^k_{ir} \; , \; u_{ki} = 0, \; y_i)$$

$$\geq p(u_{fi} = 0 | \underline{u}^k_{ir}, \; u_{ki} = 1, \; y_i) \tag{3-5}$$

$i=1,\ldots,N$; $j=1,\ldots,N$; $i \neq j$; $k=1,\ldots,N$

where

$\underline{u}^k_{ir} \triangleq \underline{u}_{ir}$ without its kth element u_{ki}.

It is a reasonable assumption since in most practical cases, one would expect that the probability that DM_i will declare H_o is higher when DM_k sends the decision $u_{ki} = 0$ as opposed to $u_{ki} = 1$.

Under (A.1) through (A.5), the problem is to obtain the optimum strategies γ_{fi}, γ_{ji}, $i=1,\ldots,N, j=1,\ldots,N, i \neq j$, so as to

$$
(P1) \begin{cases}
\text{Minimize } E\{J(\gamma_{fi}(y_i,\gamma_{ji},\ j=1,\ldots,N,\ j \neq i),\ H)\} \\
\gamma_i \ \varepsilon\ \Gamma \qquad\qquad\qquad\qquad \text{for } i=1,\ldots,N \\
\text{subject to (A.1) - (A.5)}, \\
\text{where} \\
\gamma_i \triangleq (\gamma_{fi},\ \gamma_{ji},\ j=1,2,\ldots,N,\ j \neq i) \\
\text{and} \\
\Gamma \triangleq \text{ Set of all possible decision rules}
\end{cases} \qquad (3\text{-}6)
$$

The solution to the problem (P1) is given by the following theorem.

Theorem 3.1

The optimal decision rules associated with DM_i, $i=1,\ldots,N$, are as follows.

At the detector i, D_i, the decision rules γ_{ij}, $j=1,\ldots, i-1, i+1,\ldots,N$ are given by the likelihood ratio tests

$$\Lambda_i(y_i) \underset{u_{ij}=0}{\overset{u_{ij}=1}{\gtrless}} t_{ij} \; , \tag{3-7}$$

where

$$\Lambda_i(y_i) = \frac{p(y_i|H_1)}{p(y_i|H_0)} \tag{3-8}$$

and

$$t_{ij} \triangleq \frac{\displaystyle\sum_{u_{fj},\underline{u}^i_{jr}} P_o\{p[u_{fj}|\underline{u}^i_{jr},u_{ij}=1,H_o]-p[u_{fi}|\underline{u}^i_{jr},u_{ij}=0,H_o]\} \cdot \prod_{\substack{k=1\\k\neq i,j}}^{N} p(u_{kj}|H_o)\; J(u_{fj},H_o)}{\displaystyle\sum_{u_{fj},\underline{u}^i_{jr}} P_1\{p[u_{fj}|\underline{u}^i_{jr},u_{ij}=0,H_1]-p[u_{fi}|\underline{u}^i_{jr},u_{ij}=1,H_1]\} \cdot \prod_{\substack{k=1\\k\neq i,j}}^{N} p(u_{kj}|H_1)\; J(u_{fj}|H_1)} \tag{3-9}$$

$i=1,\dots,N$; $j=1,\dots, i-1, i+1,\dots,N$.

At the data fusion center i, DF_i, the decision rule γ_{fi}, $i=1,\dots,N$, is given by the likelihood ratio test

$$\Lambda_i(y_i) \underset{u_{fi}=0}{\overset{u_{fi}=1}{\gtrless}} t_{fi}(\underline{u}_{ir}) \tag{3-10}$$

where

$$t_{fi}(\underline{u}_{ir}) \triangleq \frac{P_o \prod_{\substack{j=1 \\ j\neq i}}^{N} p(u_{ji}|H_o)[J(u_{fi}=1, H_o) - J(u_{fi}=0, H_o)]}{P_1 \prod_{\substack{j=1 \\ j\neq i}}^{N} p(u_{ji}|H_1)[J(u_{fi}=0, H_1) - J(u_{fi}=1, H_1)]} \quad (3\text{-}11)$$

Proof:

First, we obtain the decision strategy at the data fusion center of DM_i. The objective is to obtain the optimum decision rules γ_{fi} so as to minimize $E\{J(\gamma_{fi}, H)\}$ over the set of all the decisions rules. The average cost or risk function at the site i is given by

$$R_{fi} = E\{J(u_{fi}, H)\} \quad (3\text{-}12)$$

Expanding (3-12).

$$R_{fi} = \sum_{H,u_{fi},\, \underline{u}_{ir}} \int_{\underline{y}} p(H)\; p(u_{fi},\underline{u}_{ir},\underline{y}|H)\; J(u_{fi},H)\; d\underline{y} \quad (3\text{-}13)$$

where

$$\underline{y} \triangleq (y_1, y_2, \ldots, y_N)$$

Then, using (A.1) - (A.5) and noting that u_{fi} does not depend on y_j, $j=1,\ldots,N$, $j \neq i$, we have

$$R_{fi} = \sum_{H,u_{fi},\underline{u}_{ir}} \int_{\underline{y}} p(H)p(u_{fi}|\underline{u}_{ir}, y_i, H)$$

$$\cdot\ p(\underline{u}_{ir}, \underline{y}|H)\ \cdot\ J(u_{fi}, H)\ d\underline{y} \qquad (3\text{-}14)$$

Summing over all values of u_{fi}, and ignoring constant terms, we have

$$R_{fi} = \sum_{H,\underline{u}_{ir}} \int_{\underline{y}} p(H)\ p(u_{fi}=0|\underline{u}_{ir},y_i,H)\ p(\underline{u}_{ir},\underline{y}|H)\ J(u_{fi} = 0,H)\ d\underline{y}$$

$$+ \sum_{H,\underline{u}_{ir}} \int_{\underline{y}} p(H)\ p(u_{fi}=1|\underline{u}_{ir},y_i,H)\ p(\underline{u}_{ir},\underline{y}|H)\ J(u_{fi}=1,H)\ d\underline{y} \qquad (3\text{-}15)$$

Gathering appropriate terms in (3-15) and using (3-3), we have

$$R_{fi} = \sum_{\underline{u}_{ir}} \int_{y_i} p(u_{fi}=0|\underline{u}_{ir}, y_i)\ \cdot \sum_{H} \int_{\underline{y}^i} p(H)\ p(\underline{u}_{ir}, \underline{y}|H)$$

$$[J(u_{fi}=0,H)\ -\ J(u_{fi}=1,H)]\ d\underline{y} \qquad (3\text{-}16)$$

where

$$\underline{y}^i \triangleq (y_1,\ldots,y_{i-1},\ y_{i+1},\ldots,y_N) \qquad (3\text{-}17)$$

Expression (3-16) is minimized if

$$p(u_{fi}=0|\underline{u}_{ir}, y_i)$$

$$= \begin{cases} 0 & \text{if } \sum_{H} \int_{\underline{y}^i} p(H)\ p(\underline{u}_{ir}, \underline{y}|H)[J(u_{fi}=0, H) - J(u_{fi}=1, H)]\ d\underline{y}^i \geq 0 \\ 1 & \text{otherwise} \end{cases} \qquad (3\text{-}18)$$

Keeping in mind that u_{ki} depends only on y_k for k=1,2,...,N, $k \neq i$, i.e., $u_{ki} = \gamma_{ki}(y_k)$, and using (3-1), we have

$$p(\underline{u}_{ir}|H) = \sum_{\substack{j=1 \\ j\neq i}}^{N} p(u_{ji}|H) \quad . \qquad (3\text{-}19)$$

Using (3-19) and expanding the condition in (3-18) with respect to H, we have

$$p(H_o) \prod_{\substack{j=1 \\ j\neq i}}^{N} p(u_{ji}|H_o)\, p(y_i|H_o)\, [J(u_{fi}=0, H_o) - J(u_{fi}=1, H_o)]$$

$$+ p(H_1) \prod_{\substack{j=1 \\ j\neq i}}^{N} p(u_{ji}|H_1)\, p(y_i|H_1)\, [J(u_{fi}=0, H_1) - J(u_{fi}=1, H_1)] \underset{u_{fi}=0}{\overset{u_{fi}=1}{\gtrless}} 0 \qquad (3\text{-}20)$$

Rearranging equation (3-20), we obtain (3-10). Note that the decision rule of (3-10) is an LRT which is a consequence of the independence assumption. The threshold t_{fi} at DF_i is a function of the received decision vector $\underline{u}_{ir}$, which implies that at each DF_i we have 2^{N-1} thresholds. The value of each threshold is different for different values of the incoming decision vector $\underline{u}_{ir}$.

Now, we obtain the decision strategies, $\gamma_{ji}(\cdot)$ for the intermediate decisions to be transmitted. Expanding the risk function R_{fi} given in (3-14) with respect to u_{ji} and ignoring a constant term, we have

$$R_{fi} = \sum_{H,u_{fi},\underline{u}^j_{ir}} \int_{\underline{Y}} p(H)\,[p(u_{fi}|\underline{u}^j_{ir},\ u_{ji} = 0,\ y_i)$$

$$- p(u_{fi}|\underline{u}^j_{ir},\ u_{ji}=1,\ y_i)]$$

$$\cdot\ p(y_j|H)\ \cdot\ p(u_{ji}=0|y_j)\ \cdot\ p(\underline{u}^j_{ir},\ \underline{Y}^j|H)\ J(u_{fi},H)\ d\underline{Y} \qquad (3\text{-}21)$$

Rewriting (3-21) in a more appropriate form, we have

$$R_{fi} = \int_{y_j} p(u_{ji}=0|y_j) \int_{\underline{Y}^j} \sum_{H,u_{fi},\underline{u}^j_{ir}} p(H)\ p(y_j|H)\ p(\underline{u}^j_{ir},\underline{Y}^j|H)$$

$$\cdot\ J(u_{fi},H)\ \cdot\ [p(u_{fi}|\underline{u}^j_{ir},u_{ji}=0,y_i) - p(u_{fi}|\underline{u}^j_{ir},u_{ji}=1,y_i)]\ d\underline{Y} \qquad (3\text{-}22)$$

This expression is minimized if

$$p(u_{ji} = 0|y_j) =$$

$$\begin{cases} 0 & \text{if } \displaystyle\int_{\underline{Y}^j} \sum_{u_{fi},H,\underline{u}^j_{ir}} p(H)p(y_j|H)\ p(\underline{u}^j_{ir},\underline{Y}^j|H)\ J(u_{fi},H)\ \cdot \\ & \quad \cdot\ [p(u_{fi}|\underline{u}^j_{ir},u_{ji}=0,y_i) - p(u_{fi}|\underline{u}^j_{ir},u_{ji}=1,y_i)] \\ & \quad d\underline{Y}^j \geq 0 \\ 1 & \text{elsewhere} \end{cases} \qquad (3\text{-}23)$$

Expanding (3-23) over H, we get equation (3-9), and which completes the proof of Theorem 3-1. Q.E.D.

Note that the decision rule $\gamma_{ji}(y_j)$ is obtained by minimizing the risk function associated with DM_i. Intuitively, it makes sense, since the decision u_{ji} is an intermediate decision that helps DM_i make a final decision u_{fi}. None of the thresholds depend upon the observation vector $\underline{y}$, so that an off-line calculation of the thresholds is possible. The set of thresholds associated with DM_i, $\{t_{fi}, t_{ji}, j = 1, 2, \ldots, N, j \neq i\}$ can be determined through the simultaneous solution of $2^{N-1} + N-1$ nonlinear coupled equations in $2^{N-1} + N-1$ unknowns. It is important to note that these equations are necessary conditions (but not sufficient) which must be satisfied by the thresholds. They define locally optimal solutions, each must be checked to assure that a global minimum is found. In the next section, an example is presented for illustration.

3.3. Example

We consider a binary hypothesis testing problem with two decision makers DM_1 and DM_2 as shown in Figure 3.4. We assume that there are two hypotheses H_1 and H_0, where $P(H_1) = p$, and $P(H_0) = 1-p$. We choose the minimum probability of error cost function, i.e.,

$$J(u_{fi} = 1, H_1) = J(u_{fi} = 0, H_o) = 0 \quad \text{and}$$

$$J(u_{fi} = 0, H_1) = J(u_{fi} = 1, H_o) = 1 \,. \qquad (3\text{-}24)$$

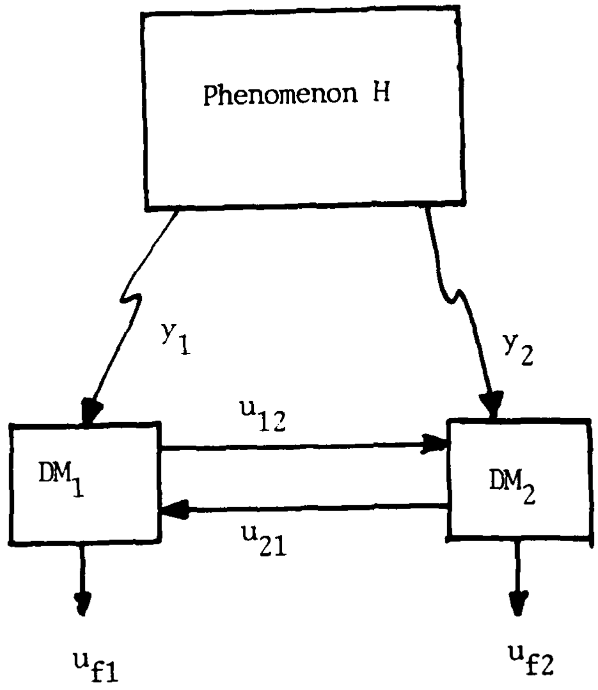


Figure 3.4 Distributed Detection System with Distributed Data Fusion for Two Sensors.

