ABSTRACT

ADAPTIVE DISTRIBUTED DETECTION WITH APPLICATIONS TO CELLULAR CDMA

by
Jian-Guo Chen

Chair and Varshney have derived an optimal rule for fusing decisions based on the Bayesian criterion. To implement the rule, probabilities of detection P_D and false alarm P_F for each detector must be known, which is not readily available in practice. This dissertation presents an adaptive fusion model which estimates the P_D and P_F adaptively by a simple counting process. Since reference signals are not given, the decision of a local detector is arbitrated by the fused decision of all the other local detectors. Adaptive algorithms for both equal probable and unequal probable sources, for independent and correlated observations are developed and analyzed, respectively. The convergence and error analysis of the system are analytically proven and demonstrated by simulations. In addition, in this dissertation, the performance of four practical fusion rules in both independent and correlated Gaussian noise is analyzed, and compared in terms of their Receiver Operating Characteristics (ROCs). Various factors that affect the fusion performance are considered in the analysis. By varying the local decision thresholds, the ROCs under the influence of the number of sensors, signal-to-noise ratio (SNR), the deviation of local decision probabilities, and correlation coefficient, are computed and plotted, respectively. Several interesting and key observations on the performance of fusion rules are drawn from the analysis.

As an application of the above theory, a decentralized or distributed scheme in which each fusion center is connected with three widely spaced base stations is proposed for digital cellular code-division multi-access communications. Detected results at each base station are transmitted to the fusion center where the final decision is made by optimal fusion. The theoretical analysis shows that this novel

structure can achieve an error probability at the fusion center which is always less than or equal to the minimum of the three respective base station. The performance comparison for binary coherent signaling in Rayleigh fading and log-normal shadowing demonstrates that the decentralized detection has a significant increased system capacity over conventional macro selection diversity. This dissertation analyzes the performance of the adaptive fusion method for macroscopic diversity combination in the wireless cellular environment when the error probability information from each base station detection is not available. The performance analysis includes the derivation of the minimum achievable error probability. An alternative realization with lower complexity of the optimal fusion scheme by using selection diversity is also proposed. The selection of the information bit in this realization is obtained either from the most reliable base station or through the majority rule from the participating base stations.

ADAPTIVE DISTRIBUTED DETECTION WITH APPLICATIONS TO CELLULAR CDMA

by
Jian-Guo Chen

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

Department of Electrical and Computer Engineering

January 1997

APPROVAL PAGE

ADAPTIVE DISTRIBUTED DETECTION WITH APPLICATIONS TO CELLULAR CDMA

Jian-Guo Chen

Dr. Nirwan Ansari, Dissertation Advisor — Date
Associate Professor of Electrical and Computer Engineering, NJIT

Dr. Zoran Siveski, Committee Member — Date
Assistant Professor of Electrical and Computer Engineering, NJIT

Dr. Joseph Frank, Committee Member — Date
Associate Professor of Electrical and Computer Engineering, NJIT

Dr. Jack Holtzman, Committee Member — Date
Professor of Electrical and Computer Engineering, Rutgers University

Dr. Zoi-Heleni Michalopoulou, Committee Member — Date
Assistant Professor of Department of Mathematics, NJIT

BIOGRAPHICAL SKETCH

Author: Jian-Guo Chen

Degree: Doctor of Philosophy

Date: January 1997

Undergraduate and Graduate Education:

- Doctor of Philosophy in Electrical Engineering,
  New Jersey Institute of Technology, Newark, NJ, 1997
- Master of Science in Electrical Engineering,
  Xidian University, Xi'an, Shaanxi, P. R. China, 1988
- Bachelor of Science in Electrical Engineering,
  Xidian University, Xi'an, Shaanxi, P. R. China, 1985

Major: Electrical Engineering

Presentations and Publications:

N. Ansari, E. Hou, B. Zhu, and J. G. Chen, "An adaptive fusion model by reinforcement learning for distributed detection systems," IEEE Trans. on Aerospace and Electronic System, vol. 32, no. 2, April, 1996, pp 524-531.

J. G. Chen, N. Ansari, and Z. Siveski, "Distributed detection for cellular CDMA," IEE Electronics Letters, vol. 32, no. 3, Feb. 1996, pp 169-171.

N. Ansari, J. G. Chen, and Y. Zhang, "Distributed detection by reinforcement learning for unequiprobable sources," Proc. 29th Conference on Information Science and Systems, Baltimore, MD, Mar. 1995, pp 612-617.

J. G. Chen and N. Ansari, "Fusion of correlated decisions by reinforcement learning," Proc. 30th Conference on Information Science and Systems, Princeton, NJ, Mar. 1996, pp 1071-1076.

J. G. Chen and N. Ansari, "Adaptive decision fusion for practical environments," Workshop on foundations of Information/Decision Fusion: Applications to Engineering Problems, Washington D. C., Aug., 1996, pp 33-39.

N. Ansari, J. G. Chen, and Y. Zhang, "Distributed detection by reinforcement learning for unequiprobable sources," accepted by IEE Proceedings—Radar and Signal Processing.

J. G. Chen, N. Ansari, and Z. Siveski, "Improving multiuser detection performance by data fusion," GLOBECOM'96, London, England, Nov. 1996, pp 1794-1798.

J. G. Chen and N. Ansari, "Adaptive fusion of correlated local decisions," submitted to IEEE Trans. on System, Man and Cybernetics.

J. G. Chen and N. Ansari, "Performance comparison of fusion rules in distributed detection," in preparation.

J. G. Chen, N. Ansari, and Z. Siveski, "Adaptive data fusion for handoffs in cellular CDMA," submitted to IEEE Trans. on Vehicular Technology.

J. G. Chen, N. Ansari, and Z. Siveski, "Performance analysis and realization of decision fusion for macroscopic diversity in cellular wireless systems," submitted to Int. Conf. on Comm., 1997.

ACKNOWLEDGMENT

I would like to thank a number of people who have made significant contributions to the development of this dissertation. To begin with, I would like to thank my advisor Dr. Nirwan Ansari, who has provided me with a lot of knowledge and encouragement. Special thanks go to Dr. Zoran Siveski, who has helped me so much to understand CDMA. I would also like to thank Dr. Joseph Frank, Dr. Jack Holtzman, and Dr. Zoi-Heleni Michalopoulou for serving on my dissertation evaluation committee. I would also like to express my sincere gratitude and thanks to Lisa Fitton, former Administrative Assistant at the Center for Communication and Signal Processing Research, NJIT, for her editorial assistance and her jovial personality which made working at the lab a very enjoyable experience.

Several of my colleagues at the lab are deserving of recognition for their help at one point or another. I would like to thank past graduates Dr. Ambalavanar Arulambalam and Dr. Susan Wu for all the helpful discussions that we have had regarding various issues in the research and simulation. I also thank Dr. Gangsheng Wang and Dr. David Chen for helping me to familiarize with the working environment quickly when I first came to the lab.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

LIST OF FIGURES

Figure		Page

CHAPTER 1

INTRODUCTION

Sensor fusion, the study of optimal information processing in distributed multi-sensor environments through intelligent integration of multisensor data, has gained popularity in recent years [16],[48]. Such a technique is expected to increase the reliability of detection, to be fairly immune to noise interference and sensor failures, and to decrease the bandwidth requirement. The demand for sophisticated distributed detection systems has generated a great deal of interest in developing new algorithms to optimally fuse the information from different sensors. *Tenney* and *Sandell* [43] were among the first to study the problem of detection with distributed sensors. They applied the classical single sensor detection theory to a two-sensor, two-hypothesis test. An optimum local decision rule was established to minimize a global cost. Sadjadi [40] generalized the work of [43] to n detectors and m hypotheses, and obtained similar conclusions. Chair and Varshney [7] assumed that each local detector had a predetermined decision rule and each local decision was independent. With these assumptions, an optimum fusion model was generated by using the minimum probability of error criterion. Optimal techniques have also been developed for other criteria. When *a priori* probabilities were unknown, Thomopoulos [44] used the Neyman-Pearson (NP) test both at the local detector level and at the decision fusion level. An optimal decision scheme was derived. Demirbas [17] applied the maximum *a posteriori* (MAP) concept for object recognition in a multi-sensor environment and showed that the maximum *a posteriori* (MAP) estimation approach minimized mean square error estimation. Reibman and Nolte [37] found the global optimal solution by combining both the sensors and the fusion processor. The fusion of correlated decisions has also been studied [1, 28, 18].

Drakopoulos and Lee [18] have developed an optimum fusion rule for correlated decisions based on N-P criteria.

When the detection rule is fixed at each sensor, the optimal fusion rule developed based on Bayesian criterion for independent local decisions [7, 44, 41] is a weighed sum of local decisions. The weight associated with each local detector indicates the detector's degree of reliability. Each weight is a function of the probability of detection P_D and the probability of false alarm P_F of the detector. The P_D and P_F can be obtained when either the distribution of the observations at each detector is given, or when some reference signals are provided to estimate the P_D and P_F by an empirical method. However, in practice, neither P_D nor P_F is known. Furthermore, since the sensors are usually exposed to a changing environment, the performance of each individual detector may not always be the same, i.e, the P_D and P_F may vary with time. To circumvent this situation, we have developed an adaptive fusion system.

In chapter 2, an adaptive system to estimate the P_D and P_F was first proposed for equiprobable sources. Without knowledge of the performance of each detector, the proposed system is capable of approximately estimating the P_D and P_F of the detector in the course of performing the decision fusion by a simple counting process. In this adaptive fusion model, the fusion result is used as a supervisor to estimate the P_M and P_F. The fusion results are classified as "reliable" and "unreliable." Reliable results will be used as a reference to update the weights in the fusion center. Unreliable results will be discarded. The decision of a local detector is arbitrated by the fused decision of all the other local detectors. The convergence and error analysis of the system are demonstrated theoretically and by simulations. Analysis on classifying the fused decisions in term of reducing the estimation error is also given. The chapter concludes with simulation results which conform to the analysis.

In Chapter 3, this method is extended to unequiprobable sources, thus enhancing its practicality.

In Chapter 4, the adaptive method is extended further by considering the correlated local decisions.

Among all the methods proposed for sensor fusion in the independent case, there are three algorithms which are most representative. The first one, proposed by Tenney and Sandell [43], seeks to optimize decisions for local sensors. The second method seeks to optimize the fusion rule instead of local decisions [7]. The third algorithm considers both the local sensors and the fusion center [44, 37], which results in a set of complicated coupling equations. Because of its complexity, the analysis of the last method is often restricted in special cases and turns out to be equivalent to the first or second method. In practical binary detection, the fusion rules most often used are the logical "AND," (referred to as AND in this dissertation), "OR," (OR) and majority (MAJ) rule. These rules may be considered special cases of the K-out-of-N rule. The reason behind their popularity is their simplicity and that they do not need any *a priori* knowledge about signal sources and sensor properties. Another fusion rule that is often cited in the literature is the Bayesian-based optimal fusion rule (referred to as the OPT rule) [7]. Theoretically, the OPT rule requires knowledge of several *a priori* probabilities. We have proposed an adaptive algorithm to implement it without any *a priori* knowledge [4, 3, 11]. Thus, it is comparable with the above three practical fusion rules, in the sense of not requiring *a priori* knowledge. Among the four rules, the OPT rule is thought to be the best one. Thorough analysis, however, indicates that this is not always true. Although a very limited performance comparison of different fusion rules was done by Reibman and Nolte [38], and Tenney and Sandell [43] for the independent case, and Aalo and Viswanathan [1] for the correlated case, none of them has provided a thorough study. Many of the important analyses presented in Chapter 5, such as the effect of deviation of local decision probabilities and sensitivity to local threshold, are not covered in their papers. In Chapter 5, a thorough study on the performance of four fusion rules (AND, OR, MAJ and OPT), and performance advantages of one over another in terms of their ROCs are given. The performance in both independent and correlated Gaussian noise environments is considered. Various factors that affect the performance of fusion, such as the number of sensors, the Signal-to-Noise Ratio (SNR), the local decision threshold, the deviation of local decision probabilities, and correlation coefficients, are investigated. Several interesting and key observations can be concluded from the analysis.

Spatial diversity is used to combat fading and shadowing effects in wireless cellular communications [29]. Usually, microscopic spatial diversity is employed to reduce the fading effect by combining signals from different receiver elements of the same base station. Since much larger spatial separation is required to achieve shadowing decorrelation, macroscopic spatial diversity, which is implemented among different base station sites or ports, has been suggested to mitigate the shadowing effect [2, 5, 15, 27, 32, 35, 39, 49, 50]. Several possible combination rules have been proposed to achieve micro diversity [6], such as maximal ratio combining, equal gain combining, and selection diversity. In selection diversity, only the most reliable one is chosen among all the received signals, and all the others are simply ignored. Compared with other combination rules, selection diversity has poor performance, relatively low complexity and bandwidth requirement. Macroscopic diversity is, however, usually realized by selection diversity, because large separation of received signals increases the difficulty of bringing them together for better performance combination. In Chapter 6, we propose an optimal fusion scheme for macroscopic diversity combination based on the minimum error probability criterion for binary signals [13]. Fusion scheme is shown to have better performance than conventional macro selection diversity. When the error probability of the local detection in each base station is not available, the adaptive fusion algorithm proposed in previous chapters is adopted to estimate the combination weights. The performance analysis of the adaptive fusion algorithm in terms of minimum achievable error probability is presented [14]. A simplified realization of the optimal fusion scheme by using selection diversity, referred to as the "improved macro selection diversity rule," that has lower complexity and less bandwidth requirement than the direct realization is also proposed. The performance comparison of the proposed fusion scheme with the conventional macroscopic selection diversity in an environment in which both the Rayleigh fading and log-normal shadowing effects are considered. In Chapter 7, the fusion scheme is applied to handle the cellular CDMA handoff problem [12], where the system capacity is considered as a performance index. The performance of fusion handoff is compared with that of hard handoff and soft handoff.

Finally, concluding remarks are discussed in Chapter 8.

CHAPTER 2

ADAPTIVE FUSION FOR EQUAL PROBABLE SOURCE

In this chapter, we propose an adaptive system to estimate the P_D and P_F for equal probable source and independent local decision. Without knowledge of the performance of each detector, the proposed system is capable of approximately estimating the P_D and P_F of the detector in the course of performing the decision fusion.

2.1 Problem Statement

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

$$
\begin{aligned}
H_0 : &\quad \text{Signal is absent;} \\
H_1 : &\quad \text{Signal is present.}
\end{aligned}
$$

The *a priori* probabilities of the two hypotheses are denoted by $P(H_0) = P_0$ and $P(H_1) = P_1$. As shown in Figure 2.1, we assume that there are n detectors, and the observations at each detector are denoted by $x_i, i = 1, ..., n$. We further assume that the observations at the individual detectors are statistically independent and that the conditional probability is denoted by $P(x_i|H_j), i = 1, ..., n, j = 0, 1$. Each detector employs a decision rule $g_i(x_i)$ to make a decision $u_i, i = 1, ..., n$, where

$$
u_i = \begin{cases} -1 & \text{if } H_0 \text{ is declared,} \\ +1 & \text{if } H_1 \text{ is declared.} \end{cases}
$$

The probabilities of false alarm and missed detection for each detector are denoted by P_{F_i} and P_{M_i}, respectively.