Before we proceed further, we state the results of Section 3.2 for the case of N=2. They can be summarized in the following:

$$\Lambda_i(y_i) \underset{u_{fi}=0}{\overset{u_{fi}=1}{\gtrless}} t_{fi}(u_{ji}) \tag{3-25}$$

for $i=1,2$, $j=1,2$, $j \neq i$,

where

$$t_{fi}(u_{ji}) \triangleq \frac{p(H_o)\; p(u_{ji}|H_o)\; [J(u_{fi}=1,\; H_o) - J(u_{fi}=0,\; H_o)]}{p(H_1)\; p(u_{ji}|H_1)\; [J(u_{fi}=0,\; H_1) - J(u_{fi}=1,\; H_1)]} \tag{3-26}$$

The decision rule $\gamma_{ji}(\cdot)$ is given by

$$\Lambda_j(y_j) \underset{u_{ji}=0}{\overset{u_{ji}=1}{\gtrless}} t_{ji} \qquad \text{for } j=1,2\,, \tag{3-27}$$

where

$$t_{ji} = \frac{\sum_{u_{fi}} p(H_o)[p(u_{fi}|u_{ji}=1,\; H_o) - p(u_{fi}|u_{ji}=0,\; H_o)]J(u_{fi},H_o)}{\sum_{u_{fi}} p(H_1)[p(u_{fi}|u_{ji}=0,\; H_1) - p(u_{fi}|u_{ji}=1,\; H_1)]J(u_{fi},H_1)}$$

$$\text{for} \quad i=1,2, \quad j=1,2, \quad i \neq j\,. \tag{3-28}$$

The optimal decision rules associated with DM_i, i=1,2,are described by three thresholds at the site i. The thresholds are determined by solving a set of three coupled nonlinear equations in three unknowns. We further assume that the conditional density $p(y_i|H_j)$ are exponential, i.e.,

$$p(y_i|H_o) = \exp(-y_i)$$

$$p(y_i|H_1) = \frac{1}{\lambda_i} \exp(-\frac{1}{\lambda_i} y_i) \; ; \; \lambda_i > 0, \quad i=1,2$$

for $y_i \geq 0$ and

$$p(y_i|H_j) = 0 \; ; \quad i=1,2 \text{ elsewhere} \tag{3-29}$$

The optimum decision rules at the site 1 are obtained from equations (3-25)-(3-28). We have the following equations.

$$\frac{p(y_i|H_1)}{p(y_i|H_o)} = \frac{1}{\lambda_i} \exp\,[y_i(1 - \frac{1}{\lambda_i})] \qquad i=1,2. \tag{3-30}$$

The thresholds $t_{f1}(u_{21})$ and t_{21} are obtained using expressions (3-26) and (3-28). We have

$$t_{f1}(u_{21}) = \frac{1-p}{p} \cdot \frac{p(u_{21}|H_o)}{p(u_{21}|H_1)} \tag{3-31}$$

$$t_{21} = \frac{[p(u_{f1}=1|u_{21}=1, H_o) - p(u_{f1}=1|u_{21}=0, H_o)](1-p)}{[p(u_{f1}=0|u_{21}=0, H_1) - p(u_{f1}=0|u_{21}=1, H_1)]p} \; . \tag{3-32}$$

We can evaluate the thresholds using the probability density functions and the decision rules. Note that expressions (3-31) and (3-32) are coupled. One needs to solve them simultaneously in order to obtain a solution. Next, we obtain the error probabilities to be used in order to compute the thresholds.

$$p(u_{21}=1|H_o) = \int_{y_i:\,\Lambda_2(y_2)\geq t_{21}} p_{y_2|H_o}(y_2|H_o)\, dy_2 \qquad (3\text{-}33)$$

Let $\lambda_i > 1$, then

$$p(u_{21}=1|H_o) = (\lambda_2\ t_{21})^{\frac{\lambda_2}{1-\lambda_2}} \qquad (3\text{-}34)$$

and

$$p(u_{21}=1|H_1) = (\lambda_2\ t_{21})^{\frac{1}{1-\lambda_2}} \qquad (3\text{-}35)$$

Furthermore, one needs to get expressions for $p(u_{f1}|\ u_{21}, H)$, we have

$$p(u_{f1}=1|u_{21},\ H_o) = [\lambda_1\ t_{f1}(u_{21})]^{\frac{\lambda_1}{1-\lambda_1}} \qquad (3\text{-}36)$$

and

$$p(u_{f1}=1|u_{21},\ H_1) = [\lambda_1\ t_{f1}(u_{21})]^{\frac{1}{1-\lambda_1}} \qquad (3\text{-}37)$$

Using expressions (3-31) and (3-32), we obtain the following coupled equations

$$t_{21} = \frac{1-p}{p} \frac{[\lambda_1 t_{f1}(u_{21}=0)]^{\frac{\lambda_1}{1-\lambda_1}} - [\lambda_1 \; t_{f1}(u_{21}=1)]^{\frac{\lambda_1}{1-\lambda_1}}}{[\lambda_1 t_{f1}(u_{21}=0)]^{\frac{1}{1-\lambda_1}} - [\lambda_1 \; t_{f1}(u_{21}=1)]^{\frac{1}{1-\lambda_1}}} ,$$

$$t_{f1}(u_{21}=1) = \frac{1-p}{p} \frac{1}{\lambda_2 t_2} , \qquad (3\text{-}38)$$

$$t_{f1}(u_{21}=0) = \frac{1-p}{p} \frac{1-(\lambda_2 t_{21})^{\frac{\lambda_2}{1-\lambda_2}}}{1-(\lambda_2 t_{21})^{\frac{1}{1-\lambda_2}}} .$$

In Figure 3-5, we present the ROC curve for this example.

3.4 Discussion

In this chapter, we have considered the problem of distributed Bayesian hypothesis testing with distributed data fusion. We have derived the optimal decision rules for the system. Due to the independence assumption, the decision rules are likelihood ratio tests (LRT). At each site, we obtain $2^{N-1}+N-1$ coupled nonlinear equations. The simultaneous solution of these equtions yield the $2^{N-1}+N-1$ thresholds. These are necessary equations which must be satisfied by all of the thresholds. There may be several

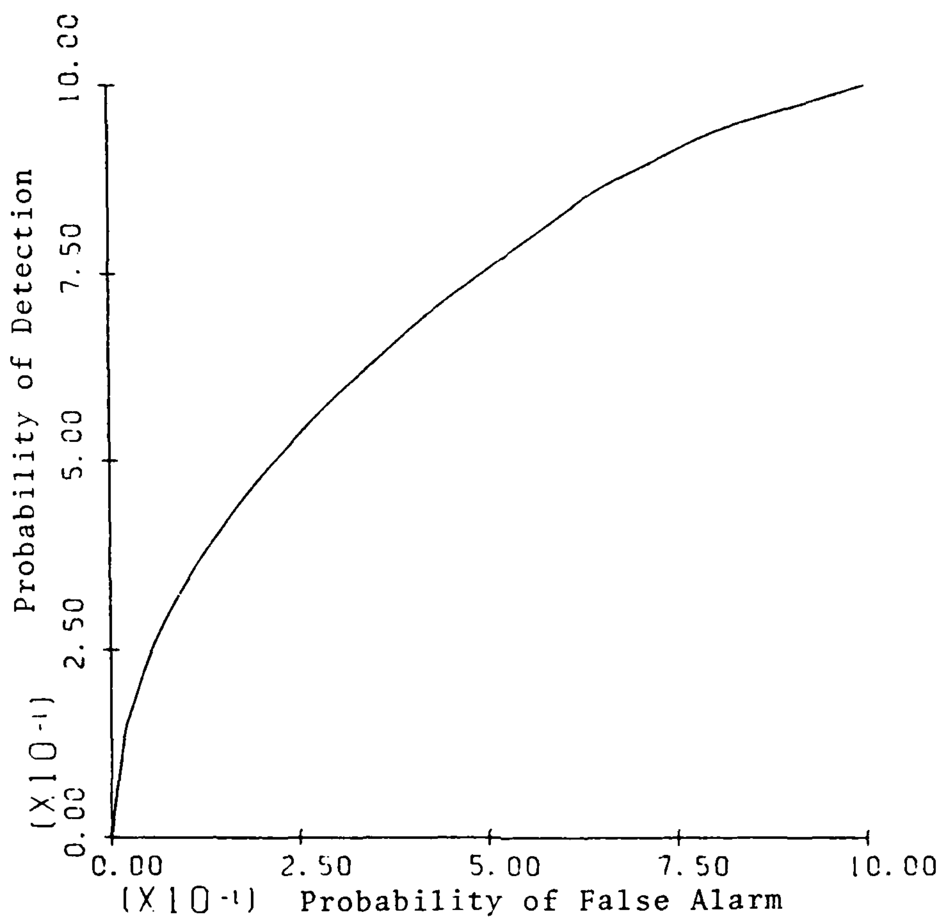


Figure 3.5. Receiver Operating Characteristic for $\lambda_1=2$, $\lambda_2=1.5$.

local minima; each must be checked to assure that a global minimum is found. In the case of N=2, we obtain the equations derived by Ekchian and Tenney [5] at each site. The results of this chapter can be generalized to the case of M hypotheses in a straightforward manner. In the next chapter, we will study the problem of distributed Neyman-Pearson hypothesis testing with distributed data fusion. Also, we formulate and solve the problem of Neyman-Pearson distributed detection for a tandem topology network.

CHAPTER FOUR

DISTRIBUTED NEYMAN-PEARSON HYPOTHESIS TESTING WITH DATA FUSION

4.1. Introduction

In Bayesian hypothesis testing, knowledge of the a priori probabilities and a cost assignment is required. In many practical situations, it is difficult to assign realistic costs or a priori probabilities may not be known. An approach to bypass this difficulty is to use the Neyman-Pearson (N-P) test. In this approach, the knowledge of a priori probabilities and a cost assignment is not required and, therefore, it has many practical applications. In the N-P detection problem, we maximize the probability of detection under a constraint on the probability of false-alarm.

In this chapter, we solve the problem of distributed hypothesis testing with data fusion using the Neyman-Pearson approach for various distributed sensor network (DSN) topologies. The problem of distributed hypothesis testing with a centralized data fusion using the Neyman-Pearson approach has been solved by Hoballah [15]. In Section 4.2, we solve the Neyman-Pearson hypothesis testing problem with data fusion for a specific two-sensor system. In Section 4.3, these results are extended to solve the problem of distributed Neyman-Pearson hypothesis testing with distributed data fusion. In Section 4.4, the results of Section 4.2 are further extended to solve the distributed Neyman-Pearson hypothesis testing problem for N detectors connected in tandem. In Section 4.5, we present an example for illustration.

4.2. A Distributed Neyman-Pearson Hypothesis Testing Problem with Two Decision Makers

In this section, we consider the system configuration as shown in Figure 4.1 . The two possible hypotheses, are denoted by H_o or H_1. Each decision maker i, DM_i, takes an observation y_i, i=1,2. Based on y_1, DM_1 makes a decision $u_1 \in \{0,1\}$ which corresponds to the hypotheses H_o or H_1. The decision u_1 is transmitted to DM_2. Then, based on u_1 and y_2, DM_2 makes a final decision u_{f2}. The joint probability density function $p(y_1,y_2|H)$ is assumed to be known a priori. The aim is to find the optimal decision rule at DM_1 and the fusion rule at DM_2 using the Neyman-Pearson approach, i.e., we maximize the overall probability of detection, $Pr[u_{f2} = 1|H_1]$, given a constraint on the overall probability of false alarm, $Pr[u_{f2} = 1|H_o]$.

The problem is formulated as follows:

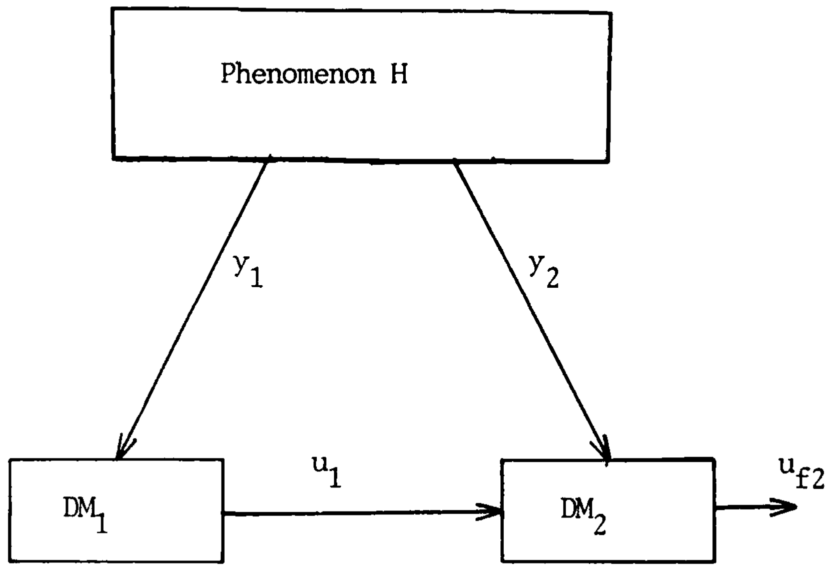


Figure 4.1 A Two Decision Maker Distributed Detection System with Data Fusion.

(A.1) Let $\{y_i, i=1,2\}$ be a set of conditionally independent random variables given the hypothesis H.

(A.2) The decision maker 2 makes a decision u_{f2} where

$$u_{f2} = \gamma_{f2}(y_2, u_1) \quad . \tag{4-1}$$

(A.3) The decision maker 1 makes a decision u_1 where

$$u_1 = \gamma_1(y_1) \tag{4-2}$$

Under (A.1)-(A.3), the distributed binary N-P hypothesis testing problem can be stated as follows:

(P2)
Find the decision rules $\gamma_{f2}(\cdot)$ and $\gamma_1(\cdot)$ so as to minimize

$$F = Pr[u_{f2}=0|H_1] + \lambda(Pr[u_{f2}=1|H_o] - \alpha) \tag{4-3}$$

under the constraint that

$$Pr[u_{f2}=1|H_o] = \alpha' \leq \alpha \tag{4-4}$$

where

$\lambda \triangleq$ the Lagrange multiplier .

The solution of problem (P2) is given by the following theorem:

Theorem 4.1

For the two-decision maker system shown in Fig. 4.1, the optimal decision rules using the Neyman-Pearson approach are given as follows:

(a) The optimum decision rule at DM_2 is described by the LRT

$$\Lambda_2(y_2) \triangleq \frac{p(y_2|H_1)}{p(y_2|H_o)} \underset{u_{f2}=0}{\overset{u_{f2}=1}{\gtrless}} t_{f2}\ (u_1) \tag{4-5}$$

where

$$t_{f2}(u_1) \triangleq \lambda \frac{p(u_1|H_o)}{p(u_1|H_1)} , \tag{4-6}$$

and

$$\alpha' = \sum_{u_1} \int_{y_2:\Lambda_2(\cdot) \geq t_{f2}(\cdot)} p(y_2|H_o)\ p(u_1|H_o)\ dy_2 . \tag{4-7}$$

(b) The optimum decision rule at DM_1 is described by the LRT

$$\Lambda_1(y_1) \triangleq \frac{p(y_1|H_1)}{p(y_1|H_o)} \underset{u_1=0}{\overset{u_1=1}{\gtrless}} t_1 , \tag{4-8}$$

where

$$t_1 \triangleq \frac{\lambda\,[p(u_{f2}=1|u_1=1,H_o) - p(u_{f2}=1|u_1=0,\ H_o)]}{[p(u_{f2}=0|u_1=0,H_1) - p(u_{f2}=0|u_1=1,\ H_1)]} \tag{4-9}$$

and

$$\frac{\alpha - p(u_{f2}=1 | u_1=0,\ H_o)}{p(u_{f2}=1 | u_1=1,\ H_o) - p(u_{f2}=1 | u_1=0,\ H_o)} = \int_{y_1 : \Lambda(y_1) \geq t_1} p(y_1 | H_o)\, dy_1 \qquad (4\text{-}10)$$

Proof:

The decision rule at DM_2 is obtained as follows. The objective is to minimize F where

$$F = Pr[u_{f2}=0 | H_1] + \lambda(Pr[u_{f2}=1 | H_o] - \alpha) \qquad (4\text{-}11)$$

Equation (4-11) leads to the classical Neyman-Pearson test where the observations are $\{y_2,\ u_1\}$. In other words, we obtain the following test [1],

$$\frac{p(u_1, y_2 | H_1)}{p(u_1, y_2 | H_o)} \underset{u_{f2}=0}{\overset{u_{f2}=1}{\gtrless}} \lambda \ . \qquad (4\text{-}12)$$

Using the independence assumption (A.1), we have

$$\frac{p(y_2 | H_1)}{p(y_2 | H_o)} \underset{u_{f2}=0}{\overset{u_{f2}=1}{\gtrless}} \lambda \frac{p(u_1 | H_o)}{p(u_1 | H_1)} \ . \qquad (4\text{-}13)$$

One needs to find an expression for the Lagrange multiplier λ to completely define the decision rule $\gamma_{f2}(\cdot)$. Since the probability of false alarm $Pr[u_{f2}=1 | H_o] \leq \alpha$, let us choose

$$Pr[u_{f2}=1 | H_o] = \alpha' \leq \alpha \ . \qquad (4\text{-}14)$$

We then have

$$\Pr[u_{f2}=1|H_o] = \alpha' = \sum_{u_1} p(u_{f2}=1|u_1, H_o)\, p(u_1|H_o) \quad , \tag{4-15}$$

where

$$p(u_{f2}=1|u_1, H_o) = \int_{y_2} p(u_{f2}=1, y_2|u_1, H_o)\, dy_2 \quad . \tag{4-16}$$

Using Bayes rule, we have

$$p(u_{f2}=1|u_1, H_o) = \int_{y_2} p(u_{f2}=1|u_1, y_2, H_o)\, p(y_2|u_1, H_o)\, dy_2 . \tag{4-17}$$

Keeping in mind that u_{f2} depends only on $\{u_1,y_2\}$, we have

$$p(u_{f2}=1|u_1, H_o) = \int_{y_2} p(u_{f2}=1|u_1,y_2)\, p(y_2|H_o)\, dy_2 \tag{4-18}$$

Expression (4-18) can be simplified as follows:

$$p(u_{f2}=1|u_1, H_o) =$$

$$\int_{y_2:\Lambda_2(\cdot)\geq t_{f2}(u_1)} p(u_{f2}=1|u_1,y_2)\, p(y_2|H_o)\, dy_2$$

$$+ \int_{y_2:\Lambda_2(\cdot)<t_{f2}(u_1)} p(u_{f2}=1|u_1,y_2)\, p(y_2|H_o)\, dy_2 \tag{4-19}$$

Since

$$p(u_{f2}=1|u_1,y_2) = \begin{cases} 1 & \text{if } \Lambda_2(y_2) \geq t_{f2}(u_1) \\ 0 & \text{otherwise} \end{cases} , \quad (4\text{-}20)$$

equation (4-15) becomes

$$p(u_{f2}=1|H_o) = \alpha' = \sum_{u_1} \int_{y_2:\Lambda_2(\cdot)\geq t_{f2}(u_1)} p(y_2|H_o)\, p(u_1|H_o)\, dy_2 \quad (4\text{-}21)$$

which completes the proof of Theorem 4-1(a). Q.E.D.

The decision rule at the decision maker 1 is obtained as follows. Starting with expression (4-11), we have

$$F = Pr[u_{f2}=0|H_1] + \lambda(Pr[u_{f2}=1|H_o] - \alpha) \quad (4\text{-}22)$$

Using (4-15) in (4-22), we have

$$F = \sum_{u_1} p(u_{f2}=0|u_1, H_1)\, p(u_1|H_1)$$

$$+ \lambda\, (\sum_{u_1} p(u_{f2}=1|u_1, H_o)\, p(u_1|H_o) - \alpha) . \quad (4\text{-}23)$$

Expanding the above expression with respect to u_1, we have

$$F = p(u_1=0|H_1)\,(D_{o|o,1} - D_{o|1,1}) + D_{o|1,1}$$

$$+ \lambda[p(u_1=1|H_o)(D_{1|1,o} - D_{1|o,o}) + D_{1|o,o} - \alpha] \quad (4\text{-}24)$$

where

$$D_{i|jk} \triangleq p(u_{f2}=i|u_1=j,\ H = H_k), \quad i,j,k \in \{0,1\} \quad (4\text{-}25)$$

Writing expression (4-24) in the sam form as (4-11), we have

$$F_{u1} = p(u_1=0|H_1) + \lambda_{u1}[p(u_1=1|H_o) - \alpha_{u1}] + C_{u1} \quad (4\text{-}26)$$

where

$$F_{u1} \triangleq \frac{F}{D_{o|o,1} - D_{o|1,1}} ,$$

$$\lambda_{u1} \triangleq \frac{\lambda\,(D_{1|1,o} - D_{1|o,o})}{D_{o|o,1} - D_{o|1,1}} ,$$

$$\alpha_{u1} \triangleq \frac{\alpha - D_{1|o,o}}{D_{1|1,o} - D_{1|o,o}} ,$$

and

$$C_{u1} \triangleq \frac{D_{o|1,1}}{D_{o|o,1} - D_{o|1,1}} .$$

The minimization of (4-26) yields the following LRT:

$$\Lambda_1(y_1) \triangleq \frac{p(y_1|H_1)}{p(y_1|H_o)} \underset{u_1=0}{\overset{u_1=1}{\gtrless}} \lambda \frac{D_{1|1,o} - D_{1|o,o}}{D_{o|o,1} - D_{o|1,1}} \triangleq t_1 \quad (4\text{-}27)$$

and

$$Pr[u_1=1|H_o] = \frac{\alpha - D_{1|o,o}}{D_{1|1,o} - D_{1|o,o}} \quad (4\text{-}28)$$

where

$$\Pr[u_1=1|H_o] = \int_{y_1:\Lambda_1(y_1)\geq t_1} p(y_1|H_o)\, dy_1 \qquad (4\text{-}29)$$

which completes the proof of Theorem 4-1(b). Q.E.D.

In the next two sections, we generalize the results of this section in two different directions. First, we formulate and solve the distributed detection problem with distributed data fusion centers. Then we solve the problem of tandem topology distributed network for N decision makers.

4.3. Distributed Neyman-Pearson Hypothesis Testing with Distributed Data Fusion

In this section, we consider the system configuration as shown in Figure 4.2. The signalling scheme for distributed data fusion has already been described in Chapter 3. In that chapter, we solved the distributed Bayesian hypothesis testing problem with distributed data fusion. In this section, we solve the distributed Neyman-Pearson hypothesis testing problem with distributed data fusion.

The aim here is to find all of the optimal decision rules at each site $i, i=1,\ldots,N$. At the site i; the objective is to maximize the probability of detection $\Pr[u_{fi}=1|H_1]$, given a constraint on the probability of false alarm $\Pr[u_{fi}=1|H_o]$. We formulate the problem as follows:

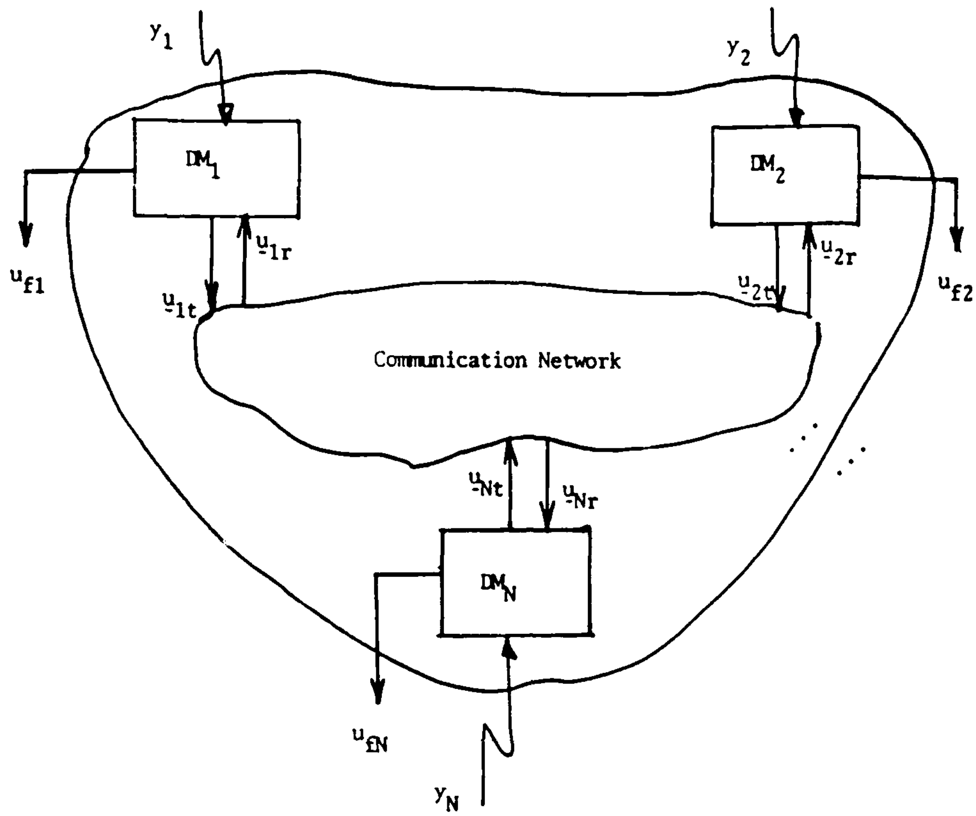


Figure 4.2 Distributed N-Sensor Detection System with Distributed Data Fusion.

(A.4) Let $\{y_i, i=1,\ldots,N\}$ be a set of independent random variable given the hypothesis H.

(A.5) At each site i, the decision maker i makes a set of intermediate decisions $\underline{u}_{it}$ where

$$\underline{u}_{it} \triangleq (u_{i1},\ldots u_{ii-1}, u_{ii+1},\ldots,u_{iN}) \tag{4-30}$$

and

$$u_{ij} = \gamma_{ij}(y_i) \quad . \tag{4-31}$$

(A.6) At each site i, the decision maker i makes a final decision u_{fi}

$$u_{fi} = \gamma_{fi}(y_i, \underline{u}_{ir}) \ , \tag{4-32}$$

where

$\gamma_{fi}(\cdot)$ is the final decision rule at the site i, i=1,...,N,

and

$\underline{u}_{ir}$ is the received decision vector from the other decision makers,

$$\underline{u}_{ir} \triangleq (u_{1i},\ldots, u_{i-1\ i}, u_{i+1\ i},\ldots, u_{Ni}) \ . \tag{4-33}$$

Under (A.4) - (A.6), the distributed binary N-P hypothesis testing problem can be stated as follows:

$$
\text{(P3)}
\begin{cases}
\text{At each } DM_i\text{, find the decision rules } \gamma_{ji}(\cdot), \\
\text{and } \gamma_{fi}(\cdot),\ i,j=1,\ldots,N,\ i \neq j \text{ so as to minimize} \\
F_i \triangleq Pr[u_{fi}=0|H_1] + \lambda_i(Pr[u_{fi}=1|H_o] - \alpha_i) \quad (4\text{-}34) \\
\text{under the constraint that} \\
Pr\ [u_{fi}=1|H_o] \leq \alpha_i\ , \quad (4\text{-}35) \\
\lambda_i \triangleq \text{the Lagrange multiplier at the site i.}
\end{cases}
$$

The solution of problem (P3) is given by the following theorem:

Theorem 4.2

For the distributed detection system shown in Fig. 4.2, the optimal decision rules using the Neyman-Pearson approach are given as follows:

(a) The optimum decision rule $\gamma_{fi}(\cdot)$ at the site i, i=1,...,N, is described by the LRT.

$$\Lambda_i(y_i)= \frac{p(y_i|H_1)}{p(y_i|H_o)} \underset{u_{fi}=0}{\overset{u_{fi}=1}{\gtrless}} t_{fi}(\underline{u}_{ir}) \ , \tag{4-36}$$

where

$$t_{fi}(\underline{u}_{ir}) \triangleq \lambda_i \prod_{\substack{j=1 \\ j\neq i}}^{N} \frac{p(u_{ji}|H_o)}{p(u_{ji}|H_1)} \tag{4-37}$$

and

$$p(u_{fi}=1|H_o) = \alpha'$$

$$= \sum_{\underline{u}_{ir}} \int_{y_i:\Lambda_{fi}(\cdot)\geq t_{fi}} p(y_i|H_o)\cdot \prod_{\substack{j=1\\ j\neq i}}^{N} p(u_{ji}|H_o)\, dy_i \tag{4-38}$$

(b) The optimum decision rule $\gamma_{ki}(\cdot)$, $k=1,2,\ldots,i-1,$ $i+1,\ldots,N$; at the detector k is given by the LRT

$$\Lambda_k(y_k) = \frac{p(y_k|H_1)}{p(y_k|H_o)} \underset{u_{ki}=0}{\overset{u_{ki}=1}{\gtrless}} t_{ki}\ , \tag{4-39}$$

where

$$t_{ki} \triangleq$$

$$\frac{\lambda_i \sum_{\underline{u}_{ji}^{k}} [p(u_{fi}=1|H_o,\underline{u}_{ir}^{k},u_{ki}=1)-p(u_{fi}=1|H_o,\underline{u}_{ir}^{k},u_{ki}=0)] \prod_{\substack{j=1\\ j\neq i,k}}^{N} p(u_{ji}|H_o)}{\sum_{\underline{u}_{ji}^{k}} [p(u_{fi}=0|H_1,\underline{u}_{ir}^{k},u_{ki}=0)-p(u_{fi}=0|H_1,\underline{u}_{ir}^{k},u_{ki}=1)] \prod_{\substack{j=1\\ j\neq i,k}}^{N} p(u_{ji}|H_1)} \tag{4-40}$$

and

$$p(u_{ki}=1|H_o) = \frac{\alpha - \sum_{\underline{u}_{ji}^{k}} p(u_{fi}=1|H_o,\underline{u}_{ir}^{k},u_{ki}=0) \prod_{\substack{j=1\\ j\neq i,k}}^{N} p(u_{ji}|H_o)}{\sum_{\underline{u}_{ji}^{k}} [p(u_{fi}=1|H_o,\underline{u}_{ir}^{k},u_{ki}=1)-p(u_{fi}=1|H_o,\underline{u}_{ir}^{k},u_{ki}=0)] \prod_{\substack{j=1\\ j\neq i,k}}^{N} p(u_{ji}|H_o)} \tag{4-41}$$

Proof:

The strategy at the data fusion center i of DM_i is obtained first. The objective is to minimize $Pr[u_{fi} = 0|H_1]$ under a constraint on the probability of false alarm, i.e., $Pr[u_{fi} = 1|H_o] = \alpha_i' < \alpha_i$. Let us define F_i as

$$F_i = Pr[u_{fi}=0|H_1] + \lambda_i[Pr[u_{fi}=1|H_o] - \alpha_i] \quad . \tag{4-42}$$

The minimization of F_i leads to the classical Neyman-Pearson test where the observations are $\{y_i, u_{ji}, j=1,2,\ldots,N, j \neq i\}$. In other words, we will obtain the following test [1]

$$\Lambda_i(y_i) \triangleq \frac{p(y_i, \underline{u}_{ir}|H_1)}{p(y_i, \underline{u}_{ir}|H_o)} \underset{u_{fi}=0}{\overset{u_{fi}=1}{\gtrless}} \lambda_i \; . \tag{4-43}$$

Using the independence assumption (A.4), we have

$$\frac{p(y_i|H_1)\; p(\underline{u}_{ir}|H_1)}{p(y_i|H_o)\; p(\underline{u}_{ir}|H_o)} \underset{u_{fi}=0}{\overset{u_{fi}=1}{\gtrless}} \lambda_i \; , \tag{4-44}$$

or

$$\frac{p(y_i|H_1)}{p(y_i|H_o)} \underset{u_{fi}=0}{\overset{u_{fi}=1}{\gtrless}} \lambda_i \prod_{\substack{j=1 \\ j\neq i}}^{N} \frac{p(u_{ji}|H_o)}{p(u_{ji}|H_1)} \; . \tag{4-45}$$

One needs to find an expression for the Lagrange multiplier λ_i in order to completely define the decision rule $\gamma_{fi}(\cdot)$. Since the probability of false alarm $Pr[u_{fi}=1|H_o] \leq \alpha_i$, we choose $Pr[u_{fi}=1|H_o] = \alpha_i' \leq \alpha_i$.

Using the theorem of total probability, we have

$$p(u_{fi}=1|H_o) = \sum_{\underline{u}_{ir}} p(u_{fi}=1|\underline{u}_{ir}, H_o)\, p(\underline{u}_{ir}|H_o) \cdot \quad (4\text{-}46)$$

Using the assumption that u_{ji} depends only on y_j, (4-46) reduces to

$$p(u_{fi}=1|H_o) = \sum_{\underline{u}_{ir}} p(u_{fi}=1|\underline{u}_{ir}, H_o) \prod_{\substack{j=1\\ j\neq i}}^{N} p(u_{ji}|H_o)\ , \quad (4\text{-}47)$$

where

$$p(u_{fi}=1|\underline{u}_{ir}, H_o) = \int_{y_i} p(u_{fi}=1|\underline{u}_{ir},y_i)\, p(y_i|H_o)dy_i\ . \quad (4\text{-}48)$$

The probability $p(u_{fi}=1|\underline{u}_{ir}, y_i)\ \varepsilon\ \{0,1\}$ depending on the values of $\{\underline{u}_{ir}, y_i\}$ which is the set of observations that the data fusion center i receives. Following the steps in equation (4-19) and (4-20), equation (4-48) can be written as

$$p(u_{fi}=1|\underline{u}_{ir}, H_o) = \int_{y_i:\Lambda_{fi}(\cdot)\geq t_{fi}(\cdot)} p(y_i|H_o)\ dy_i\ . \quad (4\text{-}49)$$

Then substituting (4-49) in (4-47), we get (4.38), and this completes the proof of Theorem 4.2(a).