After processing the observations locally, the decisions u_i are transmitted to the data fusion center. The data fusion center determines the overall decision u for the system based on the individual decisions, i.e.,

$$u = f(u_1, ..., u_n). \tag{2.1}$$

Based on the above specification, Chair and Varshney developed the optimum fusion rules as:

$$u = f(u_1, ..., u_n) = \begin{cases} +1, & \text{if} \quad w_0 + \sum_{i=1}^{n} w_i u_i > 0, \\ -1, & \text{otherwise.} \end{cases} \tag{2.2}$$

where

$$w_0 = \log \frac{P_1}{P_0}, \tag{2.3}$$

$$w_i = \begin{cases} \log \frac{1-P_{M_i}}{P_{F_i}}, & \text{if } u_i = +1, \\ \\ \log \frac{1-P_{F_i}}{P_{M_i}}, & \text{if } u_i = -1. \end{cases} \tag{2.4}$$

In case $P_0 = P_1$ and the probability of false alarm P_{Fj} is equal to the probability of miss P_{Mj}, $w_0 = 0$ and the optimal fusion rule can be simplified to

$$u = \begin{cases} +1, & \text{if } \sum_{j=1}^{n} w_j u_j > 0, \\ -1, & \text{otherwise,} \end{cases} \tag{2.5}$$

where

$$w_j = \log \frac{P_{Dj}}{P_{Fj}}, \quad \text{for each } j\ . \tag{2.6}$$

The system structure is shown in Fig. 2.1, where

$$y = \sum_{j=1}^{n} w_j u_j. \tag{2.7}$$

The structure shown in Fig. 2.1 is similar to a single neuron system, in particular, the Perceptron [33, 22, 51, 24, 25, 30]. If reference signals are given, they can be used as a "reference" to train the system such that the weights will converge to the optimal values defined by Eq. (2.6). However, in practice, such a reference is not readily available and at the same time, the P_D and P_F of a detector may vary with time. Since the fused decisions are usually better than local decisions, they

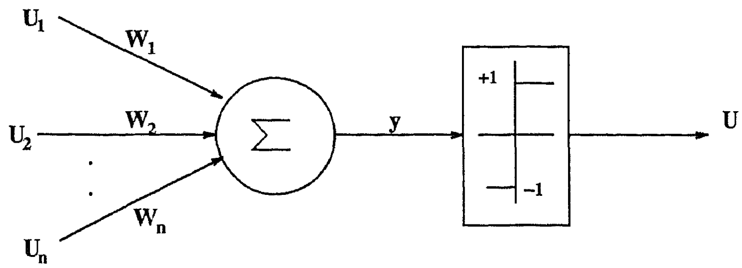

Figure 2.1 Structure of Fusion Center.

can be considered as the reference. When the ith local decision u_i is equal to the fused decision u, then u_i is considered to be correct; otherwise, u_i is considered to be incorrect. Since $u = \mathrm{sgn}(y) = \mathrm{sgn}(\sum_{j=1}^{n} w_j u_j)$, the fused decision u has already taken into account the decision of the ith detector, u_i. If u is used as a reference for u_i, a bias is established for u_i. Thus, in the proposed system, the decision of the ith local detector u_i is arbitrated by the fused decision of all the other $(n-1)$ local detectors. Denote the fused decision as $\bar{u}_i$, and define

$$y_i = \sum_{j \neq i} w_j u_j. \tag{2.8}$$

i.e, y_i is the weighed sum of all local decisions except u_i, then

$$\bar{u}_i = \mathrm{sgn}(y_i). \tag{2.9}$$

Note that $\bar{u}_i$ and u_i are conditionally independent given H_j, j =0,1. The "reference" $\bar{u}_i$ may not always be correct. To reduce the possibility of using incorrect references, the decisions $\bar{u}_i$ are further classified. The decision $\bar{u}_i$ is considered unreliable when the weighed sum defined by Eq. (2.8) is close to the decision threshold 0. Our strategy is to determine an "unreliable range" around the decision threshold such that when the weighed sum y_i falls in this range, the fused decision $\bar{u}_i$ is considered "unreliable" and will not be used for training the system. The selection of this "unreliable" range will be discussed next.

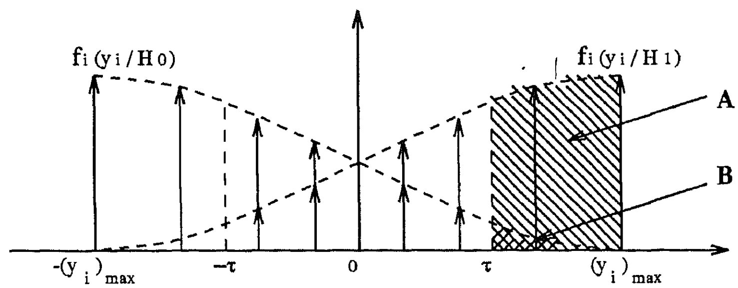

Figure 2.2 Conditional probability mass function: $f_i(y_i/H_1)$ and $f_i(-y_i/H_0)$.

2.2 The Adaptive Fusion Model Analysis

Consider the structure shown in Fig. 2.1. From Eqs. (2.7) and (2.8), we have

$$y_i = y - w_i u_i. \tag{2.10}$$

Under the assumptions that $P_0 = P_1$ and $P_{Fj} = P_{Mj}$, the conditional probability mass functions $f_i(y_i/H_1)$ and $f_i(y_i/H_0)$ are symmetric with each other, i.e., $f_i(y_i/H_1) = f_i(-y_i/H_0)$, as shown in Fig. 2.2.

We shall establish the above relationship as follows:

$$y_i = \sum_{j \neq i} w_j u_j = \sum_{S_i^+} w_j - \sum_{S_i^-} w_j, \tag{2.11}$$

where, $S_i^+ = \{j : j \neq i \text{ and } u_j = 1\}$, and $S_i^- = \{j : j \neq i \text{ and } u_j = -1\}$. By the earlier assumption of independent observations,

$$P(y_i = \xi/H_1) = \sum_{S_i} \prod_{S_i^+} P(u_j = 1/H_1) \prod_{S_i^-} P(u_j = -1/H_1), \tag{2.12}$$

where, $S_i = \{\{S_i^+, S_i^-\} :$ combinations of S_i^+ and S_i^- such that $\sum_{S_i^+} w_j - \sum_{S_i^-} w_j = \xi\}$. Since we have assumed that $P_{Mj} = P_{Fj}$, i.e.,

$$P(u_j = -1/H_1) = P(u_j = 1/H_0) = P_{Fj}, \tag{2.13}$$

which also implies that

$$P(u_j = 1/H_1) = P(u_j = -1/H_0) = P_{Dj}, \tag{2.14}$$

we have,

$$P(y_i = \xi/H_1) = \sum_{S_i} \prod_{S_i^+} P_{Dj} \prod_{S_i^-} P_{Fj}. \tag{2.15}$$

Now, assume that the local detectors, except the ith detector, make opposite decisions as compared to Eq. (2.12) such that y_i becomes $-\xi$. That is, S_i^+ and S_i^- remain the same, but the decisions are reversed. Thus,

$$\begin{aligned} P(y_i = -\xi/H_1) &= \sum_{S_i} \prod_{S_i^+} P(u_j = -1/H_1) \prod_{S_i^-} P(u_j = 1/H_1) = \sum_{S_i} \prod_{S_i^+} P_{Mj} \prod_{S_i^-} P_{Dj} \\ &= \sum_{S_i} \prod_{S_i^+} P_{Fj} \prod_{S_i^-} P_{Dj}. \end{aligned} \tag{2.16}$$

We shall next evaluate $P(y_i = \xi/H_0)$.

$$P(y_i = \xi/H_0) = \sum_{S_i} \prod_{S_i^+} P(u_j = 1/H_0) \prod_{S_i^-} P(u_j = -1/H_0) = \sum_{S_i} \prod_{S_i^+} P_{Fj} \prod_{S_i^-} P_{Dj}. \tag{2.17}$$

Since Eq. (2.16) and Eq. (2.17) are the same,

$$f_i(y_i/H_1) = f_i(-y_i/H_0). \tag{2.18}$$

Since $f_i(y_i/H_1)$ and $f_i(-y_i/H_0)$ have such a symmetric relation, let the *unreliable range* be symmetric about its decision threshold and denote the upper limit of the range as τ. We call τ the *reliability threshold.* Only the fused decisions $\bar{u}_i$ which satisfy $|y_i| > \tau$ are chosen to adapt the weight w_i. These decisions are considered as *reliable decisions*, denoted as $\bar{u}_i^*$. Other decisions are ignored. Intuitively, the bigger the value τ, the more reliable the decisions $\bar{u}_i^*$, the less the errors are between the estimates and theoretical values. Note that since u_i and $\bar{u}_i$ are conditionally independent and since i is deterministic, u_i and $\bar{u}_i^*$ are also conditionally independent.

Let $\hat{P}_{Di}, \hat{P}_{Fi}$ be the estimates of P_{Di}, P_{Fi}. When the local decision u_i agrees with the reliable decision $\bar{u}_i^*$, it is considered a detection of the local detector; otherwise, it is considered a false alarm. Using the conditional independence of u_i and $\bar{u}_i^*$, the assumption of an equiprobable source, and the definition of P_{Di} and P_{Fi},

$$\begin{aligned}
\hat{P}_{Di} &= P(u_i = 1, \bar{u}_i^* = 1) + P(u_i = -1, \bar{u}_i^* = -1) \\
&= P(H_0)P(u_i = 1, \bar{u}_i^* = 1/H_0) + P(H_1)P(u_i = 1, \bar{u}_i^* = 1/H_1) \\
&\quad + P(H_0)P(u_i = -1, \bar{u}_i^* = -1/H_0) + P(H_1)P(u_i = -1, \bar{u}_i^* = -1/H_1) \\
&= P(H_0)P(u_i = 1/H_0)P(\bar{u}_i^* = 1/H_0) \\
&\quad + P(H_1)P(u_i = 1/H_1)P(\bar{u}_i^* = 1/H_1) \\
&\quad + P(H_0)P(u_i = -1/H_0)P(\bar{u}_i^* = -1/H_0) \\
&\quad + P(H_1)P(u_i = -1/H_1)P(\bar{u}_i^* = -1/H_1) \\
&= \frac{1}{2}P_{Fi}P(\bar{u}_i^* = 1/H_0) + \frac{1}{2}P_{Di}P(\bar{u}_i^* = 1/H_1) \\
&\quad + \frac{1}{2}P_{Di}P(\bar{u}_i^* = -1/H_0) + \frac{1}{2}P_{Fi}P(\bar{u}_i^* = -1/H_1). \qquad (2.19)
\end{aligned}$$

Because of the symmetry of the conditional probability mass function (Eq. (2.18)),

$$P(\bar{u}_i^* = 1/H_0) = P(\bar{u}_i^* = -1/H_1), \qquad (2.20)$$

$$P(\bar{u}_i^* = 1/H_1) = P(\bar{u}_i^* = -1/H_0). \qquad (2.21)$$

Thus,

$$\hat{P}_{Di} = P_{Di}P(\bar{u}_i^* = 1/H_1) + P_{Fi}P(\bar{u}_i^* = 1/H_0). \qquad (2.22)$$

By the same reasoning, we have

$$\begin{aligned}
\hat{P}_{Fi} &= P(u_i = 1, \bar{u}_i^* = -1) + P(u_i = -1, \bar{u}_i^* = 1) \\
&= P_{Di}P(\bar{u}_i^* = 1/H_0) + P_{Fi}P(\bar{u}_i^* = 1/H_1). \qquad (2.23)
\end{aligned}$$

Let $\xi_1 < \xi_2 < \ldots < \xi_N$, where $\xi_N = (y_i)_{max}$, be the set of values that y_i can attain for the ith local detector. Without loss of generality, let $\xi_1 < \tau < \xi_N$, and $k \in \{1, 2, \ldots, N\}$ be the smallest integer such that $\xi_j > \tau$, $\forall j \geq k$. Define

$$A = P(\bar{u}_i^* = 1/H_1) = \sum_{j=k}^{N} P(y_i = \xi_j/H_1), \qquad (2.24)$$

$$B = P(\bar{u}_i^* = 1/H_0) = \sum_{j=k}^{N} P(y_i = \xi_j/H_0). \tag{2.25}$$

Then, Eq. (2.22) and Eq. (2.23) can be written as

$$\hat{P}_{Di} = P_{Di}A + P_{Fi}B, \tag{2.26}$$

and

$$\hat{P}_{Fi} = P_{Di}B + P_{Fi}A. \tag{2.27}$$

Let $r_i = \frac{P_{Di}}{P_{Fi}}, \hat{r}_i = \frac{\hat{P}_{Di}}{\hat{P}_{Fi}}$, then, $\log r_i = w_i$ is the weight of the ith detector defined by Eq. (2.6) and $\log \hat{r}_i = \hat{w}_i$ is the estimate for w_i.

$$\hat{r}_i = \frac{\hat{P}_{Di}}{\hat{P}_{Fi}} = \frac{P_{Di}A + P_{Fi}B}{P_{Di}B + P_{Fi}A} = r_i \frac{1 + \frac{B}{r_i A}}{1 + \frac{r_i B}{A}}. \tag{2.28}$$

$$\hat{w}_i = \log \hat{r}_i = \log r_i + \log \frac{1 + \frac{B}{r_i A}}{1 + \frac{r_i B}{A}} = w_i + \varepsilon_i. \tag{2.29}$$

As seen in Eq. (2.29), the estimate for the weight is equal to the correct weight plus an error term ε_i, where,

$$\varepsilon_i = \log \frac{1 + \frac{B}{r_i A}}{1 + \frac{r_i B}{A}}. \tag{2.30}$$

Since r_i is fixed, ε_i will approach 0 as $\frac{B}{A}$ is approaching 0. We will prove that increasing the *reliability threshold* τ will reduce the fraction $\frac{B}{A}$, and thus the error. For notational convenience, let $p_i = P_{Di}, q_i = P_{Fi}$. Since $P(y_i = \xi/H_0) = P(y_i = -\xi/H_1)$(Eq. (2.18)),

$$\frac{P(y_i = \xi/H_1)}{P(y_i = \xi/H_0)} = \frac{\sum_{S_i} \prod_{S_i^+} p_j \prod_{S_i^-} q_j}{\sum_{S_i} \prod_{S_i^+} q_j \prod_{S_i^-} p_j}. \tag{2.31}$$

From Eq. (2.11), we have

$$\exp(y_i) = \frac{\exp(\sum_{S_i^+} w_j)}{\exp(\sum_{S_i^-} w_j)} = \frac{\prod_{S_i^+} \exp(w_j)}{\prod_{S_i^-} \exp(w_j)}. \tag{2.32}$$

Applying Eq. (2.6) to Eq. (2.32) yields

$$\exp(y_i) = \prod_{S_i^+} \frac{p_j}{q_j} \prod_{S_i^-} \frac{q_j}{p_j} = \frac{\prod_{S_i^+} p_j \prod_{S_i^-} q_j}{\prod_{S_i^+} q_j \prod_{S_i^-} p_j}. \tag{2.33}$$

The above equation holds for any combination of S_i^+ and S_i^- such that

$$y_i = \sum_{j \in S_i^+} W_j - \sum_{j \in S_i^-} W_j = \xi. \tag{2.34}$$

Thus, using the following equality

$$\frac{a}{b} = \frac{c}{d} \Rightarrow \frac{a+c}{b+d} = \frac{a}{b}, \tag{2.35}$$

Eq. (2.31) becomes

$$\frac{P(y_i = \xi/H_1)}{P(y_i = \xi/H_0)} = \frac{\sum \prod_{S_i^+} p_j \prod_{S_i^-} q_j}{\sum \prod_{S_i^+} q_j \prod_{S_i^-} p_j} = \frac{\prod_{S_i^+} p_j \prod_{S_i^-} q_j}{\prod_{S_i^+} q_j \prod_{S_i^-} p_j} = \exp(\xi), \tag{2.36}$$

and

$$\frac{P(y_i = \xi/H_0)}{P(y_i = \xi/H_1)} = \exp(-\xi). \tag{2.37}$$

Thus far, we have proved that for each $y_i = \xi$, Eq. (2.37) holds. Using this equation and induction, we shall prove that $\frac{B}{A}$ is monotonically decreasing with respect to τ.

As assumed earlier, $\xi_1 < \xi_2 < \ldots < \xi_N$. From Eq. (2.37), we have

$$\frac{P(y_i = \xi_1/H_0)}{P(y_i = \xi_1/H_1)} > \frac{P(y_i = \xi_2/H_0)}{P(y_i = \xi_2/H_1)} > \ldots > \frac{P(y_i = \xi_{N-1}/H_0)}{P(y_i = \xi_{N-1}/H_1)} > \frac{P(y_i = \xi_N/H_0)}{P(y_i = \xi_N/H_1)}. \tag{2.38}$$

Repeatedly applying the inequality,

$$\frac{X}{Y} > \frac{a}{b}, \quad \Rightarrow \quad \frac{X}{Y} > \frac{X+a}{Y+b} > \frac{a}{b}, \tag{2.39}$$

to (2.38), and using the definition of A and B in Eqs. (2.24) and (2.25), it is clear that $\frac{B}{A}$ is monotonically decreasing with respect to k, and thus it is also monotonically decreasing with respect to τ. This is consistent with our intuitive reasoning. However, τ cannot go to infinity; the maximum value of τ is $(y_i)_{max}$. When τ attains its maximum, $\frac{B}{A}$ reaches its minimum value. According to the definition of A and B, the minimum of $\frac{B}{A}$ is

$$(\frac{B}{A})_{min} = \frac{P(y_i = (y_i)_{max}/H_0)}{P(y_i = (y_i)_{max}/H_1)}. \tag{2.40}$$

When P_{Di} is greater than P_{Fi} for each sensor and the learning procedure converges to its steady state, we have

$$(y_i)_{max} = \sum_{j=1, j\neq i}^{n} \log \frac{P_{Dj}}{P_{Fj}}. \tag{2.41}$$

Thus,

$$(\frac{B}{A})_{min} = \exp(-(y_i)_{max}) = \exp(-\sum_{j=1,j\neq i}^{n} \log \frac{P_{Dj}}{P_{Fj}}) = \prod_{j=1,j\neq i}^{n} \frac{P_{Fj}}{P_{Dj}}. \tag{2.42}$$

According to Eq. (2.30), the minimum error that can be achieved at steady state for a fixed i is:

$$\varepsilon_i = \log \frac{1 + \prod_{j=1}^{n} \frac{P_{Fj}}{P_{Dj}}}{1 + \frac{P_{Di}}{P_{Fi}} \prod_{j=1,j\neq i}^{n} \frac{P_{Fj}}{P_{Dj}}}. \tag{2.43}$$

Note that $(y_i)_{max}$ varies from sensor to sensor. In order to let every sensor adjust its weight and achieve the least error, the maximum value of τ is chosen to be the minimum of all $(y_i)_{max}$:

$$\tau_{max} = min\{(y_1)_{max}, (y_2)_{max}, \ldots\ldots(y_n)_{max}\}. \tag{2.44}$$

2.3 The Reinforcement Learning Algorithm

We assume that the distributed decision system has no knowledge of the probability mass functions of the observations. However, the probabilities of detection and false alarm for the ith detector $\hat{P}_{Di}$ and $\hat{P}_{Fi}$ can be approximated by relative frequencies. That is, in contrast to Eq. (2.19) and Eq. (2.23),

$$\frac{\hat{P}_{Di}}{\hat{P}_{Fi}} \approx \frac{m_i}{n_i}, \tag{2.45}$$

where m_i and n_i are, respectively, the number of decisions made by the ith detector that agree and disagree with the *reliable* fused decisions. Both m_i and n_i are simply obtained by counting in the simulations. We shall next develop the updating rule for the fusion center. Similarly,

$$\hat{w}_i = \log \frac{\hat{P}_{Di}}{\hat{P}_{Fi}} \approx \log \frac{m_i}{n_i}, \quad \Rightarrow \quad m_i \approx \exp(\hat{w}_i) n_i. \tag{2.46}$$

Taking the partial derivative with respect to m_i and n_i, respectively,

$$\frac{\partial \hat{w}_i}{\partial m_i} \approx \frac{1}{m_i}, \quad \text{and} \quad \frac{\partial \hat{w}_i}{\partial n_i} \approx -\frac{1}{n_i} = -\frac{1}{m_i}\exp(\hat{w}_i). \tag{2.47}$$

If the current local detector's decision agrees with the *reliable* fused decision, its weight $\hat{w}_i$ should be increased. In this case,

$$\Delta\hat{w}_i \approx \frac{1}{m_i}\Delta m_i = \frac{1}{m_i}. \tag{2.48}$$

On the other hand, if the current local decision disagrees with the *reliable* decision, its weight $\hat{w}_i$ should be reduced. That is,

$$\Delta\hat{w}_i \approx -\frac{1}{n_i}\Delta n_i = -\frac{1}{m_i}\exp(\hat{w}_i)\Delta n_i = -\frac{1}{m_i}\exp(\hat{w}_i). \tag{2.49}$$

Thus, we obtain the following updating rule:

$$\hat{w}_i^+ = \hat{w}_i^- + \Delta\hat{w}_i = \begin{cases} \hat{w}_i^- + \frac{1}{m_i}, & \text{if } u_i = \bar{u}_i^*, \\ \hat{w}_i^- - \frac{1}{m_i}\exp(\hat{w}_i^-), & \text{if } u_i \neq \bar{u}_i^*. \end{cases} \tag{2.50}$$

Lemma: Using the updating rule according to Eq. (2.50), $\hat{w}_i^-$ will converge to the desired steady state estimate weight $\hat{w}_i$.

Proof: Using the definition $E[X] = \sum x_i P(x_i)$ and the updating rule according to Eq. (2.50),

$$E[w_i^+ - w_i^-] = E[\Delta w] = \frac{1}{m_i}P(u_i = \bar{u}_i^*) - \frac{1}{m_i}\exp(\hat{w}_i^-)P(u_i \neq \bar{u}_i^*). \tag{2.51}$$

As the number of iterations increase, m_i will approach infinity. In this case,

$$\frac{1}{m_i}P(u_i = \bar{u}_i^*) - \frac{1}{m_i}\exp(\hat{w}_i^-)P(u_i \neq \bar{u}_i^*) = 0, \tag{2.52}$$

and from the definition of P_{Di} and P_{Fi},

$$\hat{P}_{Di} - \hat{P}_{Fi}\exp(\hat{w}_i^-) = 0. \tag{2.53}$$

Thus,

$$\hat{w}_i^- = \hat{w}_i. \tag{2.54}$$

Hence, $\hat{w}_i^- \to \hat{w}_i$, for i = 0,1, ... □

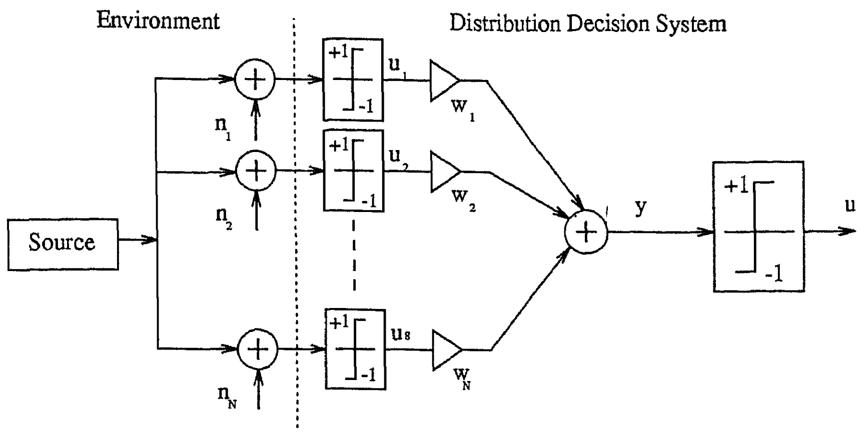

Figure 2.3 The structure of the distributed decision system.

2.4 Simulation Results

In this section, we present some computer simulation results to demonstrate the validity of our proposed adaptive scheme. Fig. 2.3 shows the simulation set-up. Here, equally likely binary signals $\{-1, 1\}$ are randomly generated as source signals. Additionally, $N_1, N_2, \ldots, N_n$ are assumed to be *i.i.d.* zero mean additive Gaussian random processes. Having selected the random noise process, the theoretical probabilities of detection and false alarm for each detector can be readily evaluated. For the Gaussian case, they can be determined by the standard deviation. They can be calculated according to:

$$P_{Fi} = Q(\frac{1}{\sigma_i}) = \int_{\frac{1}{\sigma_i}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}} d\xi, \tag{2.55}$$

where σ_i is the standard deviation of the Gaussian noise fed into the ith sensor.

$$P_{Di} = 1 - P_{Fi}. \tag{2.56}$$

Note that these theoretical probabilities and weights are calculated for comparison purposes only, and they are not readily available in practice. They are not used in the proposed adaptive fusion system. In the experiment, all the weights are first set

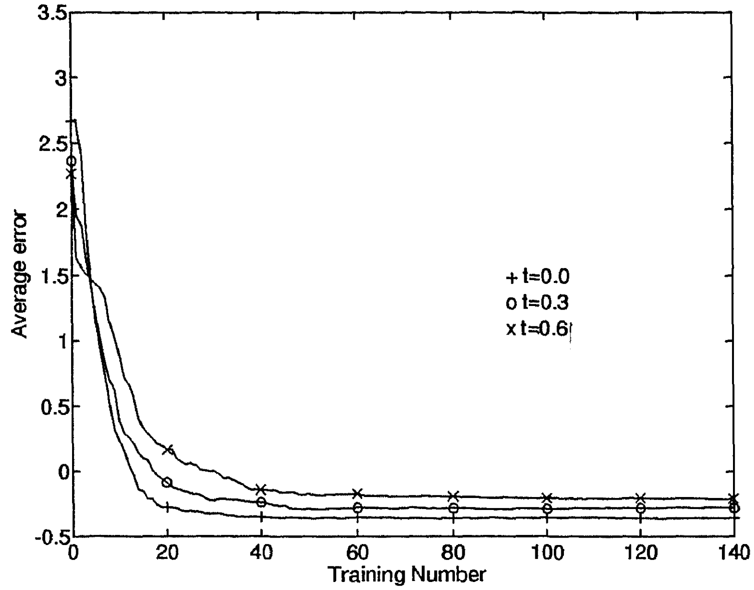

Figure 2.4 Simulation results for the case with identical detectors.

to an initial value of 1, and then updated according to Eq. (2.50). The steady state values are obtained after convergence ($\approx$ 400 iterations).

Figs. 2.4, 2.5, and Table 2.1 show the results for two different cases. The first case assumes that each local detector is identical. Here, $P_{Di} = 0.8413$, and $P_{Fi} = 0.1587$, for all $i = 1, 2, \ldots, 8$, where $w_i = \log\frac{P_{Di}}{P_{Fi}} = 1.6679$. Fig. 2.4 shows the mean error among 8 sensors between the estimate $\hat{w}_i$ and the actual weight $w_i = 1.6679$ for different values of τ, the *reliability threshold.* The figure conforms to our analytical results. That is, the larger the τ, the smaller the error. On the other hand, larger training time is needed to reach the steady state for a larger τ.

In the second case, the eight local detectors are assumed different, i.e., $P_{Di} = 0.9234$ and $P_{Fi} = 0.0766$, for i= 1,2,3,5,6,7; $P_{D4} = 0.8667$ and $P_{F4} = 0.1333$, and $P_{D8} = 0.9772$ and $P_{F8} = 0.0228$. Fig. 2.5 shows how the estimated weights approach the theoretical values. In the figure, $w_1 = 2.4895$, $w_4 = 1.8721$, $w_8 = 3.7579$. Only

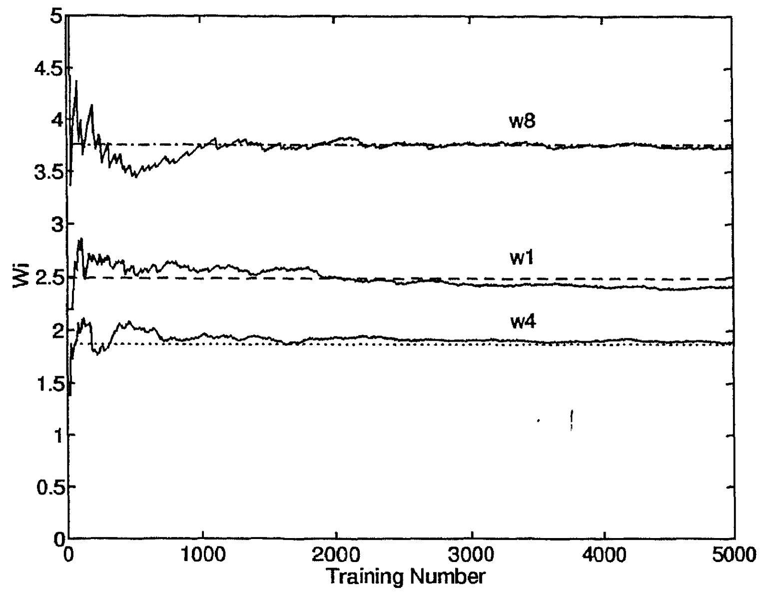

Figure 2.5 Simulation results for the case with different detectors. The straight lines represent the theoretical weight values, and the curves show the transient behavior of weight being updated.

Table 2.1 Comparison between the theoretical and steady state values of the weights.

Sensor		1st	2nd	3rd	4th
Noise Variance σ^2		0.49	0.49	0.49	0.81
SNR (dB)		3.098	3.098	3.098	0.915
Weights	Theoretical	2.4895	2.4895	2.4895	1.8721
	Steady state	2.4853	2.4916	2.4889	1.8721
Sensor		5th	6th	7th	8th
Noise Variance σ^2		0.49	0.49	0.49	0.25
SNR (dB)		3.098	3.098	3.098	6.02
Weights	Theoretical	2.4895	2.4895	2.4895	3.7579
	Steady state	2.4939	2.4971	2.4927	3.7941

three of the eight weights are shown. However, other weights also follow the same trend. Table 2.1 summaries the results for this experiment. It is readily seen that the simulation results conform closely to the theoretical results.

Though it has been shown that $\hat{w}_i^-$ converges to $\hat{w}_i$, it does not converge to w_i. The error, Eq. (2.43), depends on the number of sensors, the P_{Di} and P_{Fi}. In the Gaussian noise environment, P_{Di} and P_{Fi} are determined by the Signal-to-Noise ratio (SNR) of the ith sensor. Thus, the error ε_i is totally determined by the number and the SNRs of sensors. Fig. 2.6 shows, for case of identical sensors, the error, ε_i, versus n (the number of sensors) for various SNRs. In this case, according to Eqs. (2.55) and (2.56), the error can be simplified to:

$$\varepsilon_i = \log \frac{1 + (\frac{Q}{1-Q})^n}{1 + (\frac{Q}{1-Q})^{n-2}}, \tag{2.57}$$

where Q is the Q-function defined in Eq. (2.55) with the same standard deviation, σ, for all sensors. Note that the error is the same for every sensor.

When the SNR is different from sensor to sensor, the error can be written as

$$\varepsilon_i = \log \frac{1 + T}{1 + (\frac{1-Q_i}{Q_i})^2 T}, \tag{2.58}$$

where

$$T = \prod_{j=1}^{n} \frac{Q_j}{1 - Q_j}, \tag{2.59}$$

and

$$Q_j = Q(\frac{1}{\sigma_j}). \tag{2.60}$$

From Eq. (2.58), it is readily seen that the larger the SNR, the smaller the ε_i.

2.5 Summary

In a real-world environment, the probability mass functions of the observations at local detectors may not be known and the performance of the local detectors may

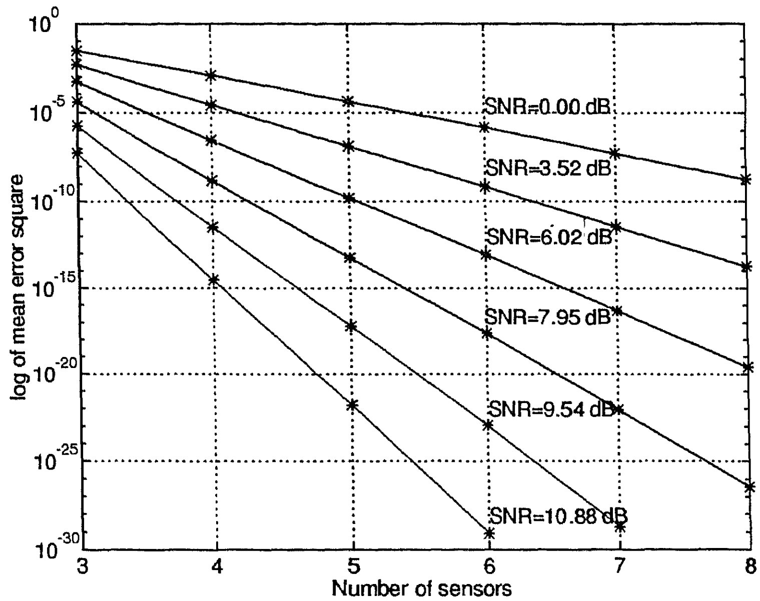

Figure 2.6 Error curve versus the number of sensors.

not be stationary. Under such circumstances, it is desirable to have a system which can adapt itself during the decision making process. This chapter proposes such an adaptive system with the assumption that $P_0 = P_1$ and $P_{Di} = P_{Fi}$. The major advantage of the system is that *a priori* knowledge of the probability mass functions of the observations is not required. The system can acquire the knowledge about the reliability of the local detectors by itself — *it can learn by doing.* A reinforcement learning rule is proposed and adopted, and its convergence is analytically proven. The simulation results conform to our theoretical analysis. If the *reliability threshold* τ can be adjusted adaptively during the process of data fusing, the system may converge faster. Future efforts will focus on adaptively adjusting the *reliability threshold*, and developing a model for unequiprobable sources.

CHAPTER 3

ADAPTIVE DECISION FUSION FOR UNEQUIPROBABLE SOURCES

3.1 Date Fusion Rule And Its Properties

The situation here is exactly the same as in Chapter 2 except the *priori* probabilities and conditional probabilities are different, *i.e.* $P_0 \neq P_1$, and $P_{M_i} \neq P_{F_i}$. Thus, the optimum fusion rule is:

$$u = f(u_1, ..., u_n) = \begin{cases} +1, & \text{if} \quad w_0 + \sum_{i=1}^{n} w_i u_i > 0, \\ -1, & \text{otherwise.} \end{cases} \tag{3.1}$$

where

$$w_0 = \log \frac{P_1}{P_0}, \tag{3.2}$$

$$w_i = \begin{cases} \log \frac{1-P_{M_i}}{P_{F_i}}, & \text{if } u_i = +1, \\ \\ \log \frac{1-P_{F_i}}{P_{M_i}}, & \text{if } u_i = -1. \end{cases} \tag{3.3}$$

The optimum data fusion rule can be implemented as shown in Figure 3.1, where

$$y = w_0 + \sum_{i=1}^{n} w_i u_i.$$

Note that the above fusion rule has the following property.

Lemma I: When each weight in the fusion center is optimum, the conditional probability mass functions $P(y - w_0 = \zeta | H_1)$ and $P(y - w_0 = \zeta | H_0)$ satisfy the

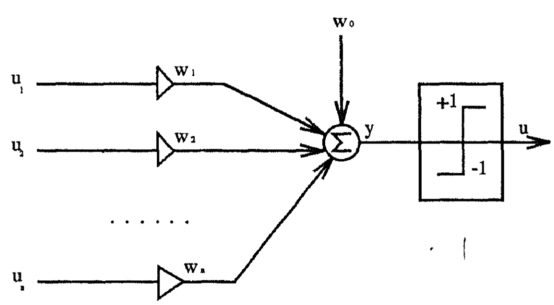

Figure 3.1 The fusion center structure.

following equation:

$$e^{\zeta} = \frac{P(y - w_0 = \zeta | H_1)}{P(y - w_0 = \zeta | H_0)}, \tag{3.4}$$

where ζ is a possible value of $y - w_0$.