The decision rule $\gamma_{ji}(\cdot)$ for the decision maker j for $j=1,\ldots,i-1,\ i+1,\ldots,N$, associated with the function F_i is derived next. Substituting (4-46) into (4-42), we obtain the expression for F_i to be

$$F_i = \sum_{\underline{u}_{ir}} p(u_{fi}=0|H_1, \underline{u}_{ir})\, p(u_{ir}|H_1)$$

$$+ \lambda_i [\sum_{\underline{u}_{ir}} p(u_{fi}=1|H_o,\underline{u}_{ir})\, p(\underline{u}_{ir}|H_o) - \alpha_i] \qquad (4\text{-}50)$$

Expanding (4-50) in terms of u_{ki}, we have

$$F_i = \sum_{\underline{u}_{ir}^k} p(u_{fi}=0|H_1 , \underline{u}_{ir}^k, u_{ki}=0) \prod_{\substack{j=1\\ j\neq i,k}}^{N} p(u_{ji}|H_1)\, p(u_{ki}=0|H_1)$$

$$+ \sum_{\underline{u}_{ir}^k} p(u_{fi}=0|H_1,\underline{u}_{ir}^k,u_{ki}=1) \prod_{\substack{j=1\\ j\neq i,k}}^{N} p(u_{ji}|H_1)\, p(u_{ki}=1|H_1)$$

$$+ \lambda_i [\sum_{\underline{u}_{ir}^k} p(u_{fi}=1|H_o,\underline{u}_{ir}^k,u_{ki}=0) \prod_{\substack{j=1\\ j\neq i,k}}^{N} p(u_{ji}|H_o)\, p(u_{ki}=0|H_o)$$

$$+ \sum_{\underline{u}_{ir}^k} p(u_{fi}=1|H_o,\underline{u}_{ir}^k, u_{ki}=1) \prod_{\substack{j=1\\ j=i,k}}^{N} p(u_{ji}|H_o)\, p(u_{ki}=1|H_o)$$

$$- \alpha_i] \qquad (4\text{-}51)$$

We can write (4-51) as

$$F_i = p(u_{ki}=0|H_1) \; D_{o1} + C_{o1}$$

$$+ \lambda_i [p(u_{ki}=1|H_o)\; D_{1o} + C_{1o} - \alpha_i] \qquad (4\text{-}52)$$

where

$$D_{o1} = \sum_{\underline{u}_{ir}^k} [p(u_{fi}=0|H_1,\underline{u}_{ir}^k,u_{ki}=0) - p(u_{fi}=0|H_1,\underline{u}_{ir}^k,u_{ki}=1)] \prod_{\substack{j=1 \\ j\neq i,k}}^{N} p(u_{ji}|H_1), \quad (4\text{-}53)$$

$$C_{o1} = \sum_{\underline{u}_{ir}^k} p(u_{fi}=0|H_1, \underline{u}_{ir}^k, u_{ki}=1) \prod_{\substack{j=1 \\ j\neq i,k}}^{N} p(u_{ji}|H_1), \quad (4\text{-}54)$$

$$D_{1o} = \sum_{\underline{u}_{ir}^k} [p(u_{fi}=1|H_o,\underline{u}_{ir}^k,u_{ki}=1) - p(u_{fi}=1|H_o,\underline{u}_{ir}^k,u_{ki}=0)] \prod_{\substack{j=1 \\ j\neq i,k}}^{N} p(u_{ji}|H_o) \quad (4\text{-}55)$$

and

$$C_{10} = \sum_{\underline{u}_{ir}^k} [p(u_{fi}=1|H_o, \underline{u}_{ir}^k, u_{ki}=0)] \prod_{\substack{j=1 \\ j\neq i,k}}^{N} p(u_{ji}|H_o) \quad (4\text{-}56)$$

Writing (4-52) in the same form as (4-11), we have

$$F_{ki} = p(u_{ki} = 0|H_1) + \lambda_{ki}[p(u_{ki} = 1|H_o) - \alpha_{ki}] + C_{ki}, \quad (4\text{-}57)$$

where

$$F_{ki} \triangleq \frac{F_i}{D_{o1}},$$

$$\lambda_{ki} \triangleq \frac{\lambda_i \; D_{1o}}{D_{o1}},$$

$$\alpha_{ki} \triangleq \frac{\alpha_i - C_{1o}}{D_{10}}$$

and

$$C_{ki} \triangleq \frac{C_{o1}}{D_{o1}}$$

The minimization of (4-57) yields the following result

$$\Lambda_k(y_k) \triangleq \frac{p(y_k|H_1)}{p(y_k|H_o)} \underset{u_{ki}=0}{\overset{u_{ki}=1}{\gtrless}} t_{ki}, \tag{4-58}$$

where t_{ki} is given by equation (4-40) and $p(u_{ki}=1|H_o)$ is given by equation (4-41). This completes the proof of Theorem 4-2. Q.E.D.

In the next section, we formulate and solve the problem of distributed Neyman-Pearson hypothesis testing in a tandem topology network consisting of N decision makers.

4.4. Distributed Neyman-Pearson Hypothesis Testing in a Tandem Topology Network

In this section, we consider the system configuration as shown in Figure 4.3. We solve the distributed Neyman-Pearson detection problem for an N-decision maker tandem topology. The distributed detection network can be viewed as a team of decision makers, where each decision maker receives a conditionally independent observation given the hypothesis H. The team of DM's desire to maximize the probability of detection $Pr[u_N=1|H_1]$ of the last decision maker, DM_N, under a constraint on the probability of false alarm $Pr[u_N=1|H_o]$ of DM_N.

We formulate the problem as follows:

(A.7) Let $\{y_i, i=1,...,N\}$ be a set of conditionally

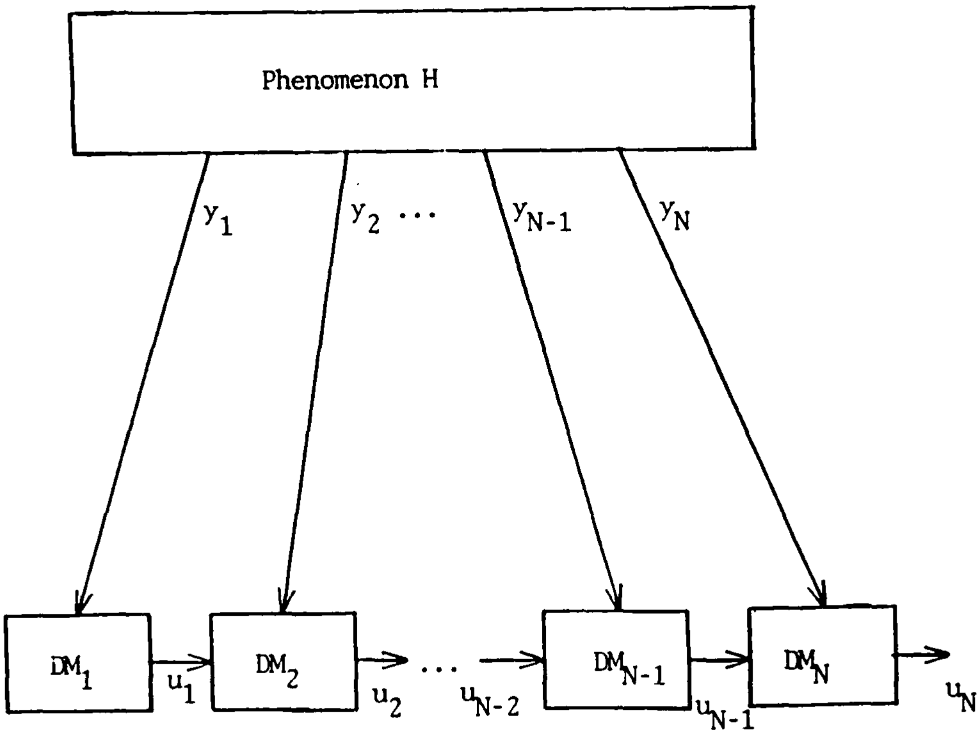


Figure 4.3 N-decision Maker Tandem Topology Detection Network.

independent random variables given the hypothesis H.

(A.8) At the first site, the decision maker 1 makes a decision u_1,

$$u_1 = \gamma_1(y_1) \quad . \tag{4-59}$$

(A.9) At the site i, i=2,...,N, the decision maker i makes a decision u_i based on the incoming decision and its own observation, i.e.,

$$u_i = \gamma_i(y_i, u_{i-1}) \quad . \tag{4-60}$$

Under (A.7)-(A.9) this problem can be formulated as follows:

(P4)
At each DM, find the decision rule $\gamma_i(\cdot)$, i=1,...,N, so as to minimize

$$F_N \triangleq Pr[u_N=0|H_1] + \lambda_N[Pr[u_N=1|H_o] - \alpha_N]$$

under the constraint that

$$Pr[u_N=1|H_o] = \alpha_N' \leq \alpha_N \ , \tag{4-61}$$

where

λ_N is the Lagrange multiplier.

The solution of problem (P4) is given by the following theorem:

Theorem 4.3

For the distributed detection system shown in

Fig. 4.3, the optimal decision rules using the Neyman-Pearson approach are given as follows:

(a) The optimum decision rule $\lambda_N(\cdot)$ at the site N, is described by the LRT

$$\Lambda_N(y_N) \triangleq \frac{p(y_N|H_1)}{p(y_N|H_o)} \underset{u_N=0}{\overset{u_N=1}{\gtrless}} t_N(u_{N-1}) , \tag{4-62}$$

where

$$t_N(u_{N-1}) \triangleq \lambda_N \frac{p(u_{N-1}|H_o)}{p(u_{N-1}|H_1)} \tag{4-63}$$

and

$$p(u_N=1|H_o) = \alpha_N'$$

$$= \sum_{u_{N-1}} \int_{y_N:\Lambda_N(\cdot)\geq t_N(\cdot)} p(y_N|H_o)\, dy_N\, p(u_{N-1}|H_o) \tag{4-64}$$

(b) The optimum decision rule $\gamma_i(\cdot)$ at the site i, i=2,3,...,N-1, is described by the LRT

$$\Lambda_i(y_i) \triangleq \frac{p(y_i|H_1)}{p(y_i|H_o)} \underset{u_i=0}{\overset{u_i=1}{\gtrless}} t_i(u_{i-1}) , \tag{4-65}$$

where

$$t_i(u_{i-1}) = \lambda_N \frac{p(u_{i-1}|H_o)[p(u_N=1|u_i=1,H_o) - p(u_N=1|u_i=0,H_o)]}{p(u_{i-1}|H_1)[p(u_N=0|u_i=0,H_1) - p(u_N=0|u_i=1,H_1)]} \tag{4-66}$$

and

$$\frac{\alpha_N - p(u_N=1|u_i=0,H_o)}{p(u_N=1|u_i=1,H_o)-p(u_N=1|u_i=0,H_o)}$$

$$= \sum_{u_{i-1}} \int_{y_i:\Lambda_i(\cdot)\geq t_i(\cdot)} p(y_i|H_o)\, dy_i\; p(u_{i-1}|H_o)\,. \tag{4-67}$$

(c) The optimum decision rule $\gamma_1(\cdot)$ at the site 1 is also described by an LRT:

$$\Lambda_1(y_1) = \frac{p(y_1|H_1)}{p(y_1|H_o)} \underset{u_1=0}{\overset{u_1=1}{\gtrless}} t_1 \;, \tag{4-68}$$

where

$$t_1 = \lambda_N \frac{p(u_N=1|u_1=1,\; H_o) - p(u_N=1|u_1=0,H_o)}{p(u_N=0|u_1=0,\; H_1) - p(u_N=0|u_1=1,H_1)} \tag{4-69}$$

and

$$\frac{\alpha_N - p(u_N=1|u_1=0,H_o)}{p(u_N=1|u_1=1,H_o) - p(u_N=1|u_1=0,H_o)} = \int_{y_1:\Lambda_1(\cdot)>t_1(\cdot)} p(y_1|H_o)\, dy_1\,. \tag{4-70}$$

<u>Proof</u>:

The decision rule at DM_N is obtained as follows. The objective of the distributed detection network is to minimize the overall probability of miss $Pr[u_N=0|H_1]$ under a constraint on the probability of false alarm $Pr[u_N=1|H_o] = \alpha'_N \leq \alpha_N$. We construct the function F_N.

$$F_N = Pr[u_N=0|H_1] + \lambda_N[Pr[u_N=1|H_o] - \alpha_N] \quad . \tag{4-71}$$

The minimization of (4-71) leads to the classical Neyman-Pearson test where the observation set for DM_N is $\{y_N, u_{N-1}\}$. In other words, we will obtain the following test [1].

$$\frac{p(y_N, u_{N-1}|H_1)}{p(y_N, u_{N-1}|H_o)} \underset{u_N=0}{\overset{u_N=1}{\gtrless}} \lambda_N \, . \tag{4-72}$$

Using the independence assumption, we have

$$\frac{p(y_N|H_1)}{p(y_N|H_o)} \underset{u_N=0}{\overset{u_N=1}{\gtrless}} \lambda_N \frac{p(u_{N-1}|H_o)}{p(u_{N-1}|H_1)} \quad . \tag{4-73}$$

One needs to find an expression for the Lagrange multipler λ_N in order to completely define the decision rule $\gamma_N(\cdot)$. We choose $Pr[u_N=1|H_o] = \alpha_N' \leq \alpha_N$ and we have

$$p[u_N=1|H_o] = \alpha_N' = \sum_{u_{N-1}} p(u_N=1, u_{N-1}|H_o) \tag{4-74}$$

Using the theorem of total probability, we have

$$Pr[u_N=1|H_o] = \sum_{u_{N-1}} p(u_N=1|u_{N-1},H_o) \, p(u_{N-1}|H_o) \tag{4-75}$$

and, as before we obtain

$$Pr[u_N=1|H_o] = \sum_{u_{N-1}} \int_{y_N:\Lambda_N(\cdot)\geq t_N(\cdot)} p(y_N|H_o) \, dy_N \, p(u_{N-1}|H_o) \tag{4-76}$$

This completes the proof of Theorem 4.3(a).

The decision rule $\gamma_i(\cdot)$, $i=2,3,\ldots,N-1$, is obtained next. We start with the function F_N,

$$F_N = \Pr[u_N=0|H_1] + \lambda_N \, [\Pr[u_N=1|H_o] - \alpha_N] \; . \qquad (4\text{-}77)$$

Expanding the probability of miss and false alarm with respect to u_i, we have

$$F_N = \sum_{u_i} p(u_N=0,u_i|H_1) + \lambda_N \, [\sum_{u_i} p(u_N=1,u_i|H_o) - \alpha_N] \, . \qquad (4\text{-}78)$$

Using Bayes rule, we obtain

$$F_N = \sum_{u_i} p(u_N=0|u_i,H_1) \; p(u_i|H_1)$$

$$+ \lambda_N[\sum_{u_i} p(u_N=1|u_i,H_o) \; p(u_i|H_o) - \alpha_N] \; . \qquad (4\text{-}79)$$

Gathering similar terms and ignoring constant terms, we have

$$F_N = p(u_i=0|H_1) \; D_{o1}$$

$$+ \lambda_N[p(u_i=1|H_o) \; D_{1o} + C_{1o} - \alpha_N] \; , \qquad (4\text{-}80)$$

where

$$D_{o1} \triangleq p(u_N=0|u_i=0,H_1) - p(u_N=0|u_i=1,H_1) \; , \qquad (4\text{-}81)$$

$$D_{1o} \triangleq p(u_N=1|u_i=1,H_o) - p(u_N=1|u_i=0,H_o) \qquad (4\text{-}82)$$

and

$$C_{1o} \triangleq p(u_N=1|u_i=0,H_o) \quad . \qquad (4\text{-}83)$$

Writing (4-80) in the same form as (4-11), we have

$$F_N = p(u_i=0|H_1) + \frac{\lambda_N D_{1o}}{D_{o1}} [p(u_i=1|H_o) - \frac{\alpha_N - C_{1o}}{D_{1o}}] \ . \qquad (4\text{-}84)$$

The minimization of expression (4-84) yields the following LRT,

$$\frac{p(y_i,\ u_{i-1}|H_1)}{p(y_i,\ u_{i-1}|H_o)} \underset{u_i=0}{\overset{u_i=1}{\gtrless}} \lambda_N \frac{D_{1o}}{D_{o1}} \ . \qquad (4\text{-}85)$$

Using the fact that y_i does not depend on u_{i-1}, we have

$$\frac{p(y_i|H_1)}{p(y_i|H_o)} \underset{u_i=0}{\overset{u_i=1}{\gtrless}} \lambda_N \frac{D_{1o}}{D_{o1}} \cdot \frac{p(u_{i-1}|H_o)}{p(u_{i-1}|H_1)} \qquad (4\text{-}86)$$

and

$$p(u_i=1|H_o) = \frac{\alpha_N - C_{1o}}{D_{1o}}$$

$$= \sum_{u_{i-1}} \int_{y_i:\Lambda_i(\cdot)\geq t_i(\cdot)} p(y_i|H_o)\ p(u_{i-1}|H_o)\ dy_i \qquad (4\text{-}87)$$

which completes the proof of Theorem 4-3(b).