Proof Consider the structure shown in Figure 3.1. We have:

$$y = w_0 + \sum_{j=1}^{n} w_j u_j, \tag{3.5}$$

$$\text{or} \quad y = w_0 + \sum_{j \in S^+} w_j - \sum_{j \in S^-} w_j, \tag{3.6}$$

where $S^+ = \{j : u_j = 1\}$, and $S^- = \{j : u_j = -1\}$. From Eq.(3.3) and Eq.(3.6):

$$\begin{aligned} y &= w_0 + \sum_{j \in S^+} \log \frac{P(u_j = 1 | H_1)}{P(u_j = 1 | H_0)} - \sum_{j \in S^-} \log \frac{P(u_j = -1 | H_0)}{P(u_j = -1 | H_1)} \\ &= w_0 + \log \left[\prod_{j \in S^+} \frac{P(u_j = 1 | H_1)}{P(u_j = 1 | H_0)} \middle/ \prod_{j \in S^-} \frac{P(u_j = -1 | H_0)}{P(u_j = -1 | H_1)} \right]. \end{aligned} \tag{3.7}$$

$$\exp(y - w_0) = \frac{\prod_{j \in S^+} P(u_j = 1 | H_1) \prod_{j \in S^-} P(u_j = -1 | H_1)}{\prod_{j \in S^+} P(u_j = 1 | H_0) \prod_{j \in S^-} P(u_j = -1 | H_0)}. \tag{3.8}$$

Let ζ be a possible value of $y - w_0$, and each local decision u_j is independent:

$$P(y - w_0 = \zeta | H_1) = \sum_{\mathbf{u} \in U} P(\mathbf{w}^T \mathbf{u} = \zeta | H1),$$

where $\mathbf{u}$ is a vector with elements $u_i, i = 1, 2, \cdots n$, $\mathbf{w}$ is a vector with elements $w_i, i = 1, 2, \cdots n$, and

$$U = \{\mathbf{u} : \mathbf{w}^T \mathbf{u} = \zeta\}.$$

By defining S as

$$\{\{S^+, S^-\}: \text{a combination of } S^+ \text{ and } S^- \text{ such that } \sum_{j \in S^+} w_j - \sum_{j \in S^-} w_j = \zeta\},$$

$$P(y - w_0 = \zeta | H_1) = \sum_{S} \prod_{j \in S^+} P(u_j = 1 | H_1) \prod_{j \in S^-} P(u_j = -1 | H_1),$$

and

$$P(y - w_0 = \zeta | H_0) = \sum_S \prod_{j \in S^+} P(u_j = 1 | H_0) \prod_{j \in S^-} P(u_j = -1 | H_0).$$

Thus,

$$\frac{P(y - w_0 = \zeta | H_1)}{P(y - w_0 = \zeta | H_0)} = \frac{\sum\limits_S \prod\limits_{j \in S^+} P(u_j = 1 | H_1) \prod\limits_{j \in S^-} P(u_j = -1 | H_1)}{\sum\limits_S \prod\limits_{j \in S^+} P(u_j = 1 | H_0) \prod\limits_{j \in S^-} P(u_j = -1 | H_0)}. \tag{3.9}$$

From Eq.(3.8) and the following equality:

$$\frac{a}{b} = \frac{c}{d} = k \quad \Rightarrow \quad \frac{a+c}{b+d} = \frac{bk + dk}{b+d} = k, \qquad (b \neq 0, d \neq 0, b + d \neq 0),$$

then

$$\begin{aligned} \frac{P(y - w_0 = \zeta | H_1)}{P(y - w_0 = \zeta | H_0)} &= \frac{\sum\limits_S \prod\limits_{j \in S^+} P(u_j = 1 | H_1) \prod\limits_{j \in S^-} P(u_j = -1 | H_1)}{\sum\limits_S \prod\limits_{j \in S^+} P(u_j = 1 | H_0) \prod\limits_{j \in S^-} P(u_j = -1 | H_0)} \\ &= \frac{\prod\limits_{j \in S^+} P(u_j = 1 | H_1) \prod\limits_{j \in S^-} P(u_j = -1 | H_1)}{\prod\limits_{j \in S^+} P(u_j = 1 | H_0) \prod\limits_{j \in S^+} P(u_j = -1 | H_0)} \end{aligned}$$

$$\frac{P(y - w_0 = \zeta | H_1)}{P(y - w_0 = \zeta | H_0)} = e^{y - w_0} = e^{\zeta}. \tag{3.10}$$

Q. E. D.

Eq.(3.4) is a very interesting result. The ratio of the conditional probabilities under H_1 and H_0 only depends on the value $y - w_0$, even the probability mass functions $P(y - w_0 = \zeta | H_1)$ and $P(y - w_0 = \zeta | H_1)$ may not be monotonic with ζ. This is illustrated in Figure 3.2.

Recall the data fusion center structure shown in Figure 3.1. If the reference signals are given, they can be used as a "reference" to train the system such that weights will converge to the optimal values defined by Eqs.(3.2) and (3.3). However, in practice such a reference is not readily available and at the same time, the P_M and P_F of a detector may vary with time. Since the fused decisions are usually

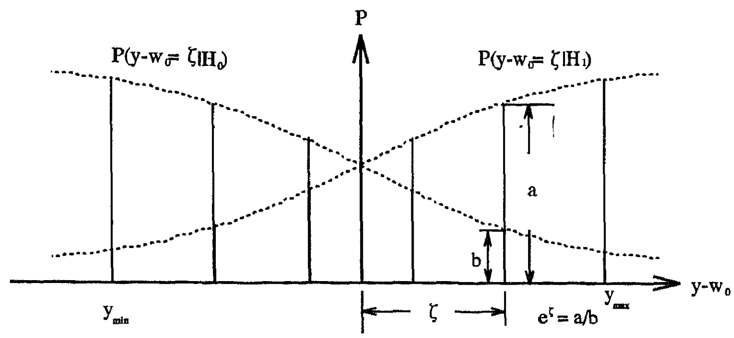

Figure 3.2 Relationship between $P(y - w_0 = \zeta|H_1)$ and $P(y - w_0 = \zeta|H_0)$.

better than local decisions, they can be considered as the reference. When the ith local decision u_i is equal to the fused decision u, then u_i is considered to be correct; otherwise, u_i is considered to be incorrect. Since $y = w_0 + \sum_{i=1}^{n} w_i u_i$, the fused decision u has already taken into account the decision of the ith detector, u_i. If u is used as a reference for u_i, a bias is established for u_i. Thus, in the proposed system, the decision of the ith local detector u_i is arbitrated by the fused decision of all the other $(n-1)$ local detectors. Denote this fused decision as $\overline{u}_i$, and define

$$y_i = y - w_0 - w_i u_i = \sum_{j=1, j\neq i}^{n} w_j u_j. \tag{3.11}$$

The decision $\overline{u}_i$ in the fusion center for updating $\widehat{w}_i$ depends on the value y_i. Here $\hat{w}_i$ is the estimated weight. Using the same procedure, it can be shown that y_i has a similar property to y in Lemma I. That is:

$$\frac{P(y_i = \zeta|H_1)}{P(y_i = \zeta|H_0)} = e^{\zeta}, \tag{3.12}$$

where ζ is a possible value that y_i takes on. The range of y_i is divided into reliable and unreliable ranges. We denote the lower and upper limit of the unreliable range as τ_1 and τ_2, as shown in Figure 3.3. Usually $\tau_2 \geq 0$, $\tau_1 \leq 0$. We call τ_1 and τ_2 the *reliability thresholds*. Only the fused decisions $\overline{u}_i$ which satisfy $y_i < \tau_1$ or $y_i > \tau_2$ are chosen to adapt each weight, denoted by $\widehat{w}_i$. These decisions are considered

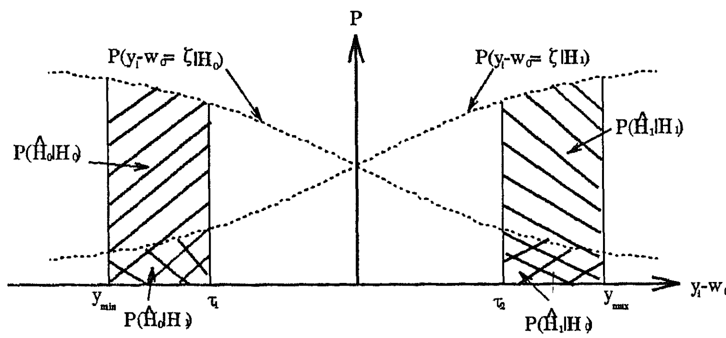

Figure 3.3 Classification of fusion results.

reliable decisions, defined by $\widehat{H_1}$ when $y_i > \tau_2$, and $\widehat{H_0}$ when $y_i < \tau_1$. The decision is considered *unreliable* when $\tau_1 < y_i < \tau_2$, denoted by $\widehat{H_x}$. Obviously, we have

$$P(\widehat{H}_1) + P(\widehat{H}_0) + P(\widehat{H}_x) = 1. \tag{3.13}$$

This type of learning belongs to the class of reinforcement learning [33].

Based on the proposed fusion rule described in eq.(3.11), we obtain the following two properties related to the steady state error.

Lemma II: If $\alpha = \frac{P(\widehat{H}_1|H_0)}{P(\widehat{H}_1|H_1)}$, $\beta = \frac{P(\widehat{H}_0|H_1)}{P(\widehat{H}_0|H_0)}$, and $\gamma = \frac{P(\widehat{H}_1|H_1)}{P(\widehat{H}_0|H_0)}$, we have following conclusions.

(1) α is monotonically decreasing when τ_2 increases, β is monotonically decreasing when τ_1 decreases, and

$$\alpha_{min} = \prod_{j=1,j\neq i}^{n} \frac{P_{Fj}}{1 - P_{Mj}} \tag{3.14}$$

$$\beta_{min} = \prod_{j=1,j\neq i}^{n} \frac{P_{Mj}}{1 - P_{Fj}} \tag{3.15}$$

(2) when $\tau_2 = (y_i)_{max}$ and $\tau_1 = (y_i)_{min}$,

$$\gamma = \frac{P(\widehat{H_1}|H_1)}{P(\widehat{H_0}|H_0)} = \prod_{j=1,j\neq i}^{n} (\frac{1 - P_{Mj}}{1 - P_{Fj}}). \tag{3.16}$$

Proof:

$$\alpha = \frac{P(\widehat{H}_1|H_0)}{P(\widehat{H}_1|H_1)}$$

$$= \frac{P(y_i > \tau_2 | H_0)}{P(y_i > \tau_2 | H_1)}$$
$$= \frac{\sum_{j=1}^{m} P(y_i = \zeta_j | H_0)}{\sum_{j=1}^{m} P(y_i = \zeta_j | H_1)}.$$

Without loss of generality, assume that $\zeta_1 > \zeta_2 > \cdots > \zeta_m > \tau_2$ are all possible y_j.

Note that as τ_2 becomes larger, m becomes smaller.

From Eq.(3.12):

$$e^{-\zeta} = \frac{P(y_i = \zeta | H_0)}{P(y_i = \zeta | H_1)}, \tag{3.17}$$

we have,

$$\frac{P(y_i = \zeta_1 | H_0)}{P(y_i = \zeta_1 | H_1)} < \frac{P(y_i = \zeta_2 | H_0)}{P(y_i = \zeta_2 | H_1)} < \cdots < \frac{P(y_i = \zeta_m | H_0)}{P(y_i = \zeta_m | H_1)}. \tag{3.18}$$

Denote $A_k = \sum_{j=1}^{k} P(y_i = \zeta_j | H_1)$, $B_k = \sum_{j=1}^{k} P(y_i = \zeta_j | H_0)$, and $\alpha_k = \frac{A_k}{B_k}$. The objective is to show that $\alpha_k > \alpha_{k-1}$ for $k = 1, 2, \cdots, n$. First we need to show $\alpha_2 > \alpha_1$.

$$\alpha_1 = \frac{P(y_i = \zeta_1 | H_0)}{P(y_i = \zeta_1 | H_1)}$$
$$\alpha_2 = \frac{P(y_i = \zeta_1 | H_0) + P(y_i = \zeta_2 | H_0)}{P(y_i = \zeta_1 | H_1) + P(y_i = \zeta_2 | H_1)}.$$

Using the following inequality and Eq.(3.18),

$$\frac{X}{Y} < \frac{a}{b} \quad \Rightarrow \quad \frac{X}{Y} < \frac{X+a}{Y+b} < \frac{a}{b} \quad (Y, b > 0), \tag{3.19}$$

we have

$$\alpha_2 > \alpha_1. \tag{3.20}$$

Next, we shall show that if $\alpha_k > \alpha_{k-1}$, then $\alpha_{k+1} > \alpha_k$.

Since

$$\alpha_{k-1} < \alpha_k \tag{3.21}$$

$$\frac{A_{k-1}}{B_{k-1}} < \frac{A_k}{B_k}$$

$$\Rightarrow \frac{A_{k-1}}{B_{k-1}} < \frac{A_{k-1} + P(y_i = \zeta_k | H_0)}{B_{k-1} + P(y_i = \zeta_k | H_1)}. \tag{3.22}$$

Using inequality Eq.(3.19) again,

$$\frac{A_{k-1}}{B_{k-1}} < \frac{A_{k-1} + P(y_i = \zeta_k | H_0)}{B_{k-1} + P(y_i = \zeta_k | H_1)} < \frac{P(y_i = \zeta_k | H_0)}{P(y_i = \zeta_k | H_1)}. \tag{3.23}$$

Applying Eq.(3.18) and Eq.(3.22) yields:

$$\frac{A_{k-1} + P(y_i = \zeta_k | H_0)}{B_{k-1} + P(y_i = \zeta_k | H_1)} < \frac{P(y_i = \zeta_k | H_0)}{P(y_i = \zeta_k | H_1)} < \frac{P(y_i = \zeta_{k+1} | H_0)}{P(y_i = \zeta_{k+1} | H_1)}.$$

Using Eq.(3.19):

$$\frac{A_{k-1} + P(y_i = \zeta_k | H_0)}{B_{k-1} + P(y_i = \zeta_k | H_1)} < \frac{A_{k-1} + P(y_i = \zeta_k | H_0) + P(y_i = \zeta_{k+1} | H_0)}{B_{k-1} + P(y_i = \zeta_k | H_1) + P(y_i = \zeta_{k+1} | H_1)}$$

$$\Rightarrow \frac{A_k}{B_k} < \frac{A_{k+1}}{B_{k+1}} \Rightarrow \alpha_k < \alpha_{k+1}. \tag{3.24}$$

From Eq.(3.20), Eq.(3.23) and Eq.(3.24), α decreases monotonically with τ_2.

However, τ_2 cannot go to infinity; the maximum value of τ_2 is $(y_i)_{max}$. When τ_2 attains its maximum, α reaches its minimum value. According to the definition of α, the minimum of α is

$$\alpha_{min} = \frac{P(y_i = (y_i)_{max} / H_0)}{P(y_i = (y_i)_{max} / H_1)} = \exp(-(y_i)_{max}). \tag{3.25}$$

When P_{Di} (P_{Di} is the probability of detection) is greater than P_{Fi} for each sensor (which is the usual case) and the learning procedure converges to its steady state, we have

$$(y_i)_{max} = \sum_{j=1, j \neq i}^{n} \log \frac{1 - P_{Mj}}{P_{Fj}}. \tag{3.26}$$

Thus,

$$\alpha_{min} = \exp(-(y_i)_{max}) = \exp(-\sum_{j=1, j \neq i}^{n} \log \frac{1 - P_{Mj}}{P_{Fj}}) = \prod_{j=1, j \neq i}^{n} \frac{P_{Fj}}{1 - P_{Mj}}. \tag{3.27}$$

By the same reasoning, we can prove that β is decreasing when τ_1 decreases, and

$$\beta_{min} = \frac{P(y_i = (y_i)_{min}|H_1)}{P(y_i = (y_i)_{min}|H_0)} = \exp((y_i)_{min}). \tag{3.28}$$

From Eq. (3.3), we know

$$(y_i)_{min} = -\sum_{j=1,j\neq i}^{n} \log\frac{1-P_{Fj}}{P_{Mj}}. \tag{3.29}$$

Thus,

$$\beta_{min} = \prod_{j=1,j\neq i}^{n} \frac{P_{Mj}}{1-P_{Fj}}. \tag{3.30}$$

When $\tau_2 = (y_i)_{max}$ and $\tau_1 = (y_i)_{min}$,

$$\begin{aligned} P(y_i \geq \tau_2|H_1) &= P(u_j = 1|H_1, \text{for all } j \text{ except } i) \\ &= \prod_{j=1,j\neq i}^{n} P(u_j = 1|H_1) = \prod_{j=1,j\neq i}^{n} (1 - P_{Mi}). \end{aligned} \tag{3.31}$$

$$\begin{aligned} P(y_i \leq \tau_1|H_0) &= P(u_j = -1|H_0, \text{ for all } j \text{ except } i) \\ &= \prod_{j=1,j\neq i}^{n} P(u_j = -1|H_0) = \prod_{j=1,j\neq i}^{n} (1 - P_{Fi}). \end{aligned} \tag{3.32}$$

Thus,

$$\gamma = \frac{P(\widehat{H_1}|H_1)}{P(\widehat{H_0}|H_0)} = \prod_{j=1,j\neq i}^{n} (\frac{1-P_{Mi}}{1-P_{Fi}}). \tag{3.33}$$

Q. E. D.

Lemma III: Let $\epsilon_i = |w_i - \widehat{w}_i|$, $i = 0, 1, \cdots n$, represent the estimation error. The minimum ϵ_i that can be achieved is:

$$\begin{cases} \epsilon_0 = \log(\frac{1+\frac{\alpha_{min}}{r_0}}{1+r_0\beta_{min}}) + \log(\gamma) \\ \epsilon_i = \begin{cases} \log\frac{1+\beta_{min}r_0}{1+\frac{\alpha_{min}}{r_0}} + \log\frac{1+\frac{\alpha_{min}}{r_0 r_i}}{1+\beta_{min}r_0 r_i} & \text{if} \quad u_i = +1 \\ \log\frac{1+\frac{\alpha_{min}}{r_0}}{1+\beta_{min}r_0} + \log\frac{1+\frac{\beta_{min}r_0}{r_i}}{1+r_i\frac{\alpha_{min}}{r_0}} & \text{if} \quad u_i = -1. \end{cases} \end{cases} \tag{3.34}$$

Proof: If $w_i = \log r_i$ and $\widehat{w}_i = \log \widehat{r}_i$, we have,

$$\begin{cases} r_0 &= \frac{P_1}{P_0} \\ r_i &= \begin{cases} \frac{P(u_i=1|H_1)}{P(u_i=1|H_0)} & \text{if} \quad u_i = +1 \\ \frac{P(u_i=-1|H_0)}{P(u_i=-1|H_1)} & \text{if} \quad u_i = -1, \end{cases} \end{cases} \tag{3.35}$$

and

$$\begin{cases} \widehat{r}_0 &= \frac{P(\widehat{H}_1)}{P(\widehat{H}_0)} \\ \widehat{r}_i &= \begin{cases} \frac{P(u_i=1|\widehat{H}_1)}{P(u_i=1|\widehat{H}_0)} & \text{if} \quad u_i = +1 \\ \frac{P(u_i=-1|\widehat{H}_0)}{P(u_i=-1|\widehat{H}_1)} & \text{if} \quad u_i = -1. \end{cases} \end{cases} \tag{3.36}$$

Using the *total probability theorem* $P(BA) = P(B|A)P(A)$:

$$P(u_i = 1|\widehat{H}_1) = \frac{P(u_i = 1, \widehat{H}_1)}{P(\widehat{H}_1)}$$
$$= \frac{P(u_i = 1, \widehat{H}_1|H_0)P(H_0) + P(u_i = 1, \widehat{H}_1|H_1)P(H_1)}{P(\widehat{H}_1)}$$
$$= \frac{P(u_i = 1|H_0)P(\widehat{H}_1|u_i = 1, H_0)P(H_0) + P(u_i = 1|H_1)P(\widehat{H}_1|u_i = 1, H_1)P(H_1)}{P(\widehat{H}_1)}.$$

Similarly,

$$P(u_i = 1|\widehat{H}_0) =$$
$$\frac{P(u_i = 1|H_0)P(\widehat{H}_0|u_i = 1, H_0)P(H_0) + P(u_i = 1|H_1)P(\widehat{H}_0|u_i = 1, H_1)P(H_1)}{P(\widehat{H}_0)},$$
$$P(u_i = -1|\widehat{H}_0) =$$
$$\frac{P(u_i = -1|H_0)P(\widehat{H}_0|u_i = -1, H_0)P(H_0) + P(u_i = -1|H_1)P(\widehat{H}_0|u_i = -1, H_1)P(H_1)}{P(\widehat{H}_0)},$$
$$P(u_i = -1|\widehat{H}_1) =$$
$$\frac{P(u_i = -1|H_0)P(\widehat{H}_1|u_i = -1, H_0)P(H_0) + P(u_i = -1|H_1)P(\widehat{H}_1|u_i = -1, H_1)P(H_1)}{P(\widehat{H}_1)}.$$

Using Eq.(3.36) and the above formulas, if $u_i = +1$,

$$\widehat{r}_i = \frac{\begin{array}{l}[P(u_i=1|H_0)P(\widehat{H}_1|u_i=1,H_0)P(H_0)+ \\ \qquad +P(u_i=1|H_1)P(\widehat{H}_1|u_i=1,H_1)P(H_1)]P(\widehat{H}_0)\end{array}}{\begin{array}{l}[P(u_i=1|H_0)P(\widehat{H}_0|u_i=1,H_0)P(H_0)+ \\ \qquad +P(u_i=1|H_1)P(\widehat{H}_0|u_i=1,H_1)P(H_1)]P(\widehat{H}_1)\end{array}}$$

$$= \frac{P(u_i=1|H_1)P(\widehat{H}_1|H_1)P(H_1)P(\widehat{H}_0)}{P(u_i=1|H_0)P(\widehat{H}_0|H_0)P(H_0)P(\widehat{H}_1)} \cdot \frac{\frac{P(u_i=1|H_0)P(\widehat{H}_1|H_0)P(H_0)}{P(u_i=1|H_1)P(\widehat{H}_1|H_1)P(H_1)}+1}{1+\frac{P(u_i=1|H_1)}{P(u_i=1|H_0)}\frac{P(\widehat{H}_0|H_1)}{P(\widehat{H}_0|H_0)}\frac{P(H_1)}{P(H_0)}}$$

$$= r_i \cdot \frac{P(\widehat{H}_1|H_1)}{P(\widehat{H}_0|H_0)} \cdot \frac{P(H_1)}{P(H_0)} \cdot \frac{P(\widehat{H}_0)}{P(\widehat{H}_1)} \cdot \frac{\frac{1}{r_i}\frac{P(\widehat{H}_1|H_0)}{P(\widehat{H}_1|H_1)}\frac{1}{r_0}+1}{1+r_i\frac{P(\widehat{H}_0|H_1)}{P(\widehat{H}_0|H_0)}r_0}. \tag{3.37}$$

Here,

$$\frac{P(\widehat{H}_0)}{P(\widehat{H}_1)} = \frac{P(\widehat{H}_0|H_1)P(H_1)+P(\widehat{H}_0|H_0)P(H_0)}{P(\widehat{H}_1|H_1)P(H_1)+P(\widehat{H}_1|H_0)P(H_0)}.$$

Thus,

$$\begin{aligned}&\frac{P(\widehat{H}_1|H_1)}{P(\widehat{H}_0|H_0)} \cdot \frac{P(H_1)}{P(H_0)} \cdot \frac{P(\widehat{H}_0)}{P(\widehat{H}_1)} \\ &= \frac{P(H_1)}{P(H_0)} \cdot \frac{\frac{P(\widehat{H}_0|H_1)}{P(\widehat{H}_0|H_0)}\cdot P(H_1)+P(H_0)}{P(H_1)+\frac{P(\widehat{H}_1|H_0)}{P(\widehat{H}_1|H_1)}\cdot P(H_0)} = \frac{\frac{P(\widehat{H}_0|H_1)}{P(\widehat{H}_0|H_0)}\cdot\frac{P(H_1)}{P(H_0)}+1}{1+\frac{P(\widehat{H}_1|H_0)}{P(\widehat{H}_1|H_1)}\cdot\frac{P(H_0)}{P(H_1)}}.\end{aligned} \tag{3.38}$$

Substituting α, β and r_0 into Eq.(3.37), after simplification, we get

$$\widehat{r}_i = r_i \cdot \frac{1+\beta r_0}{1+\frac{\alpha}{r_0}} \cdot \frac{1+\frac{\alpha}{r_0 r_i}}{1+\beta r_0 r_i}. \tag{3.39}$$

Similarly, if $u_i = -1$,

$$\begin{aligned}\widehat{r}_i &= \frac{P(\widehat{H}_1)P(H_0)P(\widehat{H}_0|H_0)P(u_i=-1|H_0)}{P(\widehat{H}_0)P(H_1)P(\widehat{H}_1|H_1)P(u_i=-1|H_1)} \cdot \frac{1+\frac{P(u_i=-1|H_1)P(\widehat{H}_0|H_1)P(H_1)}{P(u_i=-1|H_0)P(\widehat{H}_0|H_0)P(H_0)}}{\frac{P(u_i=-1|H_0)P(\widehat{H}_1|H_0)P(H_0)}{P(u_i=-1|H_1)P(\widehat{H}_1|H_1)P(H_1)}+1} \\ &= r_i \cdot \frac{1+\frac{\alpha}{r_0}}{1+\beta r_0} \cdot \frac{1+\frac{\beta r_0}{r_i}}{1+r_i\frac{\alpha}{r_0}}.\end{aligned} \tag{3.40}$$

From Eq.(3.36),

$$\widehat{r}_0 = \frac{P(\widehat{H}_1)}{P(\widehat{H}_0)} = \frac{P(\widehat{H}_1|H_1)P_1+P(\widehat{H}_1|H_0)P_0}{P(\widehat{H}_0|H_0)P_0+P(\widehat{H}_0|H_1)P_1}$$

$$= \frac{P_1 + P_0 \frac{P(\widehat{H}_1|H_0)}{P(\widehat{H}_1|H_1)}}{p_0 + P_1 \frac{P(\widehat{H}_0|H_1)}{P(\widehat{H}_0|H_0)}} \cdot \frac{P(\widehat{H}_1|H_1)}{P(\widehat{H}_0|H_0)}$$

$$= \frac{P_1 + P_0\alpha}{P_1\beta + P_0}\gamma = r_0 \frac{1 + \frac{\alpha}{r_0}}{1 + r_0\beta}\gamma. \tag{3.41}$$

According to the definitions of ϵ_i, $\widehat{r}_i$ and r_i, the following weight error is obtained:

$$\begin{cases} \epsilon_0 &= \log(\frac{1+\frac{\alpha}{r_0}}{1+r_0\beta}) + \log(\gamma), \\ \epsilon_i &= \begin{cases} \log\frac{1+\beta r_0}{1+\frac{\alpha}{r_0}} + \log\frac{1+\frac{\alpha}{r_0 r_i}}{1+\beta r_0 r_i} & \text{if} \quad u_i = +1, \\ \log\frac{1+\frac{\alpha}{r_0}}{1+\beta r_0} + \log\frac{1+\frac{\beta r_0}{r_i}}{1+r_i\frac{\alpha}{r_0}} & \text{if} \quad u_i = -1. \end{cases} \end{cases} \tag{3.42}$$

From Eq. (3.42), we know that when $\alpha = 0$, β=0 and γ=1 , ϵ_i (for i=0,1, ... , n) would achieve its minimum. In Lemma II, we have proved that α and β are monotonically decreasing with τ_1 and τ_2. Thus, when α and β achieve their minimum, ϵ_i (for i=0,1, ..., n) also get its minimum, and thus:

$$\begin{cases} \epsilon_0 &= \log(\frac{1+\frac{\alpha_{min}}{r_0}}{1+r_0\beta_{min}}) + \log(\gamma), \\ \epsilon_i &= \begin{cases} \log\frac{1+\beta_{min} r_0}{1+\frac{\alpha_{min}}{r_0}} + \log\frac{1+\frac{\alpha_{min}}{r_0 r_i}}{1+\beta_{min} r_0 r_i} & \text{if} \quad u_i = +1, \\ \log\frac{1+\frac{\alpha_{min}}{r_0}}{1+\beta_{min} r_0} + \log\frac{1+\frac{\beta_{min} r_0}{r_i}}{1+r_i\frac{\alpha_{min}}{r_0}} & \text{if} \quad u_i = -1, \end{cases} \end{cases} \tag{3.43}$$

where α_{min}, β_{min} and γ are defined in Lemma II, r_i and r_0 are defined in Eq. (3.35). Note that the minimum error is uniquely determined by P_1 , P_0 and the parameters of sensors (P_{Fi} and P_{Mi}) . Q. E. D.

Note that $(y_i)_{max}$ and $(y_i)_{min}$ vary from sensor to sensor. In order to enable every sensor to adjust its weight and achieve the least error, the maximum value of τ_2 is chosen to be the minimum of all $(y_i)_{max}$, and the minimum value of τ_1 is chosen to be the maximum of all $(y_i)_{min}$. That is:

$$(\tau_2)_{max} = min\{(y_1)_{max}, (y_2)_{max}, ...(y_n)_{max}\}, \tag{3.44}$$

$$(\tau_1)_{min} = max\{(y_1)_{min}, (y_2)_{min}, \ldots\ldots, (y_n)_{min}\}. \tag{3.45}$$

Lemmas II and III illustrate how close the estimated weights can reach actual weights.

3.2 The Reinforcement Updating Rule

The distributed decision system is assumed to have no knowledge of the probability mass functions of the observations. Thus, the estimated probability of detection and false alarm for the ith detector $\widehat{P}_{Di}$ and $\widehat{P}_{Fi}$ can be approximated by relative frequencies. Let m be the number of $\widehat{H}_1$, n the number of $\widehat{H}_0$, and

m_{1i} the number of $u_i = +1$ and $\bar{u}_i = +1$,

m_{0i} the number of $u_i = -1$ and $\bar{u}_i = -1$,

n_{1i} the number of $u_i = +1$ and $\bar{u}_i = -1$,

n_{0i} the number of $u_i = -1$ and $\bar{u}_i = +1$.

Hence, m, n, m_{1i}, m_{0i}, n_{1i} and n_{0i} can simply be obtained by counting in the simulations. That is,

$$\begin{aligned} \frac{m}{n} &\approx \frac{P(\widehat{H}_1)}{P(\widehat{H}_0)}, \\ \frac{m_{1i}}{n_{1i}} &\approx \frac{P(u_i = +1, \widehat{H}_1)}{P(u_i = +1, \widehat{H}_0)}, \\ \frac{m_{0i}}{n_{0i}} &\approx \frac{P(u_i = -1, \widehat{H}_0)}{P(u_i = -1, \widehat{H}_1)}. \end{aligned} \tag{3.46}$$

We shall next develop the updating rule for the fusion center. From Eq.(3.36),

$$\widehat{w}_0 = \log \frac{P(\widehat{H}_1)}{P(\widehat{H}_0)}.$$

$$\widehat{w}_i = \begin{cases} \log \frac{P(u_i=+1|\widehat{H}_1)}{P(u_i=+1|\widehat{H}_0)}, & \text{if} \quad u_i = +1, \\ \\ \log \frac{P(u_i=-1|\widehat{H}_0)}{P(u_i=-1|\widehat{H}_1)}, & \text{if} \quad u_i = -1. \end{cases}$$

Using the Bayes rule, $P(x,y) = p(x|y)P(y)$,

$$\widehat{w}_i = \begin{cases} \log \frac{P(u_i=+1,\widehat{H}_1)P(\widehat{H}_0)}{P(u_i=+1,\widehat{H}_0)P(\widehat{H}_1)}, & \text{if} \quad u_i = +1, \\ \\ \log \frac{P(u_i=-1,\widehat{H}_0)P(\widehat{H}_1)}{P(u_i=-1,\widehat{H}_1)P(\widehat{H}_0)}, & \text{if} \quad u_i = -1. \end{cases} \tag{3.47}$$

Applying Eq.(4.1) and Eq.(3.36) yields

$$\widehat{w}_0 \approx \log \frac{m}{n}$$

$$\widehat{w}_i \approx \begin{cases} \log \frac{m_{1i}}{n_{1i}} - \widehat{w}_0, & \text{if} \quad u_i = +1, \\ \\ \log \frac{m_{0i}}{n_{0i}} + \widehat{w}_0, & \text{if} \quad u_i = -1, \end{cases} \tag{3.48}$$

and

$$\begin{aligned} m &\approx e^{\widehat{w}_0} n, \\ m_{1i} &\approx n_{1i}\exp(\widehat{w}_i + \widehat{w}_0) \quad \text{if } u_i = +1, \\ m_{0i} &\approx n_{0i}\exp(\widehat{w}_i - \widehat{w}_0) \quad \text{if } u_i = -1. \end{aligned} \tag{3.49}$$

Taking the partial derivative of Eq. (6.13) with respect to m, n, m_{1i}, m_{0i}, n_{1i} and n_{0i}, respectively,

$$\begin{aligned} \frac{\partial \widehat{w}_0}{\partial m} &\approx \frac{1}{m}, \\ \frac{\partial \widehat{w}_0}{\partial n} &\approx -\frac{1}{n} = -\frac{1}{m} e^{\widehat{w}_0}, \end{aligned} \tag{3.50}$$

and

$$\frac{\partial \widehat{w}_i}{\partial m_{1i}} \approx \frac{1}{m_{1i}}, \quad \frac{\partial \widehat{w}_i}{\partial n_{1i}} \approx -\frac{1}{m_{1i}} e^{\widehat{w}_i + \widehat{w}_0} \qquad \text{if } u_i = +1, \tag{3.51}$$

$$\frac{\partial \widehat{w}_i}{\partial m_{0i}} \approx \frac{1}{m_{0i}}, \quad \frac{\partial \widehat{w}_i}{\partial n_{0i}} \approx -\frac{1}{m_{0i}} e^{\widehat{w}_i - \widehat{w}_0} \qquad \text{if } u_i = -1. \tag{3.52}$$

If the current local detector's decision conforms to the reliable fusion, its weight $\widehat{w}_i$ should be reinforced. In this case,

$$\Delta \widehat{w}_i \approx \begin{cases} \frac{1}{m_{1i}} \Delta m_{1i} = \frac{1}{m_{1i}}, & \text{if} \quad u_i = +1 \quad \text{and} \quad \widehat{H}_1, \\ \\ \frac{1}{m_{0i}} \Delta m_{0i} = \frac{1}{m_{0i}} & \text{if} \quad u_i = -1 \quad \text{and} \quad \widehat{H}_0. \end{cases} \tag{3.53}$$

Table 3.1 Adaptive fusion rule for independent source.