The decision rule $\gamma_1(\cdot)$ is obtained as follows. Using the function F_N, we have

$$F_N = Pr[u_N=0|H_1] + \lambda_N[Pr[u_N=1|H_o] - \alpha_N] \quad . \tag{4-88}$$

Expanding F_N with respect to u_1, we have

$$F_N = \sum_{u_1} p(u_N=0,u_1|H_1) + \lambda_N[\sum_{u_1} p(u_N=1,u_1|H_o) - \alpha_N] \quad . \tag{4-89}$$

Using the total probability theorem, we have

$$F_N = \sum_{u_1} p(u_N=0|u_1,H_1)\, p(u_1|H_1) + \lambda_N[\sum_{u_1} p(u_N=1|u_1,H_o)\, p(u_1|H_o) - \alpha_N] \tag{4-90}$$

Gathering similar terms and ignoring constant terms, we have

$$F_N = p(u_1=0|H_1) + \lambda_N \frac{A_{1o}}{A_{o1}} [p(u_1=1|H_o) - \frac{\alpha_N - B_{1o}}{A_{1o}}] \tag{4-91}$$

where

$$\lambda'_{u_1} \triangleq \lambda_N \frac{A_{1o}}{A_{o1}} \quad , \tag{4-92}$$

$$\alpha'_{u_1} \triangleq \frac{\alpha_N - B_{1o}}{A_{1o}} \quad , \tag{4-93}$$

$$A_{1o} \triangleq p(u_N=1|u_1=1,H_o) - p(u_N=1|u_1=0,H_o) \quad , \tag{4-94}$$

$$A_{o1} \triangleq p(u_N=0|u_1=0,H_1) - p(u_N=0|u_1=1,H_1) \tag{4-95}$$

and

$$B_{1o} = p(u_N=1|u_1=0,H_o) \quad . \tag{4-96}$$

The minimization of expression (4-91) yields the following LRT,

$$\frac{p(y_1|H_1)}{p(y_1|H_o)} \underset{u_1=0}{\overset{u_1=1}{\gtrless}} \lambda'_{u_1} \tag{4-97}$$

and

$$p(u_1=1|H_o) = \alpha'_{u_1} \tag{4-98}$$

This completes the proof of Theorem 4-3. Q.E.D.

4.5. Example

We consider a simple binary hypothesis problem with two decision makers DM_1 and DM_2 as shown in Figure 4-1. We assume that there are two hypotheses H_1 and H_o. The observations at both decision makers are exponentially distributed, i.e.,

$$p(y_i|H_o) = \exp(-y_i) \quad , \tag{4-99}$$

$$p(y_i|H_1) = \frac{1}{\theta_i} \exp(- \frac{1}{\theta_i} y_i) \; ;$$

$$\theta_i > 0 \; , \; i = 1,2 \; ; \; y_i \geq 0 \tag{4-100}$$

and

$$p(y_i|H_j) = 0 \qquad i=1,2 \quad \text{otherwise} \; . \tag{4-101}$$

The optimum decision rules $\gamma_{f2}(\cdot)$ and $\gamma_1(\cdot)$ are obtained using Theorem 4.1, we have

$$\Lambda_i(y_i) \triangleq \frac{p(y_i|H_1)}{p(y_i|H_o)} = \frac{1}{\theta_i} \exp[y_i(1 - \frac{1}{\theta_i})] \tag{4-102}$$

$$i = 1,2 \; .$$

The decision rule is given by

$$\Lambda_i(y_i) \underset{u_i=0}{\overset{u_i=1}{\gtrless}} t_i \ , \qquad (4\text{-}103)$$

which is equivalent to

$$y_i \underset{u_i=0}{\overset{u_i=1}{\gtrless}} \frac{\theta_i}{\theta_{i-1}} \ln(\theta_i t_i) \triangleq t_i' \qquad (4\text{-}104)$$

The probability of false alarm and the probability of miss for the two detectors are given by

$$P_{Fi} = (\theta_i t_i)^{\frac{\theta_i}{1-\theta_i}} = \exp(-\ t_i') \qquad (4\text{-}105)$$

and,

$$P_{Mi} = 1 - (\theta_i t_i)^{\frac{1}{1-\theta_i}} = 1 - \exp(-t_i'/\theta_i) \ . \qquad (4\text{-}106)$$

From (4-6) and (4-9), the two thresholds satisfy the following set of equations

$$t_2(u_1=1) = \lambda \frac{p(u_1=1|H_o)}{p(u_1=1|H_1)} = \lambda \frac{P_{F1}}{1-P_{M1}}$$

$$= \lambda \frac{\exp(-\ t_1')}{\exp(-\ t_1'/\theta_1)}$$

$$= \exp[-\ t_1' \ (\frac{\theta_1 - 1}{\theta_1})] \qquad (4\text{-}107)$$

and

$$t_2(u_1=0) = \lambda \frac{p(u_1=0|H_o)}{p(u_1=0|H_1)}$$

$$= \lambda \frac{1 - \exp(-\ t_1')}{1 - \exp(-\ t_1'/\theta_1)} \quad , \qquad (4\text{-}108)$$

and

$$t_1 = \lambda \frac{p(u_{f2}=1|u_1=1,H_o) - p(u_{f2}=1|u_1=0,H_o)}{p(u_{f2}=0|u_1=0,H_1) - p(u_{f2}=0|u_1=1,H_1)}$$

$$= \lambda \frac{\exp[-\ t_2'(u_1=1)] - \exp[-\ t_2'(u_1=0)]}{\exp[-\ \frac{t_2'(u_1=1)}{\theta_2}] - \exp[-\ \frac{t_2'(u_1=0)}{\theta_2}]} \qquad (4\text{-}109)$$

where

$$t_i = \frac{1}{\theta_i} \exp(\frac{\theta_i - 1}{\theta_i}\ t_i') \quad . \qquad (4\text{-}110)$$

Finally, using (4.7) we obtain

$$\alpha' = \sum_{u_1} p(u_1|H_o) \int_{y_2:\Lambda_2(\cdot)\geq t_2(u_1)} p(y_2|H_o)\ dy_2 \qquad (4\text{-}111)$$

$$= p(u_1=0|H_o)\ p(u_2=1|u_1=0,H_o) + p(u_1=1|H_o)\ p(u_2=1|u_1=1,H_o)$$

(4-112)

then

$$\alpha = [1 - \exp(-\ t_1')]\ \exp(-\ t_2'(u_1=0))$$

$$+ \exp(-\ t_1')\ \exp(-\ t_2'(u_1=1)) \qquad (4\text{-}113)$$

Solving (4-107), (4-108), (4-109) and (4-113) simultaneously, we obtain the desired thresholds. We show the ROC for the case when $\theta_1=2$ and $\theta_2=1.5$ in Figure 4.4.

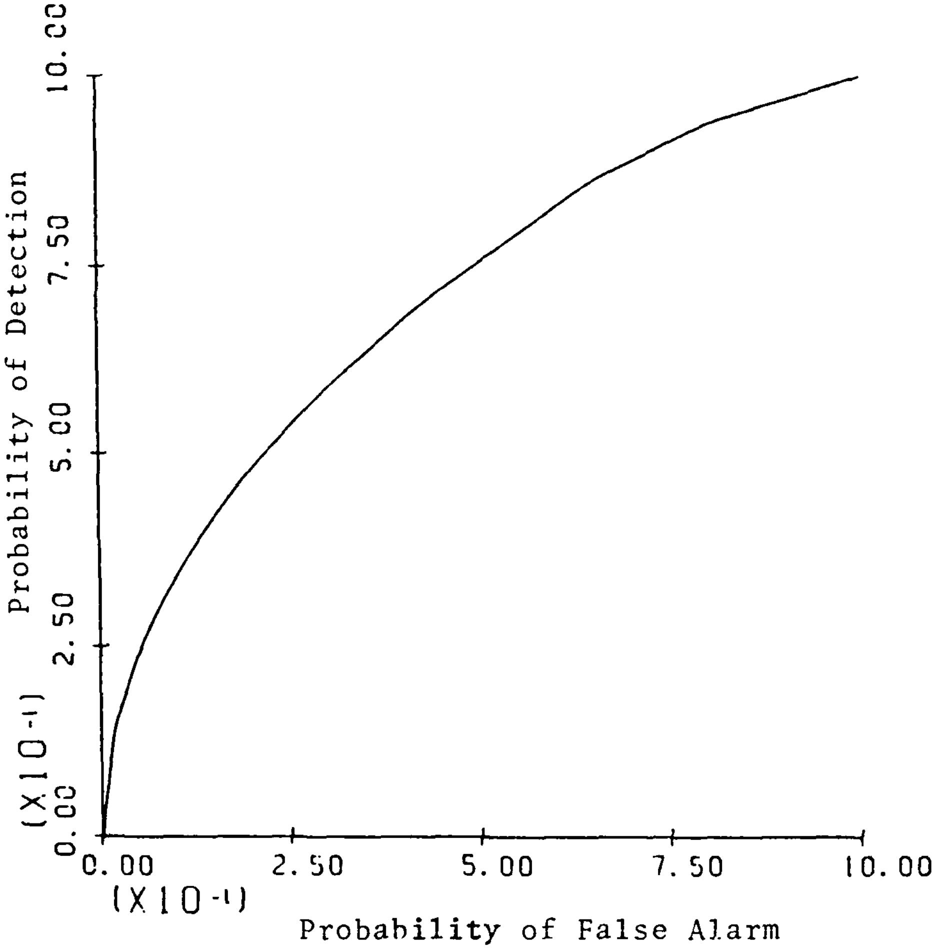


Figure 4.4. Receiver Operating Characteristic for $\theta_1=2$ and $\theta_2=1.5$.

4.6. Discussion

In this chapter, we have considered the problem of distributed Neyman-Pearson hypothesis testing for two different topologies. First, we solved the problem of distributed hypothesis testing with distributed data fusion. Then, we solved the problem of distributed detection for an N-decision maker tandem topology. In both cases, we derived the optimum decision rules for each decision maker. We have shown that the threshold equations are coupled. One needs to solve these equations simultaneously in order to obtain a solution for the thresholds. There may be several solutions for these coupled nonlinear equations. Each solution must be checked to assure that the global solution is found. We have also shown that the decision rules are likelihood ratio tests, this is due to the independence assumption. The present analysis can be extended in many directions. For instance, the case of M-ary hypothesis testing with N decision makers can be considered. In the next chapter, we consider the problem of distributed sequential detection.

CHAPTER FIVE

DECENTRALIZED SEQUENTIAL HYPOTHESIS TESTING

5.1 Introduction

The decentralized sequential detection problem has received some recent attention [10,11]. In [10], Teneketzis considered a decentralized version of the Wald problem with two decision makers. In his model, each detector was given the flexibility of either stopping and making a decision or continuing to the next time stage. The coupling between the two detectors was introduced through a common cost function. His results show that coupling causes considerable complexity in the computation of the optimal stopping rules. In [11], Hashemipour and Rhodes examined a two-step, two-detector, hypothesis-testing problem with a data fusion center. They also discussed its straightforward extension to a decentralized multi-stage sequential detection problem. Their model is different from the one examined in [10]. In [11], the sequential test is performed at the data fusion center and local detectors have no control over the termination of the test.

Here, we make some further contributions in the area of distributed sequential hypothesis testing. First, we solve the distributed sequential detection problem for a tandem topology. This is a multi-stage extension of the work of Ekchian and Tenney [5]. Then, we solve the sequential probability ratio test problem for the same model as in [11].Here, however, we use the Neyman-Pearson approach for the solution of the problem.

In Section 5.2, we briefly discuss the centralized Wald problem. The results are used in Section 5.3. In Section 5.3, we formulate and solve the decentralized Wald problem for a tandem topology network. In Section 5.4, we formulate and derive the decision rules for the decentralized sequential probability ratio test. Finally, in Section 5.5 we discuss the results of this chapter.

5.2. The Centralized Wald Problem

In this section, we present briefly the centralized Wald Problem and its solution. The Wald Problem is a well known problem in statistical sequential analysis [13,18]. It can be formulated as follows:

(A.1) Consider two hypotheses $\{H_0, H_1\}$ where the a priori probability of H_0 is given by

$$\mathrm{Prob}[H_0] = P_0 \ . \tag{5-1}$$

(A.2) Consider a detector whose observation at time t is denoted by y_t. Let $\{y_t;\ t=1,2,\ldots\}$ be a set of independent random variables. Based on $y^t \triangleq (y_1,\ldots,y_t)$, the detector makes a decision u given by

$$u = \gamma(y^t) \ , \quad u \in \{0,1\} \tag{5-2}$$

where $\gamma(\cdot)$ is the decision rule.

(A.3) The cost of making a decision u at the detector is $J(u,H)$, where H is the true hypothesis.

(A.4) The detector must make a decision u no later than t=N. Then the Centralized Wald Problem is to

$$\text{(P5)} \left\{ \begin{array}{l} \text{Minimize } E[J(\gamma(y^t),\ H)] \\ \gamma \in \Gamma \\ \text{Subject to (A.1) - (A.4)} \\ \text{where} \\ \Gamma \triangleq \text{Set of all stopping rules} \end{array} \right. \tag{5-3}$$

The solution to problem (P5) is given in [13]. It can be summarized in the following theorem:

Theorem 5.1

The optimal decision rule at the detector is described by (2N-1) thresholds ℓ_N, α_1, $\beta_1,\ldots,\alpha_{N-1}$, β_{N-1}. The thresholds can be obtained from the following equations.

At t = k (k=1,2,...,N-1)

$$\alpha_k J(0,1) + (1 - \alpha_k) J(1,1)$$

$$= E_{y_{k+1}} [\tilde{J}_{k+1} (\frac{\alpha_k p(y_{k+1}|H_o)}{\alpha_k p(y_{k+1}|H_o) + (1-\alpha_k) p(y_{k+1}|H_1)})] \tag{5-4}$$

and

$$\beta_k J(0,0) + (1 - \beta_k) J(1,0)$$

$$= E_{y_{k+1}} [\tilde{J}_{k+1} (\frac{\beta_k p(y_{k+1}|H_o)}{\beta_k p(y_{k+1}|H_o) + (1-\beta_k) p(y_{k+1}|H_1)})] \tag{5-5}$$

where

$$\tilde{J}_k(p(H_o|y^k)) = \min\{\min_{u\in[0,1]} E_{y_k}\{ J(u,H)|y^k\},$$

$$E_{y_{k+1}} \{\tilde{J}_{k+1}(p(H_o|y^{k+1}))\}\} \tag{5-6}$$

and

$E_{y_{k+1}}(\cdot)$ denotes the expectation with respect to

$p(y_{k+1}|y^k)$.

At t = N

$$\ell_N = \frac{J(0,1) - J(1,1)}{J(1,0) + J(0,1) - J(0,0) - J(1,1)} \tag{5-7}$$

At each time t, t=1,...,N, the detector evaluates $p(H_o|y^k)$ using its observation and compares $p(H_o|y^k)$ with α_t and β_t (or ℓ_N, if t=N) and decides to stop or continue according to the following rule:

At t = k (k=1,...,N-1)

$$\text{if } p[H_o|y^k] \leq \alpha_k \text{ , decide } H_1,$$

$$\text{if } p[H_o|y^k] \geq \beta_k \text{ , decide } H_o,$$

$$\text{if } \alpha_k < p[H_o|y^k] < \beta_k \text{ , continue to take observations.} \qquad (5\text{-}8)$$

At t = N

$$\text{if } p(H_o|y^N) \geq \ell_N \text{ , decide } H_o \text{ ,}$$

$$\text{if } p(H_o|y^N) < \ell_N \text{ , decide } H_1 \text{ ,} \qquad (5\text{-}9)$$

In this section,we have presented the centralized version of the Wald Problem - its formulation and its solution. In the next section, we formulate and solve the decentralized Wald Problem for a tandem topology network.

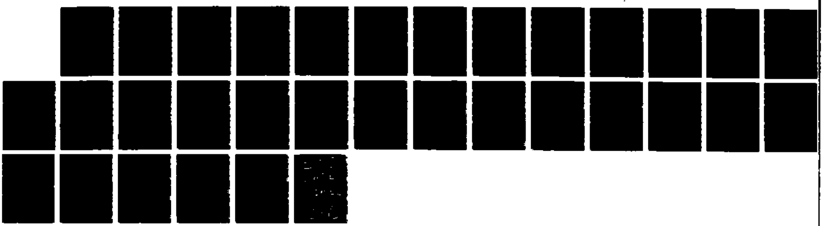
5.3. The Decentralized Wald Problem for a Tandem Topology Network

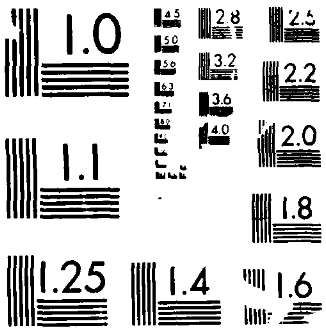
Consider the tandem topology network as shown in Figure 5.1. It consists of two decision makers DM_1 and DM_2 connected in series. In this section, we formulate and solve the decentralized Wald problem for the system shown in Figure 5.1.

The problem is formulated as follows:

AD-A195 910 ON HYPOTHESIS TESTING IN DISTRIBUTED SENSOR NETWORKS (U) SYRACUSE UNIV NY DEPT OF ELECTRICAL AND COMPUTER ENGINEERING Z CHAIR ET AL. NOV 87 RADC-TR-87-180 2/2

UNCLASSIFIED F30602-81-C-0169 F/G 17/8 NL





MICROCOPY RESOLUTION TEST CHART

DS-1963-A

(A.5) Consider two hypotheses $\{H_o, H_1\}$ where the a priori probability of H_o is given by

$$\text{Prob }[H_o] = P_o \quad . \tag{5-10}$$

(A.6) The decision maker 1 makes a decision u_t at time t given by

$$u_t = \gamma_u(x^t), \qquad u_t \in \{0, 1\} , \tag{5-11}$$

where $\gamma_u(\cdot)$ is the decision rule at DM_1, and $x^t \triangleq (x_1,\ldots,x_t)$ represents the set of observations at DM_1 up to time t.

(A.7) Define $u^t \triangleq (u_1,\ldots,u_t)$ where u_t is the decision of decision maker 1 at time t.

(A.8) The decision maker 2 makes a decision v_t at time t given by

$$v_t = \gamma_v(y_u^t) \ , \qquad v_t \in \{0, 1\} \tag{5-12}$$

where $\gamma_v(\cdot)$ is the decision rule at DM_2, and $y_u^t \triangleq (y_1,\ldots,y_t;\ u^t)$ represents the set of observations at DM_2 up to time t. Note that y_u^t also contains u^t, the set of incoming decisions from DM_1.

(A.9) Define $v^t \triangleq (v_1,\ldots,v_t)$ where v_t is the decision of DM_2 at time t.

(A.10) The observations $\{x_1,\ldots,x_t\}$ and $\{y_1,\ldots,y_t\}$ are assumed to be mutually independent given the hypothesis H, i.e.,

$$p(x_1,\ldots,x_t,\ y_1,\ldots,y_t|H) = \prod_{i=1}^{t} p(x_i|H) \prod_{j=1}^{t} p(y_j|H). \tag{5-13}$$

(A.11) The cost of making the final decision v_t at DM_2 is $J(v_t,H)$ where H is the true hypothesis. Furthermore, we assume that

$$J(0,\ H_1) \geq J(1,\ H_1)\ , \tag{5-14}$$

$$J(1,\ H_o) \geq J(0,\ H_o)\ . \tag{5-15}$$

(A.12) DM_1 performs a nonsequential test and yields a decision $u_t \in \{0,1\}$. DM_2 performs the centralized Wald test and yields a decision $v_t \in \{0,1,\text{ continue}\}$. DM_2 must make a final decision $v_t \in \{0,1\}$ no later than $t = N$.

Under (A.5) - (A.12), the decentralized sequential detection problem is to obtain all the decision rules so as to

(P6) $\left\{ \begin{array}{l} \text{Minimize } E\{J(\gamma_v(y_u^t),\ H)\} \\ \gamma \triangleq \{\gamma_v(\cdot)\ ,\ \gamma_u(\cdot)\} \in \Gamma_t \\ \text{subject to (A.5) - (A.12)} \\ \text{where } t \text{ is the stopping time of } DM_2 \text{ and } \Gamma_t \\ \text{is the set of all the stopping rules} \end{array} \right.$ (5-16)

The solution of the problem (P6) is given by the following theorem:

Theorem 5.2

For the system shown in Figure 5.1, the optimal decision rules for the two-stage decentralized sequential detection problem (N=2) are as follows:

(a) The optimum local decision rule at DM_1 at time 1 is described by the LRT,

$$\Lambda_{u_1}(x^1) \triangleq \frac{p(H_1|x^1)}{p(H_o|x^1)} \underset{u_1=0}{\overset{u_1=1}{\gtrless}} t_{u_1} , \tag{5-17}$$

where

$$t_{u_1} \triangleq \frac{\Sigma_1(1, H_o) - \Sigma_1(0, H_o)}{\Sigma_1(0, H_1) - \Sigma_1(1, H_1)} , \tag{5-18}$$

$$\Sigma_1(u_1, H) \triangleq E_{u_2|u_1,H}\Sigma_2(u^2, H)$$

and

$$\Sigma_2(u^2, H) \triangleq \sum_{v^2} p(v^2|u^2, H) J(v, H) . \tag{5-19}$$

(b) The optimum local decision rule at DM_1 at time 2 is described by the LRT,

$$\Lambda_{u_2}(x^2) = \frac{p(H_1|x^2)}{p(H_o|x^2)} \underset{u_2=0}{\overset{u_2=1}{\gtrless}} t_{u_2}(u_1) , \tag{5-20}$$

where

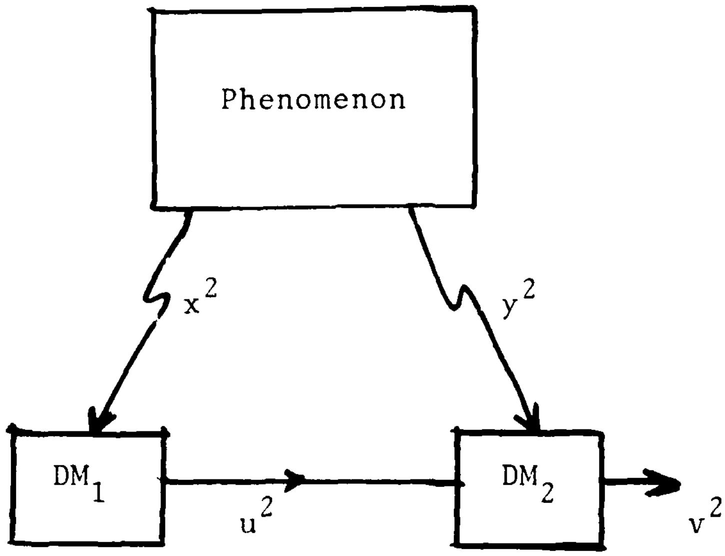


Figure 5.1   A Two-Detector Tandem topology Configuration.

$$t_{u_2}(u_1) \triangleq \frac{\Sigma_2(u_1,u_2=1,\ H_o) - \Sigma_2(u_1,u_2=0,\ H_o)}{\Sigma_2(u_1,u_2=0,\ H_1) - \Sigma_2(u_1,u_2=1,\ H_1)} \,. \quad (5\text{-}21)$$

(c) The optimum decision rules at DM_2 are described by three thresholds ℓ_2, α_1, β_1. The thresholds satisfy the following equations:

At t=1

$$\alpha_1\ J(0,1) + (1-\alpha_1)\ J(1,1)$$

$$= E_{y_2,u_2}\ [\tilde{J}_2\ [\frac{\alpha_1\ p(u_2,y_2|H_o)}{\alpha_1 p(u_2,y_2|H_o) + (1-\alpha_1)\ p(u_2,y_2|H_1)}]] \quad (5\text{-}22)$$

and

$$\beta_1\ J(0,0) + (1-\beta_1)\ J(1,0)$$

$$= E_{y_2,u_2}\ [\tilde{J}_2\ [\frac{\beta_1\ p(u_2,y_2|H_o)}{\beta_1 p(u_2,y_2|H_o) + (1-\beta_1)\ p(u_2,y_2|H_1)}]] \quad (5\text{-}23)$$

where

$$\tilde{J}_1\ [p(H_o|y_u^1)] =$$

$$\min\{\min_{v_t\in[0,1]} E_{y_1,u_1}\ \{J(v_t,H)\,|y_u^1\}\ , \quad (5\text{-}24)$$

$$E_{y_2,u_2}\ \{\tilde{J}_2(p(H_o|y_u^2))\}\}$$

At t=2

$$\ell_2 = \frac{J(0,1) - J(1,1)}{J(1,0) + J(0,1) - J(0,0) - J(1,1)} \quad (5\text{-}25)$$

At each time t, DM_2 evaluates $p(H_o|y_u^t)$ using its observation and compares $p(H_o|y_u^t)$ with α_1 and β_1 (or ℓ_2 if t=2). Decision maker 2 decides to stop or continue according to the following rule:

At t=1

if $p(H_o|y_1,u_1) \le \alpha_1$, decide H_1 ;

if $p(H_o|y_1,u_1) \ge \beta_1$, decide H_o ;

if $\alpha_1 < p(H_o|y_1,u_1) < \beta_1$, continue to take observations.

At t=2

if $p(H_o|y_u^2) \ge \ell_2$, decide H_o ;

if $p(H_o|y_u^2) < \ell_2$, decide H_1 ; (5-26)

where

$$y_u^2 \triangleq \{y_1,y_2, u_1,u_2\} \text{ and } y_u^1 \triangleq \{y_1,u_1\}$$

Proof:

The objective is to minimize

$$R \triangleq E\{J(v_t,H)\} \quad (5\text{-}27)$$

The decision rule at time 1 for DM_1 is obtained as follows. Writing (5-27) explicitly, we have

$$R = \sum_{H,u^2,v^2} \int_{x^2,y^2} p(u^2,v^2,x^2,y^2,H)\ J(v_t,H)\ dx^2dy^2 \quad (5\text{-}28)$$

We define

$$\Sigma_2(u^2,H) = \sum_{v^2} \int_{y^2} p(v^2,y^2|u^2,H)\ J(v_t,H)\ dy^2 \tag{5-29}$$

Substituting (5-29) in (5-28), we get

$$R = \sum_{H,u^2} \int_{x^2} p(u^2,x^2,H)\ dx^2\ \Sigma_2(u^2,H) \quad , \tag{5-30}$$

where

$$p(u^2,x^2,H) = p(u_2,x_2|u_1,x_1,H)\ p(u_1|x_1,H) \cdot p(H|x_1)\ p(x_1) \quad . \tag{5-31}$$

Using (5-11), (5-31) becomes

$$p(u^2,x^2,H) = p(u_2,x_2|u_1,x_1,H)\ p(u_1|x_1)\ p(H|x_1)\ p(x_1) \quad . \tag{5-32}$$

Substituting (5-32) in (5-30), and integrating with respect to all values of x_2, we have

$$R = \sum_{H,u^2} \int_{x_1} p(u_2|u_1,x_1,H)\ p(u_1|x_1)\ p(H|x_1)\ p(x_1) \cdot dx_1\ \Sigma_2(u^2,H) \tag{5-33}$$

Expanding with respect to u_1 and ignoring a constant term, we have

$$\int_{x_1} p(x_1)\ p(u_1=0|x_1) \sum_H p(H|x_1) \{ \sum_{u_2} p(u_2|u_1=0,H)\ \Sigma_2(u_1=0,u_2,H) - \sum_{u_2} p(u_2|u_1=1,H)\ \Sigma_2(u_1=0,u_2,H) \} \tag{5-34}$$

Expression (5-34) is minimized if

$$p(u_1=0|x_1) = \begin{cases} 0 \text{ if } \sum_H p(H|x_1)\{\sum_{u_2} p(u_2|u_1=0,H)\ \Sigma_2(u_1=0,u_2,H) \\ \qquad - \sum_{u_2} p(u_2|u_1=1,H)\ \Sigma_2(u_1=1,u_2,H)\} \geq 0 \\ 1 \quad \text{otherwise} \end{cases} \tag{5-35}$$

Expression (5-35) can be written in a more convenient form as

$$p(u_1=0|x_1) = \begin{cases} 0 \text{ if } E_{H|x_1} \{E_{u_2|u_1=0,H}\ \Sigma_2(u_1=0,u_2,H) \\ \qquad - E_{u_2|u_1=1,H}\Sigma_2(u_1=1,u_2,H)\} \geq 0 \\ 1 \quad \text{otherwise}\ . \end{cases} \tag{5-36}$$

We define

$$\Sigma_1(u_1,H) \triangleq E_{u_2|u_1,H}\ \Sigma_2(u_1,u_2,H) \tag{5-37}$$

From (5-14), it follows that

$$\begin{aligned} \Sigma_1(1,H_o) &\geq \Sigma_1(0,H_o) \\ \Sigma_1(0,H_1) &\geq \Sigma_1(1,H_1) \end{aligned} \tag{5-38}$$

Using (5-38), we then obtain the following decision rule at DM_1 at time 1,

$$\Lambda_{u_1}(x^1) \underset{u_1=0}{\overset{u_1=1}{\gtrless}} t_{u_1} , \tag{5-39}$$

which is the desired result as stated in Theorem 5.2(a).