	$\widehat{H}_1$		$\widehat{H}_0$	
	u_i=+1	u_i=-1	u_i=+1	u_i=-1
$\Delta\widehat{w}_0$	$\frac{1}{m}$		$-\frac{1}{m}e^{\widehat{w}_0^-}$	
$\Delta\widehat{w}_i$	$\frac{1}{m_{1i}}$	$-\frac{1}{m_{0i}}e^{\widehat{w}_i^- - \widehat{w}_0^-}$	$-\frac{1}{m_{1i}}e^{\widehat{w}_i^- + \widehat{w}_0^-}$	$\frac{1}{m_{0i}}$

On the other hand, if the current local decision contradicts the reliable decision, its weight $\widehat{w}_i$ should be reduced. That is,

$$\Delta\widehat{w}_i \approx \begin{cases} -\frac{1}{n_{1i}}\Delta n_{1i} = -\frac{1}{m_{1i}}e^{\widehat{w}_i+\widehat{w}_0}\Delta n_{1i} = -\frac{1}{m_{1i}}e^{\widehat{w}_i+\widehat{w}_0}, & \text{if} \quad u_i = +1 \quad \text{and} \quad \widehat{H}_0, \\ -\frac{1}{n_{0i}}\Delta n_{0i} = -\frac{1}{m_{0i}}e^{\widehat{w}_i-\widehat{w}_0}\Delta n_{0i} = -\frac{1}{m_{0i}}e^{\widehat{w}_i-\widehat{w}_0}, & \text{if} \quad u_i = -1 \quad \text{and} \quad \widehat{H}_1, \end{cases} \tag{3.54}$$

and

$$\Delta\widehat{w}_0 \approx \begin{cases} \frac{1}{m}\Delta m = \frac{1}{m} & \text{when} \quad \widehat{H}_1 \quad \text{occurs}, \\ -\frac{1}{n}\Delta n = -\frac{1}{m}e^{\widehat{w}_0}\Delta n = -\frac{1}{m}e^{\widehat{w}_0} & \text{when} \quad \widehat{H}_0 \quad \text{occurs}. \end{cases} \tag{3.55}$$

Thus, we obtain the following updating rule:

$$\widehat{w}_i^+ = \widehat{w}_i^- + \Delta\widehat{w}_i, \; i = 0, 1, 2, \cdots, \tag{3.56}$$

where $\widehat{w}_i^+$ and $\widehat{w}_i^-$ represent the weight after and before each update. Since the steady state $\widehat{w}_i's$ are what we are trying to compute, for actual implementation, we use the current estimated weight $\widehat{w}_i^-$ to compute $\Delta\widehat{w}_i$. That is, to update the weights according to Eq.(3.56), $\Delta\widehat{w}_i$ is computed according to the Table 3.1.

Lemma IV Using the updating rule according to Eq.(3.56) and the Table 3.1, $\widehat{w}_i^-$ will converge to the desired steady state estimated weight $\widehat{w}_i$.

Proof: At steady state,

$$E[\widehat{w}_i^+ - \widehat{w}_i^-] = 0. \tag{3.57}$$

Using the definition $E[X] = \sum x_i P(x_i)$, the updating rule according to Eq.(3.56) and the above table, with u_i=+1, Eq.(3.57) becomes,

$$\frac{1}{m_{1i}}P(u=+1,\widehat{H}_1)-\frac{1}{m_{1i}}e^{\widehat{w}_i^-+\widehat{w}_0^-}P(u=+1,\widehat{H}_0)=0.$$

Using Eq.(3.47) for further simplification yields,

$$\widehat{w}_i^- + \widehat{w}_0^- = \widehat{w}_i + \widehat{w}_0. \tag{3.58}$$

Similarly, if $u_i=-1$, we have

$$\widehat{w}_i^- - \widehat{w}_0^- = \widehat{w}_i - \widehat{w}_0. \tag{3.59}$$

For $i=0$, the following condition can similarly be obtained at steady state:

$$\tfrac{1}{m}P(\widehat{H}_1)-\tfrac{1}{m}e^{\widehat{w}_0^-}P(\widehat{H}_0)=0.$$

Thus,

$$\widehat{w}_0^- = \widehat{w}_0. \tag{3.60}$$

Hence, $\widehat{w}_i^- \to \widehat{w}_i$, for $i=0,1,\cdots$., n. Q. E. D.

Convergence is a crucial criterion in determining whether an adaptive algorithm is workable. Lemma IV, which proves the convergence of our proposed algorithm, analytically justifies the validity of the algorithm.

3.3 Simulation

Figure 3.4 shows the simulation set up to validate the proposed adaptive fusion model. In the simulation presented here, the source produces a binary signal with $P(H_1)=0.3$ and $P(H_0)=0.7$, where $H_1:+1$ and $H_0:-1$. Eight sensors are used. The probabilities of false alarm and missing, P_F and P_M, of each sensor are fixed, but not known to the system. The channel is additive Gaussian noise. The Gaussian random variables are generated according to the following transformation:

$$\begin{cases} x=(-2\ln r_1)^{1/2}\cos 2\pi r_2 \\ y=(-2\ln r_1)^{1/2}\sin 2\pi r_2, \end{cases}$$

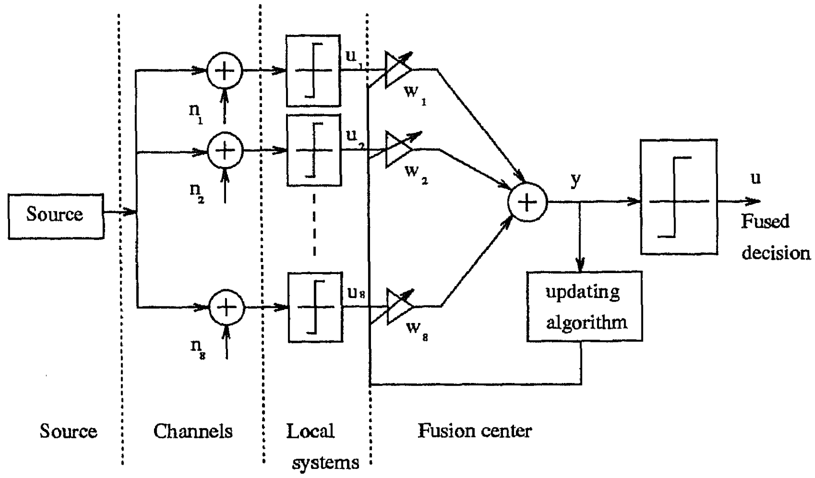

Figure 3.4 Computer simulation diagram.

where r_1 and r_2 are uniformly distributed on $(0, 1)$, and (x, y) becomes a pair of orthogonally normalized *Gaussian* random variables. The additive Gaussian variable for each sensor is zero-mean with a standard deviation ranging from 0.5 to 1.2.

3.3.1 Conditional Probability Mass Function of y

Figure 3.5 shows the histograms of $P(y = \zeta|H_0)$ and $P(y = \zeta|H_1)$ for 8 sensors and 250000 samples. We can see that the they are not monotonic. Figure 3.6, which illustrates $\log \frac{P(y=\zeta|H_1)}{P(y=\zeta|H_0)}$, is almost a straight line, conforming to Lemma I:

$$e^{\zeta} = \frac{P(y-w_0=\zeta|H_1)}{P(y-w_0=\zeta|H_0)}.$$

3.3.2 Convergence of Weights

Figure 3.7 shows average errors of weights $|w_i - \widehat{w}_i|$ for different τ, $\tau = 0$, 0.25y_{max}, and 0.5y_{max}. Here, $\tau = |\tau_1| = |\tau_2|$. As shown in the figure, the larger the τ, the smaller the error, which agrees with Lemma II. As the number of unreliable samples increases, the training time becomes longer.

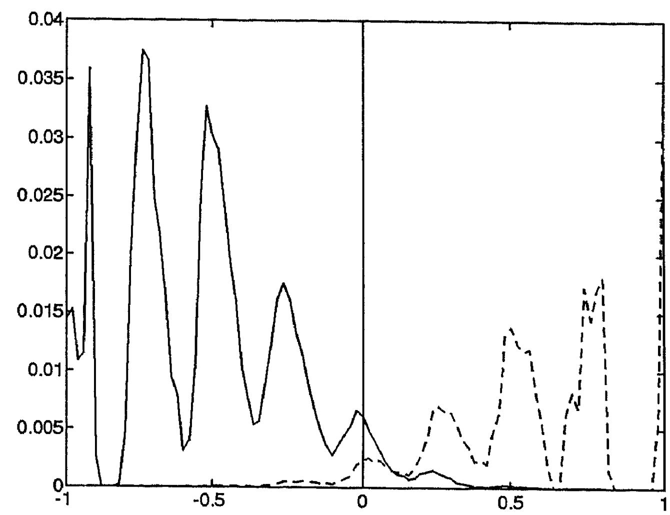

Figure 3.5 Probability mass functions $P(y|H_1)$ and $P(y|H_0)$.

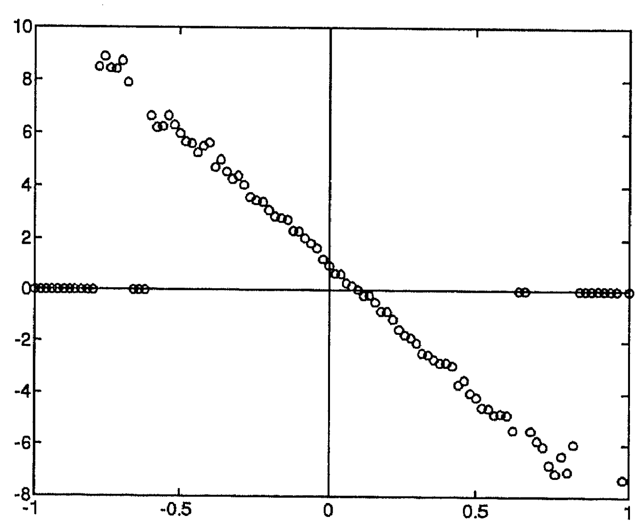

Figure 3.6 The log-ratio of probability mass functions $logP(y|H_1)/P(y|H_0)$.

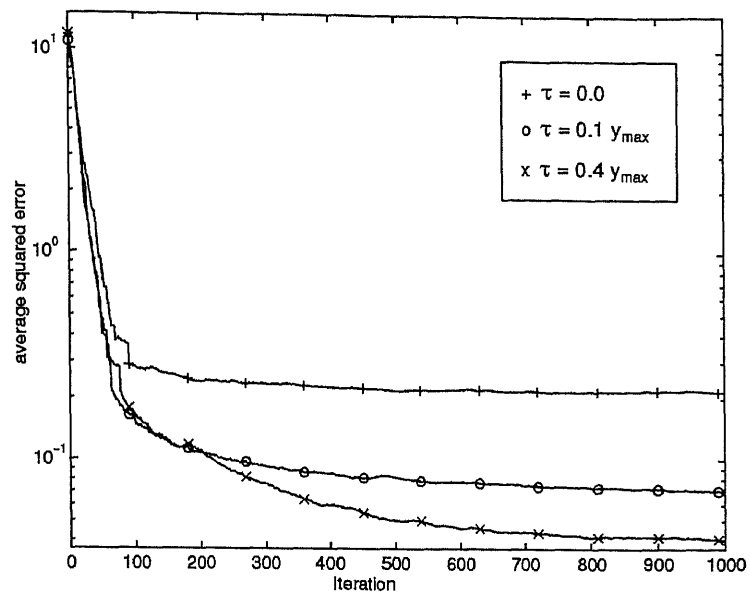

Figure 3.7 The error with various reliability thresholds.

3.4 Summary

In the real-world environment, the probability mass functions of the observations at local detectors may not be known and the performance of the local detectors may not be consistent. Under such circumstances, a system which can adapt itself during the decision making process is needed. The major advantage is that the system can still have smaller error and does not need *a priori* knowledge of the probability mass functions of the observations. Simulation results conform to our theoretical analysis.

CHAPTER 4

ADAPTIVE FUSION OF CORRELATED LOCAL DECISIONS

In this chapter, we derive another form of the MAP-based optimal fusion rule and extend our adaptive algorithm by considering dependent/correlated decisions. The chapter is arranged as follows. In Section 3.1, we develop and derive the optimal fusion rule for correlated decisions. The adaptive fusion rule and the proof of its convergence are discussed in Section 3.2. The residue between error probabilities obtained using the optimal fusion rule and the adaptive fusion rule is analyzed in Section 3.3. Section 3.4 illustrates the effect of the number of sensors and correlation coefficients on the error probability in a Gaussian noise environment. Simulations are presented in Section 3.5. Conclusions are drawn in Section 3.6.

4.1 The Optimal Fusion Rule for Correlated Decisions

Consider the binary hypothesis testing problem with N sensors in which each sensor employs a predetermined decision rule. The two hypotheses have *a priori* probabilities, $P(H_1)$ and $P(H_0)$, respectively. In binary detection theory, one of the most popular detection criteria is the likelihood ratio criterion. The likelihood ratio is expressed as

$$\begin{aligned}\Lambda(u_1, u_2, ...u_N) &= \frac{P(u_1, u_2, ...u_N|H_1)}{P(u_1, u_2, ...u_N|H_0)} \\ &= \frac{P_1(u_1)P_1(u_2|u_1)P_1(u_3|u_1,u_2)...P_1(u_N|u_1,u_2,...u_{N-1})}{P_0(u_1)P_0(u_2|u_1)P_0(u_3|u_1,u_2)...P_0(u_N|u_1,u_2,...u_{N-1})}, \end{aligned} \tag{4.1}$$

where

$$P_i(u_k|u_1, u_2, ...u_{k-1}) = P(u_k|u_1, u_2, ...u_{k-1}, H_i), \quad i = 0, 1 \tag{4.2}$$

are conditional probabilities, and $u_1, u_2, \cdots, u_N$ are local decisions that are binary random variables. H_0 and H_1 represent the following two hypotheses:

H_0 : Signal is absent;

$$H_1: \quad \text{Signal is present.}$$

Here, u_k , for $k = 1, 2, \cdots, N$, is defined by

$$u_k = \begin{cases} -1 & \text{if } H_0 \text{ is declared,} \\ +1 & \text{if } H_1 \text{ is declared.} \end{cases}$$

Since

$$\begin{aligned} P_i(u_k|u_1, u_2, ...u_{k-1}) &= \frac{P_i(u_1, u_2, ...u_k)}{P_i(u_1, u_2, ...u_{k-1})} \\ &= \frac{P_i(u_1, u_2, ...u_k)}{P_i(u_1, u_2, ..., u_{k-1}, u_k = +1) + P_i(u_1, u_2, ..., u_{k-1}, u_k = -1)} \\ &= \frac{1}{\frac{P_i(u_1, u_2, ..., u_{k-1}, u_k = +1)}{P_i(u_1, u_2, ...u_k)} + \frac{P_i(u_1, u_2, ..., u_{k-1}, u_k = -1)}{P_i(u_1, u_2, ...u_k)}}, \end{aligned}$$

we have

$$P_i(u_k|u_1, u_2, ...u_{k-1}) = \begin{cases} \frac{1}{1 + \frac{P_i(u_1, u_2, ..., u_{k-1}, u_k = -1)}{P_i(u_1, u_2, ..., u_{k-1}, u_k = +1)}} & \text{if } u_k = +1 \\ \frac{1}{1 + \frac{P_i(u_1, u_2, ..., u_{k-1}, u_k = +1)}{P_i(u_1, u_2, ..., u_{k-1}, u_k = -1)}} & \text{if } u_k = -1. \end{cases} \tag{4.3}$$

Let

$$p_k = \frac{P_1(u_1, u_2, ..., u_{k-1}, u_k = -1)}{P_1(u_1, u_2, ..., u_{k-1}, u_k = +1)}, \quad q_k = \frac{P_0(u_1, u_2, ..., u_{k-1}, u_k = -1)}{P_0(u_1, u_2, ..., u_{k-1}, u_k = +1)}. \tag{4.4}$$

So

$$P_1(u_k|u_1, u_2, ...u_{k-1}) = \begin{cases} \frac{1}{1+p_k} & \text{if } u_k = +1 \\ \frac{p_k}{1+p_k} & \text{if } u_k = -1, \end{cases} \tag{4.5}$$

$$P_0(u_k|u_1, u_2, ...u_{k-1}) = \begin{cases} \frac{1}{1+q_k} & \text{if } u_k = +1 \\ \frac{q_k}{1+q_k} & \text{if } u_k = -1. \end{cases} \tag{4.6}$$

By defining the weight W_k for $k = 0, 1, \cdots, N$ as:

$$W_k = \begin{cases} \log \frac{P(H_1)}{P(H_0)} & \text{for } k = 0 \\ \begin{cases} \log \frac{P_1(u_1)}{P_0(u_1)} & \text{if } u_k = +1 \\ \log \frac{P_0(u_1)}{P_1(u_1)} & \text{if } u_k = -1 \end{cases} & \text{for } k = 1 \\ \begin{cases} \log \frac{P_1(u_k|u_1, u_2, ..., u_{k-1})}{P_0(u_k|u_1, u_2, ..., u_{k-1})} & \text{if } u_k = +1 \\ \log \frac{P_0(u_k|u_1, u_2, ..., u_{k-1})}{P_1(u_k|u_1, u_2, ..., u_{k-1})} & \text{if } u_k = -1 \end{cases} & \text{for } k > 1, \end{cases} \tag{4.7}$$

we have

$$W_k = \begin{cases} \log\frac{1+q_k}{1+p_k} & \text{if } u_k = +1 \text{ and } k > 1 \\ \log\frac{q_k(1+p_k)}{p_k(1+q_k)} & \text{if } u_k = -1 \text{ and } k > 1 \end{cases}. \tag{4.8}$$

The Maximum *A Posteriori* Probability (MAP) or Minimum Error Probability detection rule is

$$\Lambda(u_1, u_2, ...u_N) = \frac{P_1(u_1, u_2, ...u_N)}{P_0(u_1, u_2, ...u_N)} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \frac{P(H_0)}{P(H_1)}. \tag{4.9}$$

According to the above equation, the optimal fusion rule based on the MAP detection criterion for correlated local decisions is equivalent to

$$\sum_{i=0}^{N} W_i \cdot u_i \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} 0, \tag{4.10}$$

where u_0 is always set to 1.

Comparing Eq. (4.7) with the results developed by Chair and Varshney in [7] for the independent case, it can be seen that when the local decisions are independent, Eq. (4.7) is the same as that in [7]. Thus, Eq. (4.7) is a generalization of Chair and Varshney's result for the correlated case.

4.2 The Adaptive Fusion Rule

The optimal fusion rule derived in Section 2 requires the knowledge of *a priori* probabilities and conditional probabilities that are either difficult to acquire or time-varying. To realize the optimal fusion, an adaptive algorithm is necessary to estimate these probabilities.

4.2.1 Adaptive Fusion Rule

Similar to the independent case [4, 3], denote the events of the fusion results being +1 and -1 by $\widehat{H_1}$ and $\widehat{H_0}$, respectively. In addition, let m be the number of events in which $\widehat{H_1}$ occurs, n the number of events in which $\widehat{H_0}$ occurs, and

$m_{k,1}$ the number of events in which $(u_1, u_2, \ldots\ldots, u_{k-1}, u_k = +1, \widehat{H_1})$ occurs,

$m_{k,0}$ the number of events in which $(u_1, u_2, \ldots\ldots, u_{k-1}, u_k = -1, \widehat{H_1})$ occurs,

$n_{k,1}$ the number of events in which $(u_1, u_2, \ldots\ldots, u_{k-1}, u_k = +1, \widehat{H_0})$ occurs,

$n_{k,0}$ the number of events in which $(u_1, u_2, \ldots\ldots, u_{k-1}, u_k = -1, \widehat{H_0})$ occurs.

Similar to Eq. (4.7), define

$$\widehat{W_k} = \begin{cases} \log\frac{\widehat{P(H_1)}}{\widehat{P(H_0)}} & \text{for } k = 0 \\ \begin{cases} \log\frac{\widehat{P_1(u_1)}}{\widehat{P_0(u_1)}} & \text{if } u_k = +1 \\ \log\frac{\widehat{P_0(u_1)}}{\widehat{P_1(u_1)}} & \text{if } u_k = -1 \end{cases} & \text{for } k = 1 \\ \begin{cases} \log\frac{\widehat{P_1}(u_k|u_1, u_2, \ldots\ldots, u_{k-1})}{\widehat{P_0}(u_k|u_1, u_2, \ldots\ldots, u_{k-1})} & \text{if } u_k = +1 \\ \log\frac{\widehat{P_0}(u_k|u_1, u_2, \ldots\ldots, u_{k-1})}{\widehat{P_1}(u_k|u_1, u_2, \ldots\ldots, u_{k-1})} & \text{if } u_k = -1 \end{cases} & \text{for } k > 1, \end{cases} \tag{4.11}$$

$$\widehat{p_k} = \frac{P(u_1, u_2, \ldots\ldots, u_{k-1}, u_k = -1|\widehat{H_1})}{P(u_1, u_2, \ldots\ldots, u_{k-1}, u_k = +1|\widehat{H_1})} = \frac{P(u_1, u_2, \ldots\ldots, u_{k-1}, u_k = -1, \widehat{H_1})}{P(u_1, u_2, \ldots\ldots, u_{k-1}, u_k = +1, \widehat{H_1})},$$
$$\widehat{q_k} = \frac{P(u_1, u_2, \ldots\ldots, u_{k-1}, u_k = -1|\widehat{H_0})}{P(u_1, u_2, \ldots\ldots, u_{k-1}, u_k = +1|\widehat{H_0})} = \frac{P(u_1, u_2, \ldots\ldots, u_{k-1}, u_k = -1, \widehat{H_0})}{P(u_1, u_2, \ldots\ldots, u_{k-1}, u_k = +1, \widehat{H_0})},$$

where the symbols with "hat" are estimations of symbols without "hat." $\widehat{p_k}$ and $\widehat{q_k}$ can be approximated by

$$\widehat{p_k} \approx \frac{m_{k,0}}{m_{k,1}} \qquad \widehat{q_k} \approx \frac{n_{k,0}}{n_{k,1}}. \tag{4.12}$$

When u_k=+1,

$$\widehat{W_k} = \log\frac{1 + \widehat{q_k}}{1 + \widehat{p_k}} \approx \log\frac{m_{k,1}}{n_{k,1}}\frac{n_{k,1} + n_{k,0}}{m_{k,1} + m_{k,0}}.$$

Note that

$$n_{k,1} + n_{k,0} = n_{k-1,j} \qquad m_{k,1} + m_{k,0} = m_{k-1,j},$$

where j is the output of the $(k-1)$th local sensor; that is,

$$j = \begin{cases} 1 & \text{if } u_{k-1} = +1 \\ 0 & \text{if } u_{k-1} = -1 \ . \end{cases} \tag{4.13}$$

Thus,

$$\widehat{W_k} \approx \log \frac{m_{k,1}}{n_{k,1}} - \log \frac{m_{k-1,j}}{n_{k-1,j}}. \tag{4.14}$$

Following the same reasoning, the approximated weight for $u_k = -1$ is

$$\widehat{W_k} \approx \log \frac{m_{k-1,j}}{n_{k-1,j}} - \log \frac{m_{k,0}}{n_{k,0}}. \tag{4.15}$$

$\widehat{W_k}$ exhibits the following property

$$\widehat{W_k}|_{u_k=+1} + \widehat{W_k}|_{u_k=-1} = \log \frac{m_{k,1}}{n_{k,1}} - \log \frac{m_{k,0}}{n_{k,0}} = \log \frac{\widehat{q_k}}{\widehat{p_k}}. \tag{4.16}$$

The partial derivatives of $\widehat{W_k}$ with respect to $m_{k,1}$, $m_{k,0}$, $n_{k,1}$ and $n_{k,0}$ are

$$\begin{aligned} \frac{\partial \widehat{W}_k}{\partial m_{k,1}} &\approx \frac{1}{m_{k,1}}, \quad \frac{\partial \widehat{W}_k}{\partial n_{k,1}} \approx -\frac{1}{m_{k,1}} \frac{m_{k-1,j}}{n_{k-1,j}} e^{\widehat{W}_k} \qquad \text{if } u_k = +1, \\ \frac{\partial \widehat{W}_k}{\partial m_{k,0}} &\approx -\frac{1}{m_{k,0}}, \quad \frac{\partial \widehat{W}_k}{\partial n_{k,0}} \approx \frac{1}{m_{k,0}} \frac{m_{k-1,j}}{n_{k-1,j}} e^{-\widehat{W}_k} \qquad \text{if } u_k = -1. \end{aligned}$$

According to the concept of reinforcement learning [33], if the current local detector's decision conforms to that of the fusion center, its weight $\widehat{W_k}$ should be reinforced. In this case,

$$\Delta \widehat{W}_k \approx \begin{cases} \frac{1}{m_{k,1}} \Delta m_{k,1} = \frac{1}{m_{k,1}}, & \text{if} \quad u_k = +1 \quad \text{and} \quad \widehat{H}_1, \\ \\ \frac{1}{n_{k,0}} \Delta n_{k,0} = \frac{1}{m_{k,0}} \frac{m_{k-1,j}}{n_{k-1,j}} e^{-\hat{W}_k} & \text{if} \quad u_k = -1 \quad \text{and} \quad \widehat{H}_0. \end{cases} \tag{4.17}$$

On the other hand, if the current local decision contradicts that of the fusion, its weight $\widehat{W}_i$ should be reduced. That is,

$$\Delta \widehat{W}_k \approx \begin{cases} -\frac{1}{m_{k,0}} \Delta m_{k,0} = -\frac{1}{m_{k,0}}, & \text{if} \quad u_k = -1 \quad \text{and} \quad \widehat{H}_1, \\ \\ -\frac{1}{n_{k,1}} \Delta n_{k,1} = -\frac{1}{m_{k,1}} \frac{m_{k-1,j}}{n_{k-1,j}} e^{\hat{W}_k} & \text{if} \quad u_k = +1 \quad \text{and} \quad \widehat{H}_0. \end{cases} \tag{4.18}$$

Hence, the adaptive fusion rule is

$$\widehat{W}_k^+ = \widehat{W}_k^- + \Delta\widehat{W}_k,\ k = 0, 1, 2, \cdots, N, \tag{4.19}$$

where $\widehat{W}_k^+$ and $\widehat{W}_k^-$ represent the weight after and before each updating. The change of weights ΔW_k is summarized by the following table:

	$\widehat{H_1}$		$\widehat{H_0}$	
	u_k=+1	u_k=-1	u_k=+1	u_k=-1
$\Delta\widehat{W}_0$	$\frac{1}{m}$		$-\frac{1}{m}e^{\widehat{W}_0^-}$	
$\Delta\widehat{W}_k$	$\frac{1}{m_{k,1}}$	$-\frac{1}{m_{k,0}}$	$-\frac{1}{m_{k,1}}\frac{m_{k-1,j}}{n_{k-1,j}}e^{\widehat{W}_k^-}$	$\frac{1}{m_{k,0}}\frac{m_{k-1,j}}{n_{k-1,j}}e^{-\widehat{W}_k^-}$

4.2.2 Proof of Convergence

Since $\widehat{W}_0$ is the same as that in the independent case, its convergence can be proved similarly to our previous work [4]. Here, we only consider the convergence of $\widehat{W}_k$ for $k > 1$. From Eq. (4.19), it is easily seen that convergence of $\widehat{W}_k$ is equivalent to $\Delta\widehat{W}_k \to 0$. According to Eqs. (4.17) and (4.18), the expected value of $\Delta\widehat{W}_k$ can be expressed as

$$\begin{aligned} E[\Delta\widehat{W}_k] &= +\frac{1}{m_{k1}}P(u_k = +1, \widehat{H_1}) + \frac{1}{n_{k0}}P(u_k = -1, \widehat{H_0}) \\ &\quad -\frac{1}{m_{k0}}P(u_k = -1, \widehat{H_1}) - \frac{1}{n_{k1}}P(u_k + = 1, \widehat{H_0}). \end{aligned}$$

When the number of iterations increases, m_{k1}, n_{k1}, m_{k0} and n_{k0} will approach infinity, while $P(u_k = +1, \widehat{H_1})$, $P(u_k = -1, \widehat{H_1})$, $P(u_k = +1, \widehat{H_0})$ and $P(u_k = -1, \widehat{H_0})$ are always between zero and one. Thus,

$$\lim_{n+m\to\infty} E[\Delta\widehat{W}_k] = 0,$$

where the number of iterations equals $n + m$. Following the same reasoning, it can be shown that the variance and higher moments of $\Delta\widehat{W}_k$ approach zero when the number of iterations goes to infinity. According to the theory of probability [20], it can be concluded that the following equation holds with probability 1.

$$\lim_{n+m\to\infty} \Delta\widehat{W}_k = 0.$$

Thus, $\widehat{W_k}$ converges asymptotically to a real number with probability 1. This completes our proof. □

4.3 Error Analysis

It has been shown in Section 4.2 that the adaptive fusion rule converges asymptotically. However, the algorithm is not guaranteed to converge to the optimal weights. To compare the performance between the optimal and adaptive algorithms, error probabilities obtained by these two methods are investigated. Based on the previous analysis, the ideal optimum decision is

$$y = \log\frac{P(U,H_1)}{P(U,H_0)} \mathop{\gtrless}_{H_0}^{H_1} 0. \tag{4.20}$$

Using the proposed adaptive algorithm, the decision becomes

$$\widehat{y} = \log\frac{P(U,\widehat{H_1})}{P(U,\widehat{H_0})} \mathop{\gtrless}_{H_0}^{H_1} 0, \tag{4.21}$$

where y and $\widehat{y}$ are linear combinations of the local decisions. Let $U = (u_1, u_2, \cdots, u_N)$ be the vector representation of the local decisions.

Since

$$P(\widehat{H_1},U) = P(H_0,U)P(\widehat{H_1}|U,H_0) + P(H_1,U)P(\widehat{H_1}|U,H_1) \tag{4.22}$$

$$P(\widehat{H_0},U) = P(H_0,U)P(\widehat{H_0}|U,H_0) + P(H_1,U)P(\widehat{H_0}|U,H_1), \tag{4.23}$$

the adaptive fusion algorithm can be written as

$$\widehat{y} = y + \log\frac{P(\widehat{H_1}|U,H_1) + e^{-y}P(\widehat{H_1}|U,H_0)}{P(\widehat{H_0}|U,H_0) + e^{+y}P(\widehat{H_0}|U,H_1)}. \tag{4.24}$$

Thus, the decision rule derived in Eq. (4.21) becomes

$$y \underset{H_0}{\overset{H_1}{\gtrless}} T(U), \tag{4.25}$$

where

$$T(U) = \log \frac{P(\widehat{H_0}|U, H_0) + e^{+y} P(\widehat{H_0}|U, H_1)}{P(\widehat{H_1}|U, H_1) + e^{-y} P(\widehat{H_1}|U, H_0)}. \tag{4.26}$$

In comparing Equations (4.20) and (4.25), the decision rule using the adaptive algorithm is equivalent to the optimal decision rule offset by T(U).

The error probability using the optimum decision rule is defined by

$$P_e = P(H_0|H_1)P(H_1) + P(H_1|H_0)P(H_0), \tag{4.27}$$

where

$$P(H_0|H_1) = \sum_{y<0} P(U|H_1), \quad P(H_1|H_0) = \sum_{y>0} P(U|H_0). \tag{4.28}$$

The error probability using the adaptive decision rule is

$$P_e^{'} = P^{'}(H_0|H_1)P(H_1) + P^{'}(H_1|H_0)P(H_0), \tag{4.29}$$

where

$$P^{'}(H_0|H_1) = \sum_{y<T(U)} P(U|H_1), \quad P^{'}(H_1|H_0) = \sum_{y>T(U)} P(U|H_0). \tag{4.30}$$

Since the optimum detection rule achieves the minimum error probability, $P_e^{'}$ is usually larger than P_e. The degradation in performance can be measured by the absolute difference between the two error probabilities, that is

$$\begin{aligned} \Delta P_e &= \left|P_e^{'} - P_e\right| \\ &= \left|P(H_1)\left(\sum_{y<T(U)} P(U|H_1) - \sum_{y<0} P(U|H_1)\right) + \right. \\ &\quad \left. P(H_0)\left(\sum_{y>T(U)} P(U|H_0) - \sum_{y>0} P(U|H_0)\right)\right|. \end{aligned} \tag{4.31}$$

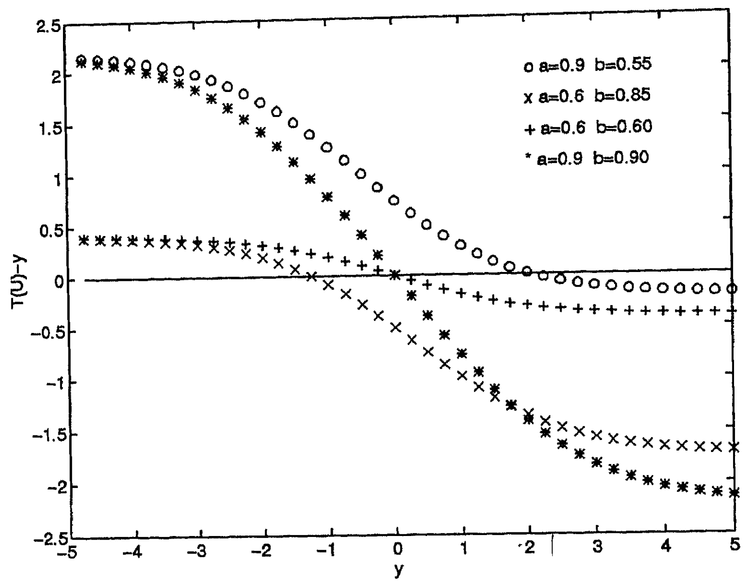

Figure 4.1 T(U)-y versus y, where T(U) is the offset and y is the linear combination of local decisions.