Next, the decision rule at time t=2 for DM_1 is obtained. Starting with expression (5-30), we have

$$R = \sum_{H,u^2} \int_{x^2} p(u^2,x^2,H)\, dx^2\, \Sigma_2(u^2,H) . \tag{5-40}$$

Using Bayes rule, we have

$$p(u^2,x^2,H) = p(u^2|x^2,H)\, p(H|x^2)\, p(x^2) . \tag{5-41}$$

Using (5-11), we have

$$p(u^2,x^2,H) = p(u^2|x^2)\, p(H|x^2)\, p(x^2) . \tag{5-42}$$

Furthermore,

$$p(u^2|x^2) = p(u_2|u_1,x^2)\, p(u_1|x^2) . \tag{5-43}$$

Using (5-11), we have

$$p(u^2|x^2) = p(u_2|u_1,x^2)\, p(u_1|x_1) \tag{5-44}$$

Substituting (5-44) and (5-42) in (5-40), we obtain

$$R = \sum_{H,u^2} \int_{x^2} p(H|x^2)\, p(u_2|u_1,x^2)\, p(u_1|x_1)\, p(x^2) \cdot \Sigma_2(u^2,H)\, dx^2 . \tag{5-45}$$

Explicitly summing over u_2, and neglecting a constant term, we have

$$R = \sum_{u_1} \int_{x^2} p(u_2=0|u_1,x^2)\ p(u_1|x_1)\ p(x^2)\ dx^2$$

$$\{ \sum_{H} p(H|x^2)\ [\Sigma_2(u_1,u_2=0,H) - \Sigma_2(u_1,u_2=1,H)]\} \qquad (5\text{-}46)$$

Expression (5-46) is minimized by setting

$$p(u_2=0|u_1,x^2) = \begin{cases} 0 & \text{if } \sum_H p(H|x^2)[\Sigma_2(u_1,u_2=0,H) - \Sigma_2(u_1,u_2=1,H)] \geq 0 \\ 1 & \text{otherwise} \end{cases} \qquad (5\text{-}47)$$

We assume that

$$\begin{aligned} \Sigma_2(u_1,u_2=0,H_1) &\geq \Sigma_2(u_1,u_2=1,H_1) \\ \Sigma_2(u_1,u_2=1,H_o) &\geq \Sigma_2(u_1,u_2=0,H_o) \end{aligned} \qquad (5\text{-}48)$$

Note that (5-48) follows from (5-14). Expanding (5-47) over H, we obtain the decision rule for DM_1 at time 2

$$\Lambda_{u_2}(x^2) \underset{u_2=1}{\overset{u_2=1}{\gtrless}} t_{u_2}(u_1) \qquad (5\text{-}49)$$

which is the desired result of Theorem 5.2(b).

We need not prove part (c) of the theorem in detail because the results can be obtained by a straightforward application of the results of Theorem 5-1. This completes the proof of Theorem 5-2. Q.E.D.

In Theorem 5-2, we presented the results for the two-stage decentralized sequential detection problem for a tandem configuration. The above results can easily be extended to the multi-stage decentralized sequential detection problem. The results are stated in Theorem 5-3 without proof.

Theorem 5-3

Consider the configuration shown in Fig. 5.1. The optimal decision rules for the multi-stage decentralized sequential detection problem are given as follows:

(a) The optimal local decision rules at DM_1 for $t=1,\ldots,N$, are described by the LRT's given below.

At $t=1$,
$$\Lambda_{u_1}(x^1) \underset{u_1=0}{\overset{u_1=1}{\gtrless}} t_{u_1} \qquad (5\text{-}50)$$

For $1<t<N$
$$\Lambda_{u_t}(x^t) \underset{u_t=0}{\overset{u_t=1}{\gtrless}} t_{u_t}(u^{t-1}) \qquad (5\text{-}51)$$

At $t=N$,
$$\Lambda_{u_N}(x^N) \underset{u_N=0}{\overset{u_N=1}{\gtrless}} t_{u_N}(u^{N-1}) \qquad (5\text{-}52)$$

where

$$t_{u_1} \triangleq \frac{\Sigma_1(1, H_o) - \Sigma_1(0, H_o)}{\Sigma_1(0, H_1) - \Sigma_1(1, H_1)} \quad t=1$$

$$t_{u_t}(u^{t-1}) \triangleq \frac{\Sigma_t(u^{t-1}, 1, H_o) - \Sigma_t(u^{t-1}, 0, H_o)}{\Sigma_t(u^{t-1}, 0, H_1) - \Sigma_t(u^{t-1}, 1, H_1)}$$

$$1 < t \leq N \tag{5-53}$$

$$\Sigma_t(u^t, H) \triangleq E_{u_N, \ldots, u_{t+1}|u^t, H} \Sigma_N(u^N, H) \tag{5-54}$$

$$\Sigma_N(u^N, H) \triangleq \sum_{v^N} p(v^N|u^N, H) J(v_t, H) \tag{5-55}$$

$$= E_{v^N|u^N, H} J(v_t, H) . \tag{5-56}$$

and

$$\Lambda_{u_t}(x^t) = \frac{p(H_1|x^t)}{p(H_o|x^t)} . \tag{5-57}$$

(b) The optimum decision rules at DM_2 are described by $2N-1$ thresholds $\alpha_1, \alpha_2, \ldots, \alpha_{N-1}, \beta_1, \beta_2, \ldots, \beta_{N-1}$, and ℓ_N. The thresholds satisfy the following equations.

At t=k (k=1,...,N-1)

$$\alpha_k J(0,1) + (1 - \alpha_k) J(1,1)$$

$$= E_{y_{k+1}, u_{k+1}} [\tilde{J}_{k+1} (\frac{\alpha_k p(y_{k+1}, u_{k+1}|H_o)}{\alpha_k p(y_{k+1}, u_{k+1}|H_o) + (1-\alpha_k) p(y_{k+1}, u_{k+1}|H_1)})] \tag{5-58}$$

and

$$\beta_k J(0, 0) + (1 - \beta_k) J(1, 1)$$

$$= E_{y_{k+1}, u_{k+1}}[\hat{J}_{k+1} (\frac{\beta_k p(y_{k+1}, u_{k+1}|H_o)}{\beta_k p(y_{k+1}, u_{k+1}|H_o) + (1 - \beta_k) p(y_{k+1}, u_{k+1}|H_1)})] \tag{5-59}$$

<u>At t = N</u>

$$\ell_N = \frac{J(0, 1) - J(1, 1)}{J(1, 0) + J(0, 1) - J(0, 0) - J(1, 1)} \tag{5-60}$$

At each time t, DM_2 evaluates $p(H_o|y_u^t)$ with α_k and β_k (or ℓ_N if t=N) and decides to stop or continue according to the following rule.

<u>At t = k</u>

if	$p(H_o\|y_u^t) \leq \alpha_t$	decide H_o
if	$p(H_o\|y_u^t) \geq \beta_t$	decide H_1
if	$\alpha_t \leq p(H_o\|y_u^t) \leq \beta_t$	continue to take observations (5-61)

<u>At t = N</u>

if	$p(H_o\|y_u^t) \geq \ell_N$	decide H_o
if	$p(H_o\|y_u^t) < \ell_N$	decide H_1 (5-62)

Next, we consider the decentralized sequential probability ratio test problem using the Neyman-Pearson approach.

5.4. A Decentralized Sequential Probability Ratio Test

Consider a distributed detection network with a data fusion center as shown in Figure 5.2. We solve the problem of decentralized sequential probability ratio test. We use the Neyman-Pearson approach for system optimization.

The problem is formulated as follows:

(A.13) Consider two hypotheses $\{H_o, H_1\}$ where the a priori probability of H_o is given by

$$\text{Prob } [H_o] = P_o$$

(A.14) DM_1 makes a decision u_t at time t given by

$$u_t = \gamma_u(x^t) \quad , \quad u_t \in \{0, 1\} \tag{5-63}$$

where $\gamma_u(\cdot)$ is the decision rule at DM_1, and $x^t \triangleq (x_1, \ldots, x_t)$ represents the set of observations at DM_1 up to time t.

(A.15) DM_2 makes a decision v_t at time t given by

$$v_t = \gamma_v(y^t) \quad , \quad v_t \in \{0, 1\} \tag{5-64}$$

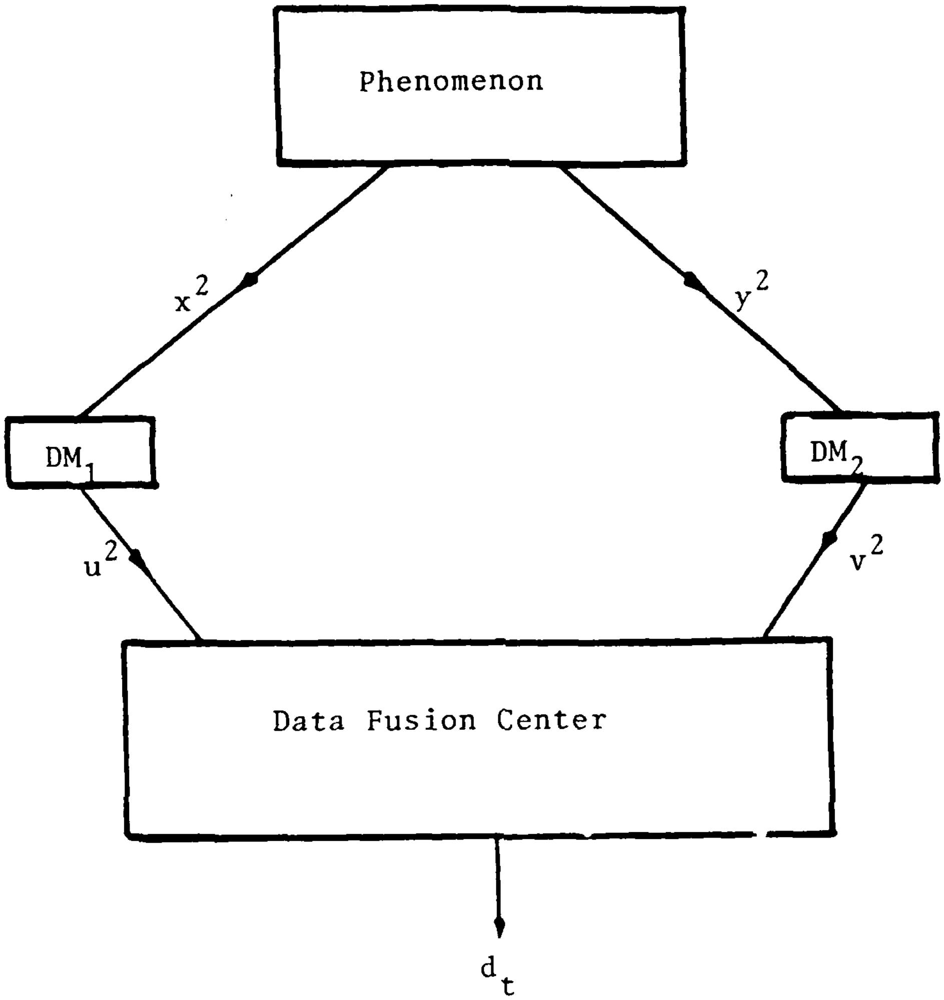


Figure 5.2 A Distributed Detection System with a Data Fusion Center.

where $\gamma_v(\cdot)$ is the decision rule at DM_2 and $y^t \triangleq (y_1,\ldots,y_t)$ represents the set of observations at DM_2 up to time t.

(A.16) The observations $\{x_1,\ldots,x_t\}$ and $\{y_1,\ldots,y_t\}$ are assumed to be mutually independent given the hypothesis H, i.e.,

$$p(x_1,\ldots,x_t, y_1,\ldots,y_t|H) = \prod_{i=1}^{t} p(x_i|H) \prod_{j=1}^{t} p(y_j|H) . \tag{5-65}$$

(A.17) The data fusion center performs the centralized sequential probability ratio test which yields the final decision d_t given by

$$d_t = \gamma_{d_t}(u_t, v_t) , \tag{5-66}$$

where (u_t, v_t) are the decisions at DM_1 and DM_2 at time t. Under (A.13)-(A.17), the decentralized sequential probability ratio test problem is stated as follows:

(P7)
At each decision maker and at the data fusion center, find the decision rules $\gamma_u(\cdot)$, $\gamma_v(\cdot)$, and $\gamma_{d_t}(\cdot)$ so as to minimize the function F given by

$$F \triangleq Pr[d_t=0|H_1] + \lambda\ [Pr[d_t=1|H_o] - \alpha]$$

under the constraint that (5-67)

$$Pr[d_t=1|H_o] = \alpha' \leq \alpha$$

and

$\lambda \triangleq$ the Lagrange multiplier

The solution to problem (P7) is given by the following theorem:

Theorem 5.4

For the system shown in Figure 5.2, the optimal decision rules for the two-stage decentralized sequential probability ratio test problem are given by

(a) The optimal decision rule at DM_1 at time 1 is described by the LRT.

$$\Lambda_{u_1}(x^1) \triangleq \frac{p(x^1|H_1)}{p(x^1|H_o)} \underset{u_1=0}{\overset{u_1=1}{\gtrless}} \lambda_{u_1} , \tag{5-68}$$

where

$$\lambda_{u_1} \triangleq \lambda \frac{\Sigma_1^0(u_1=1,\ H_o) - \Sigma_1^0(u_1=0,\ H_o)}{\Sigma_1^1(u_1=0,\ H_1) - \Sigma_1^1(u_1=1,\ H_1)} \tag{5-69}$$

and

$$\Sigma_1^i(u_1,\ H_j) = E_{u_2|u_1,H_j}\ \Sigma_2^i(u_1,u_2,H_j) ,$$

$$\Sigma_2^i(u_1,u_2,H_j) = \Sigma_{v^2}\ p(v^2|H_j)\ p(d_t=i|u^2,v^2) . \tag{5-70}$$

We also have

$$p(u_1=0|H_1) = \alpha_{u_1}$$

$$= \frac{\alpha - \Sigma_1^0(u_1=0,\ H_o)}{\Sigma_1^0(u_1=1,\ H_o) - \Sigma_1^0(u_1=0,\ H_o)} . \tag{5-71}$$

(b) The optimal decision rule of DM_1 at time 2 is described by the LRT

$$\Lambda_{u_2}(x^2) \triangleq \frac{p(x^2|H_1)}{p(x^2|H_o)} \underset{u_2=0}{\overset{u_2=1}{\gtrless}} \lambda_{u_2} , \quad (5\text{-}72)$$

where

$$\lambda_{u_2} = \lambda \frac{\Sigma_2^0(u_1,u_2=1,H_o) - \Sigma_2^0(u_1,u_2=0,H_o)}{\Sigma_2^1(u_1,u_2=0,H_1) - \Sigma_2^1(u_1,u_2=1,H_1)} \quad (5\text{-}73)$$

$$p(u_1,u_2=1|H_o) = \frac{\alpha - \Sigma_2^0(u_1,u_2=0,H_o)}{\Sigma_2^0(u_1,u_2=1,H_o) - \Sigma_2^0(u_1,u_2=0,H_o)} \quad (5\text{-}74)$$

and

$$p(u_1=0,u_2=0|H) = \iint_D p(x_1|H)\, p(x_2|H)\, dx_1 dx_2$$

where $D \triangleq \{x^2 : \Lambda_{u_2}(x^2) \leq t_{u_2}(\cdot) \text{ and } \Lambda_{u_1}(x^1) \leq t_{u_1}\}$ (5-75)

(c) The optimal decision rules at DM_2 at times 1 and 2 are obtained by interchanging x with y and u with v in equations (5-68) through (5-75).

(d) The decision rule at the data fusion center is given by the following equation :

<u>At t=k</u> (k=1,2,...,N-1)

If $\Lambda_{d_t} \triangleq \frac{p(u_t, v_t|H_1)}{p(u_t, v_t|H_o)} \geq T_1$ decide H_1

If $\Lambda_{d_t} \leq T_o$ decide H_o

$$\text{If} \quad T_o < \Lambda_{d_t} < T_1 \qquad \text{continue to take observations} \tag{5-76}$$

where

$$T_o \triangleq \frac{P_M}{1 - P_F} \tag{5-77}$$

and

$$T_1 \triangleq \frac{1 - P_M}{P_F} \tag{5-78}$$

$\{P_M, P_F\}$ are the pre-assigned probabilities of miss and false alarm respectively.

At t=N

$$\text{If} \quad \Lambda_{d_N} \triangleq \frac{p(u_N, v_N | H_1)}{p(u_N, v_N | H_o)} \underset{d_t = 0}{\overset{d_t = 1}{\gtrless}} T_f \tag{5-79}$$

where T_f is given and satisfies

$$T_o \leq T_f \leq T_1 .$$

Proof:

The decision rule at time 1 for DM_1 is obtained as follows:

The objective function

$$F = Pr\{d_t = 0 | H_1\} + \lambda [Pr\{d_t = 1 | H_o\} - \alpha] \tag{5-80}$$

is expanded with respect to u^2 and v^2 to get

$$F = \sum_{u^2,v^2} p(d_t=0,u^2,v^2|H_1) + \lambda[\sum_{u^2,v^2} p(d_t=1,u^2,v^2|H_o) - \alpha] \tag{5-81}$$

Using Bayes rule, (5-81) yields

$$F = \sum_{u^2,v^2} p(d_t=0|u^2,v^2)\ p(u^2,v^2|H_1)$$

$$+ \lambda[\sum_{u^2,v^2} p(d_t=1|u^2,v^2)\ p(u^2,v^2|H_o) - \alpha] \tag{5-82}$$

Using (5-63) through (5-65), we have

$$F = \sum_{u^2} p(u^2|H_1) \sum_{v^2} p(v^2|H_1)\, p(d_t=0|u^2,v^2)$$

$$+ \lambda[\sum_{u^2} p(u^2|H_o) \sum_{v^2} p(v^2|H_o)\ p(d_t=1|u^2,v^2)-\alpha] \tag{5-83}$$

We define

$$\Sigma_2^0(u^2,H_o) = \sum_{v^2} p(v^2|H_o)\ p(d_t=0|u^2,v^2) \tag{5-84}$$

and

$$\Sigma_2^1(u^2,H_1) = \sum_{v^2} p(v^2|H_1)\ p(d_t=1|u^2,v^2) \tag{5-85}$$

Substituting (5-84) and (5-85) in (5-83), we get

$$F = \sum_{u_2,u_1} p(u_1,u_2|H_1)\ \Sigma_2^1(u_1,u_2,H_1)$$

$$+ \lambda\ [\sum_{u_2,u_1} p(u_1,u_2|H_o)\Sigma_2^0(u_1,u_2,H_o) - \alpha] \tag{5-86}$$

Expanding with respect to u_1, we have

$$F = \sum_{u_2} p(u_1=0|H_1)\ p(u_2|u_1=0,H_1)\ \Sigma_2^1(u_1=0,u_2,H_1)$$

$$+ \sum_{u_2} p(u_1=1|H_1)\ p(u_2|u_1=1,H_1)\ \Sigma_2^1(u_1=1,u_2,H_1)$$

$$+ \lambda[\sum_{u_2} p(u_1=1|H_o)\ p(u_2|u_1=1,H_o)\ \Sigma_2^0(u_1=1,u_2,H_o)$$

$$+ \sum_{u_2} p(u_1=0|H_o)\ p(u_2|u_1=0,H_o)\ \Sigma_2^0(u_1=0,u_2,H_o)$$

$$- \alpha] \qquad (5\text{-}87)$$

Expression (5-87) can be written as

$$F = p(u_1=0|H_1)\ D_{o1} + C_{o1}$$

$$+\lambda\ [p(u_1=1|H_o)\ D_{1o} + C_{1o} - \alpha] \qquad (5\text{-}88)$$

where

$$D_{o1} \triangleq \sum_{u_2} p(u_2|u_1=0,H_1)\ \Sigma_2^1(u_1=0,u_2,H_1)$$

$$- \sum_{u_2} p(u_2|u_1=1,H_1)\ \Sigma_2^1(u_1=1,u_2,H_1) \qquad (5\text{-}89)$$

or

$$D_{o1} = E_{u_2|u_1=0,H_1}\ \Sigma_2^1(u_1=0,u_2,H_1)$$

$$- E_{u_2|u_1=1,H_1}\ \Sigma_2^1(u_1=1,u_2,H_1) \qquad (5\text{-}90)$$

$$C_{o1} \triangleq E_{u_2|u_1=1,H_1} \Sigma_2^1(u_1=1,u_2,H_1) \qquad (5\text{-}91)$$

$$D_{1o} = \sum_{u_2} p(u_2|u_1=1,H_o) \Sigma_2^0(u_1=1,u_2,H_o)$$

$$- \sum_{u_2} p(u_2|u_1=0,H_o) \Sigma_2^0(u_1=0,u_2,H_o) \qquad (5\text{-}92)$$

or

$$D_{1o} = E_{u_2|u_1=1,H_o} \Sigma_2^0(u_1=1,u_2,H_o)$$

$$- E_{u_2|u_1=0,H_o} \Sigma_2^0(u_1=0,u_2,H_o) \qquad (5\text{-}93)$$

$$C_{1o} = E_{u_2|u_1=0,H_o} \Sigma_2^0(u_1=0,u_2,H_o) \qquad (5\text{-}94)$$

Ignoring the constant term C_{o1}, and dividing by D_{o1}, equation (5-87) gives

$$F_{u_1} = p(u_1=0|H_1) + \lambda_{u_1}[p(u_1=1|H_o) - \alpha_{u_1}], \qquad (5\text{-}95)$$

where

$$\lambda_{u_1} \triangleq \lambda \frac{D_{1o}}{D_{o1}} \qquad (5\text{-}96)$$

and

$$\alpha_{u_1} \triangleq \frac{\alpha - C_{1o}}{D_{1o}} \quad . \qquad (5\text{-}97)$$

By minimizing (5-95), we obtain the following LRT:

$$\frac{p(x^1|H_1)}{p(x^1|H_1)} \underset{u_1=0}{\overset{u_1=1}{\gtrless}} \lambda_{u_1} \quad , \tag{5-98}$$

$$p(u_1=0|H_1) = \alpha_{u_1} \tag{5-99}$$

This completes the proof of Theorem 5-4 (a).

The decision rule at DM_1 at time 2 is obtained as follows. Starting with expression (5-86), and using the fact that decision u_1 is known at time 2, we get

$$F = \sum_{u_2} p(u_1,u_2|H_1)\ \Sigma_2^1(u_1,u_2,H_1)$$

$$+ \lambda[\sum_{u_2} p(u_1,u_2|H_o)\ \Sigma_2^0(u_1,u_2,H_o) - \alpha] \tag{5-100}$$

Expanding F with respect to u_2 and ignoring the constant term, we have

$$F = p(u_1,u_2=0|H_1)\ [\Sigma_2^1(u_1,u_2=0,H_1) - \Sigma_2^1(u_1,u_2=1,H_1)]$$

$$+ \lambda[p(u_1,u_2=1|H_o)\ [\Sigma_2^0(u_1,u_2=1,H_o) - \Sigma_2^0(u_1,u_2=0,H_o)]$$

$$+ \Sigma_2^0(u_1,u_2=0,H_o) - \alpha] \tag{5-101}$$

Expression (5-101) can be written as

$$F_{u_2} = p(u_1,u_2=0|H_1) + \lambda_{u_2} [p(u_1,u_2=1|H_o) - \alpha_{u_2}] \quad (5\text{-}102)$$

where

$$\lambda_{u_2} \triangleq \lambda \cdot \frac{\Sigma_2^0(u_1,u_2=1,H_o) - \Sigma_2^0(u_1,u_2=0,H_o)}{\Sigma_2^1(u_1,u_2=0,H_1) - \Sigma_2^1(u_1,u_2=1,H_1)} \quad (5\text{-}103)$$

$$\alpha_{u_2} \triangleq \frac{\alpha - \Sigma_2^0(u_1,u_2=0, H_o)}{\Sigma_2^0(u_1,u_2=1,H_o) - \Sigma_2^0(u_1,u_2=0,H_o)} \quad (5\text{-}104)$$

Minimizing (5-102) yields the result in Theorem 5.4 (b).

The decision rules at DM_2 at time 1 and 2 are obtained using similar derivations as in part (a) and (b).

We need not prove part (d) of the theorem because the results are a direct application of the classical results on the sequential probability ratio test. This completes the proof of Theorem 5-4. Q.E.D.

In Theorem 5-4, we solved the two-stage decentralized sequential probability ratio test problem. The above results can easily be extended to the multi-stage decentralized sequential probability ratio test. The results are stated in Theorem 5-5 without proof.

Theorem 5-5

Consider the configuration shown in Figure 5.2. The optimal decision rules for the multi-stage decentralized sequential detection problem are given as follows:

(a) The optimal local decision rules at DM_1 for $t=1,\dots,N$ are described by the LRT's given below :

$$\Lambda_{u_1}(x^1) \underset{u_1=0}{\overset{u_1=1}{\gtrless}} \lambda_{u_1} \qquad t=1 \quad , \tag{5-105}$$

$$\Lambda_{u_t}(x^t) \underset{u_t=0}{\overset{u_t=1}{\gtrless}} \lambda_{u_t}(u^{t-1}) \quad 1 < t < N \, , \tag{5-106}$$

$$\Lambda_{u_N}(x^N) \underset{u_N=0}{\overset{u_N=1}{\gtrless}} \lambda_{u_N}(u^{N-1}) \qquad t=N \quad , \tag{5-107}$$

where

$$\lambda_{u_1} = \lambda \frac{\Sigma_1^0(1,H_o) - \Sigma_1^0(0,H_o)}{\Sigma_1^1(0,H_1) - \Sigma_1^1(1,H_1)} \quad , \tag{5-108}$$

$$\lambda_{u_t}(u^{t-1}) \triangleq \lambda \frac{\Sigma_t^0(u^{t-1},1,H_o) - \Sigma_t^0(u^{t-1},0,H_o)}{\Sigma_t^1(u^{t-1},0,H_1) - \Sigma_t^1(u^{t-1},1,H_1)} \tag{5-109}$$

$$1 \le t \le N \quad ,$$

$$\Sigma_t^i(u^t,H) \triangleq E_{u_N,\dots,u_{t+1}|u^t,H} \; \Sigma_N^i(u^N,H) \tag{5-110}$$

and

$$\Sigma_N^i(u^N,H) \triangleq \sum_{v^N} p(v^N|H) \; p(d_t=i|u^N,v^N) \, , \tag{5-111}$$

or

$$\Sigma_N^i(u^N,H) = E_{v^N|H} \; p(d_t=i|u^N,v^N) \quad . \tag{5-112}$$

Finally ,

$$\alpha_{u_t} \triangleq \frac{\alpha - \Sigma_t^0(u^{t-1},0,H_o)}{\Sigma_t^0(u^{t-1},1,H_o) - \Sigma_t^0(u^{t-1},0,H_o)} . \qquad (5\text{-}113)$$

(b) The optimal decision rules at DM_2 at all times are obtained by interchanging x with y and u with v in equations (5-105) through (5-113).

(c) The decision rule at the data fusion center is given by Theorem 5.4(d).

5-5. Discussion

In this chapter, we have considered two decentralized sequential detection problems namely: the Bayesian sequential hypothesis testing problem for a tandem topology network, and the decentralized sequential probability ratio test problem. In both problems, we derived the decision rules at each decision maker at each time stage. Under the assumption that the observations are conditionally independent; the decision rules are likelihood ratio test (LRT). Furthermore, the threshold of each decision maker at time t depends on all of its previous decisions. However, all threshold equations are coupled at all times. One needs to solve these equations simultaneously in order to obtain a solution for the thresholds. The coupling between the threshold equations causes considerable complexity in the computation of the optimal thresholds. However, the results obtained here could be used to obtain some simpler suboptimal solutions for these decentralized sequential detection problems by making suitable simplifications.

CHAPTER SIX

SUMMARY AND SUGGESTIONS FOR FUTURE RESEARCH

6.1 Summary

In this Report, we have considered some hypothesis testing problems where the detection network consists of a multitude of geographically distributed decision makers. We investigated distributed detection problems for various network topologies. First, we considered two schemes for data fusion namely, centralized data fusion and distributed data fusion. Next, the distributed detection problem with distributed data fusion was solved using the Bayesian approach as well as the Neyman-Pearson approach. Then, the Neyman-Pearson detection problem and the sequential hypothesis testing problem for a tandem topology was solved. Finally, the distributed sequential probability ratio test was investigated. In all cases, we have derived the optimal strategies at each decision maker. The decision rules at each decision maker turn out to be LRT's where the thresholds are described by a set of coupled nonlinear equations. The simultaneous solution of these equations yields the set of thresholds. There may be several local solutions. Each must be checked to assure that a global minimum is found.

6.2 Suggestions for Future Research

Optimal decisions rules have been derived for various distributed detection problems in this Report. The thresholds associated with the decision rules are obtained through the solution of coupled nonlinear equations. The computation of the numerical values of the thresholds is extremely difficult and involved. Therefore, efficient computational algorithms for the solution of the threshold equations should be developed.

Throughout this Report, we have assumed that the observations are conditionally independent as is the case in most of the published literature. Hypothesis testing problems with dependent observations should be investigated due to its wide practical applications.

Different structures (topology, communication protocols, etc...) of distributed detection systems should also be studied, i.e., how should one organize the various decision makers in a detection network in order to maximize the performance of the overall system?

REFERENCES

[1] Van Trees, H. L., Detection, Estimation and Modulation Theory, New York, Wiley, 1969.

[2] Tenney, R. R., and Sandell, N. R., "Detection with Distributed Sensors," IEEE Transactions on Aerospace and Electronic Systems, Vol. AES 17, pp. 501-509, 1986.

[3] Sadjadi, F. A., "Hypothesis Testing in Distributed Environment," IEEE Trans. on Aerospace and Electronic Systems, Vol. AES 32, pp. 134-137, March 1986.

[4] Lauer, G. and Sandell, Jr., N.R., "Distributed Detection of Known Signals in Correlated Noise," Report 160, ALPHATECH, Burlington, Mass., March 1982.

[5] Ekchian, L. K. and Tenney, R. R., "Detection Networks," Proceeding of the 21st. IEEE Conference on Decision and Control, pp. 686-691, Orlando, Fla., December 1982.

[6] Ekchian, L. K., and Tenney, R. R., "Recursive Solution of Distributed Detection," American Control Conference, pp. 1338-1339, June 1983.

[7] Ho, Y. C., "Team Decision Theory and Information Structures," Proceedings of the IEEE, Vol. 68, No. 6, pp. 644-654, June 1980.

[8] Kushner, H. J., and Pacut, A., "A Simulation Study of a Decentralized Detection Problem," IEEE Transactions on Automatic Control, 27, 5, pp. 1116-1119, Oct. 1982.

[9] Tsitsiklis, J. N., and Athans, M., "On the Complexity of Decentralized Decision Making and Detection Problems," IEEE Trans. on Automatic Control, Vol. AC-30, No. 5, pp. 440-446, May 1985.

[10] Teneketzis, D., "The Decentralized Wald Problem," IEEE International Large-Scale Systems Symposium, Virginia Beach, VA., October 1982.

[11] Hashemipour, H. R., and Rhodes, I. B., "Decentralized Sequential Detection," Department of Electrical and Computer Engineering, University of California, Santa Barbara, California.

[12] Teneketzis, D., "The Decentralized Quickest Detection Problem," IEEE International Large-Scale Systems Symposium, Virginia Beach, VA, October 1982.

[13] Wald, A., Sequential Analysis, Dover, 1973.

[14] Tsitsiklis, J. N., "On Threshold Rules in Decentralized Detection," Proceedings at the 25th Conference on Decision and Control, Athens, Greece, December 1986, pp. 232-236.

[15] Hoballah, I., "On the Design and Optimization of Distributed Signal Detection and Parameter Estimation Systems," Syracuse University, Ph.D. Thesis, 1986.

[16] Sandell, N. R., and Athans, M., "Solution of Some Nonclassical LQG Stochastic Decision Problems," IEEE Trans. on Automatic Control, Vol. AC-19, pp. 108-116, April 1974.

[17] Washburn, R. B., and Teneketzis, D., "Asymptotic Agreement among Communicating Decision Makers," ALPHATECH Report, ALPHATECH Inc., 1984.

[18] Bertsekas, D. P., Dynamic Programming and Stochastic Control, Academic Press, 1976.

[19] DeGroot, M. H., Optimal Statistical Decisions, McGraw-Hill, 1970.

[20] Papoulis, A., Probability,Random Variables and Stochastic Processes, McGraw-Hill, 1984.

[21] Srinivasan, R., "Distributed Radar Detection Theory," IEE Proceedings, Vol. 133, Pt. F, No. 1, Feburary 1986, pp. 55-60.

☆ U.S. Government Printing Office 1988 511-117/64063

END
DATE
FILMED
8-88
DTIC