Comparing the decision rules using optimal and adaptive fusion, the following is desirable

when y>0, T(U)-y<0, and
when y<0, T(U)-y>0,

such that $\Delta P_e = 0$. From Eq. (4.26), the offset increases with y monotonically. If $P(\widehat{H_0}|U, H_0) > 0.5$ and $P(\widehat{H_1}|U, H_1) > 0.5$ (which is the usual case), $T(U) - y$ decreases with y monotonically (see Figure 4.1).

Proposition: Let a=$P(\widehat{H_0}|U, H_0)$, b=$P(\widehat{H_1}|U, H_1)$, 1-a=$P(\widehat{H_1}|U, H_0)$, 1-b=$P(\widehat{H_0}|U, H_1)$, and y_T denote the zero-crossing point of $T(U) - y$ (i.e., $y_T = \{y : T(U) - y = 0\}$). Since $T(U) - y$ decreases with y monotonically, y_T is unique:

$$y_T = log(\frac{a - 0.5}{b - 0.5}). \tag{4.32}$$

□

The closer the y_T to zero, the smaller the degradation in error probability. The adaptive rule behaves as well as the optimal decision rule when a=b, in which case $y_T = 0$. Figure 4.1 shows plots of $T(U) - y$ versus y for different values of a, b. By using the following relationship between y and $\Lambda(U)$:

$$y = \log(\Lambda(U)\frac{P(H_1)}{P(H_0)}), \tag{4.33}$$

many of the properties discussed in this section can also be expressed in term of the likelihood ratio function, $\Lambda(U)$.

4.4 Performance Analysis in Gaussian Noise

To gain an insight into the proposed adaptive fusion algorithm, performance analysis in Gaussian noise which is both theoretically tractable and computationally feasible is examined in this section. Suppose all of the sensors are corrupted by Gaussian noise that has a zero mean and the same variance of σ^2. Let the correlation coefficient, ρ, between different sensors be the same. Thus, the observation vector X at the local sensors is Gaussian-distributed. Let $\underline{\mu_1}$ =[1 1 1 ... 1 1] and $\underline{\mu_0}$ =[-1 -1 -1 ... -1 -1] be its mean vectors for H_1 and H_0, respectively. The correlation matrix of X is

$$C = \sigma^2 \begin{bmatrix} 1 & \rho & \rho & \cdots\cdots & \rho \\ \rho & 1 & \rho & \cdots\cdots & \rho \\ \cdots & \cdots & \cdots & \cdots\cdots & \\ \rho & \rho & \cdots & \cdots\cdots & 1 \end{bmatrix}.$$

In addition, suppose all the local sensors adopt the same decision threshold, t, implying that the optimal fusion rule is the same as the k out of N rule [43]. Further assume that $P(\widehat{H_0}|U, H_0)= P(\widehat{H_1}|U, H_1)$, i.e., $P_e{=}P_e^{'}$. Let $A_{N-k,k}(t,\rho)$ denote the joint probability of N random variables with the correlation coefficient ρ of which k out of N random variables are greater than t, and the other $N-k$ are less than t. If all the random variables are identical, it can be shown that

$$A_{N-k,k}(t,\rho) = \binom{N}{k} P(x_1 < t, x_2 < t, \cdots, x_{N-k} < t, x_{N-k+1} > t, x_{N-k+2} > t, \cdots, x_N > t), \tag{4.34}$$

and

$$A_{N-k,k}(t,\rho) = \sum_{j=0}^{k}(-1)^j\binom{k}{j}A_{N-k+j,0}(t,\rho). \tag{4.35}$$

When $x_i's$ are Gaussian random variables with a zero mean and correlation matrix C, $P(x_1 < t, x_2 < t, \cdots, x_{N-k+j} < t)$, can be expressed as [21]

$$P(x_1 < t, x_2 < t, \cdots, x_{N-k+j} < t) = \int_{-\infty}^{\infty} Q^{N-k+j}\left(\frac{t-\sqrt{\rho}y}{\sqrt{1-\rho}}\right)f(y)dy, \tag{4.36}$$

where $f(\cdot)$ and $Q(\cdot)$ are the standard normal density and cumulative distribution functions. Eq. (4.36) can be computed numerically. It follows from Eq. (4.34) that the two likelihood functions can be expressed as

$$P(U|H_1) = P(k \text{ out of } N \text{ sensors decide } +1|H_1) = A_{N-k,k}(-1,\rho),$$
$$P(U|H_0) = P(k \text{ out of } N \text{ sensors decide } +1|H_0) = A_{N-k,k}(+1,\rho). \tag{4.37}$$

To study the effect of N and ρ on the the error probability, consider the case that $P(H_0) = P(H_1)$=0.5 and σ^2=1 (i.e. SNR=0 dB). Thus, the fusion rule is simplified to

$$\Lambda(U) = \frac{P(U|H_1)}{P(U|H_0)} = \frac{P(k \text{ out of } N \text{ sensors decide } +1|H_1)}{P(k \text{ out of } N \text{ sensors decide } +1|H_0)} = \frac{A_{N-k,k}(-1,\rho)}{A_{N-k,k}(+1,\rho)} \underset{H_0}{\overset{H_1}{\gtrless}} 1, \tag{4.38}$$

where $A_{N-k,k}(-1,\rho)$ and $A_{N-k,k}(+1,\rho)$ can be computed numerically using Eqs. (4.34), (4.35) and (4.36). From this decision equation, there exists a K for given N and ρ such that:

$$\text{when } k \geq K, \quad \frac{A_{N-k,k}(-1,\rho)}{A_{N-k,k}(+1,\rho)} \geq 1; \tag{4.39}$$

$$\text{when } k < K, \quad \frac{A_{N-k,k}(-1,\rho)}{A_{N-k,k}(+1,\rho)} < 1. \tag{4.40}$$

In this case, the error probability defined in Eq. (4.27) can be expressed as

$$P_e = 0.5\left[\sum_{k=0}^{K-1} A_{N-k,k}(-1,\rho) + \sum_{k=K}^{N} A_{N-k,k}(+1,\rho)\right]. \tag{4.41}$$

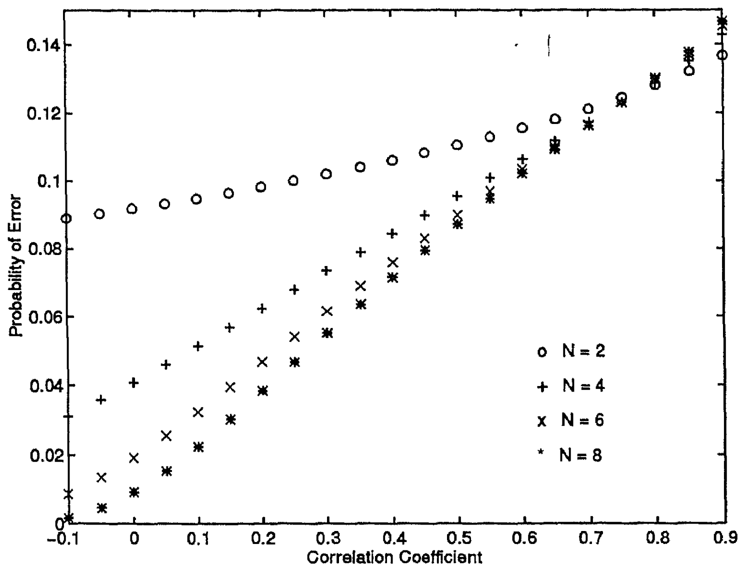

Figure 4.2 P_e versus ρ, for N = 2, 4, 6, 8.

For this special case, the error probability using the optimal decision and adaptive decision rules can be determined. Figures 4.2 and 4.3 show the error probability versus ρ based on Eq. (4.41) when N is even and odd, respectively. When N is even, there exists a k such that $\Lambda(u)$=1, but when N is odd, no such k exists. $\Lambda(U)$=1 corresponds to an undetermined case which can be considered as either H_0 or H_1. The contribution to the error probability for the undetermined case is considered as half of the probability it occurs.

It can be seen that better performance is achieved with smaller correlation coefficient between sensors. This agrees with the conclusion of other fusion rules [1, 18]. Also, better performance can be achieved by increasing the number of sensors, but this advantage diminishes as the correlation coefficient ρ increases.

4.5 Simulations

Figure 4.4 shows the set-up for our simulations.

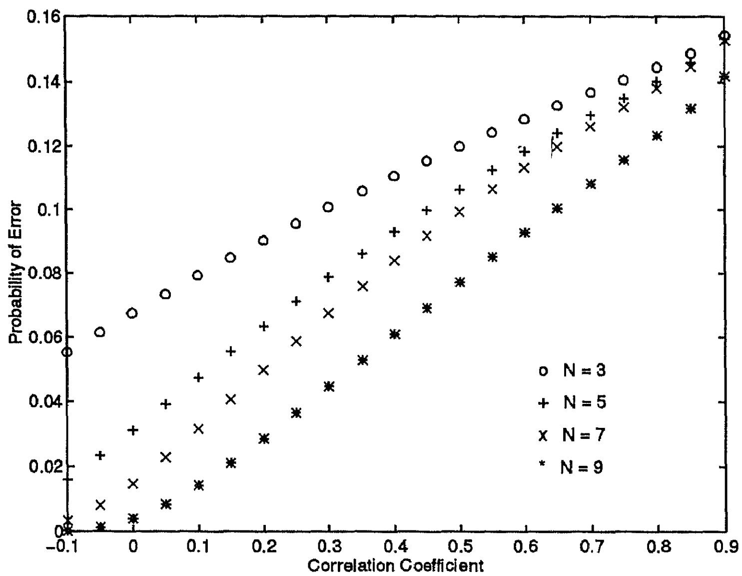

Figure 4.3 P_e versus ρ, for N = 3, 5, 7, 9.

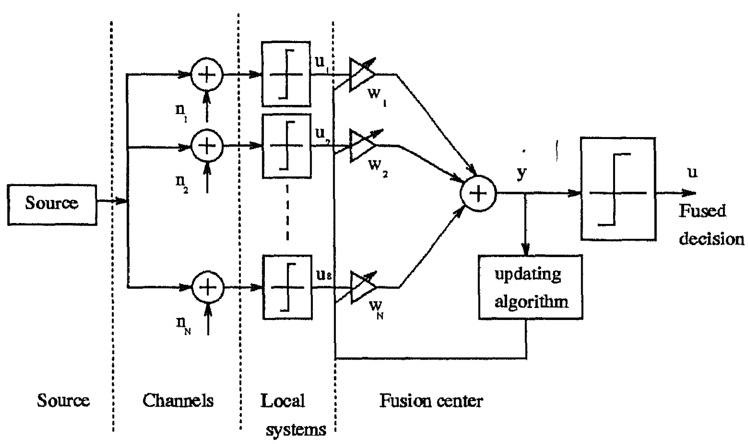

Figure 4.4 The set-up of computer simulations.

The source emits a sequence of +1 and -1. The probability of emitting +1 is $P(H_1)$, and that of emitting -1 is $P(H_0)$. The additive noise is zero mean Gaussian with a variance of 1. Each local sensor makes its decision u_i and transmits it to the fusion center. The fusion center computes the linear combination of local decisions to produce $\mathbf{y}$, and then compares it with a threshold (here, zero is used). If $\mathbf{y}$ is greater than 0, the final decision is +1, otherwise, -1.

4.5.1 Generation of Correlated Gaussian Noise

In our simulations, we need to generate Gaussian noise with the specified correlation coefficient. The usual random number generator can only produce independent and identically distributed noise. Correlated noise can be obtained through some linear transformations. Let Z denote an N-dimensional correlated noise vector whose correlation matrix is C_z. Y is another N-dimensional noise vector defined by

$$Y = A^T Z. \tag{4.42}$$

If A is a square matrix whose column vectors are the eigenvectors of C_z, the correlation matrix C_y of Y becomes a diagonal matrix whose diagonal elements are the eigenvalue of C_z [23]. Denote C_y as

$$C_y = \begin{bmatrix} \lambda_1 & 0 & 0 & \ldots\ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots\ldots & 0 \\ \ldots & \ldots & \ldots & \ldots\ldots & \\ 0 & 0 & \ldots & \ldots\ldots & \lambda_N \end{bmatrix},$$

where λ_i is an eigenvalue of C_z. Since λ_i are distinct even when each element in Z has the same variance and the same correlation coefficient, by introducing the next transformation

$$X = B^T Y, \tag{4.43}$$

where

$$B = \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & 0 & 0 & \ldots\ldots & 0 \\ 0 & \frac{1}{\sqrt{\lambda_2}} & 0 & \ldots\ldots & 0 \\ \ldots & \ldots & \ldots & \ldots\ldots & \\ 0 & 0 & \ldots & \ldots\ldots & \frac{1}{\sqrt{\lambda_N}} \end{bmatrix},$$

the correlation matrix C_x of X becomes an identity matrix. Combining the above two transformations, we have

$$Z = (BA^T)^{-1} X. \tag{4.44}$$

According to the above definitions, X is an independent zero mean, unit variance Gaussian random vector which can be generated easily. Z becomes a zero mean Gaussian random vector with correlation matrix C_z. B and A are determined by the eigenvalues and eigenvectors of C_z. In our experiment, C_z is specified to have the same correlation coefficient and variance. MATLAB software is employed to implement the noise generation and eigenanalysis.

4.5.2 Simulation Results

Consider the same situation described in Section 4.4. Theoretical analysis has shown that the number of sensors and correlation coefficient greatly affect the performance of fusion (see Figures 4.2 and 4.3). These effects are also observed in the simulation results. Figure 4.5 shows the plots of error probability versus the iteration for different N with a fixed correlation coefficient. Figure 4.6 shows the plot for different correlation coefficient with a fixed N. Tables 4.1 and 4.2 summarize the corresponding theoretical and simulated values at the 400th iteration of the error probability.

From these two figures and tables, it can be seen that the proposed algorithm converges, and the steady state error probabilities obtained from the simulations are very close to the theoretical values.

It is interesting to note, as illustrated by Figure 4.7 and summarized in Table 4.3, that the adaptive algorithm developed for the correlated case always outperforms the one we previously developed for the independent case [4, 3] regardless of whether or not the local decisions are actually correlated. This may be attributed to the fact

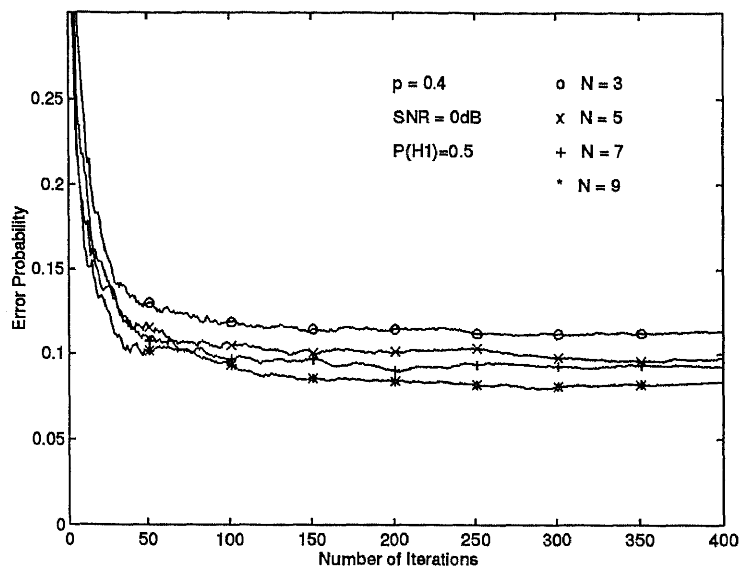

Figure 4.5 The effect of the number of sensors on error probability.

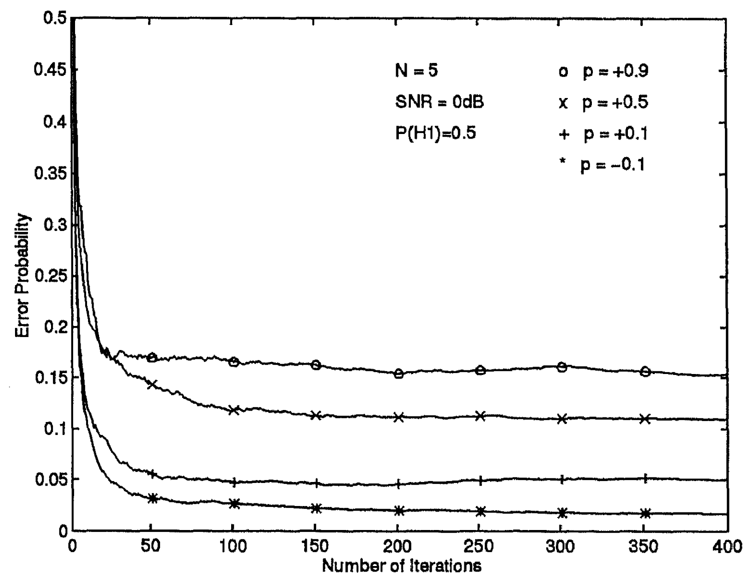

Figure 4.6 The effect of the correlation coefficient on error probability.

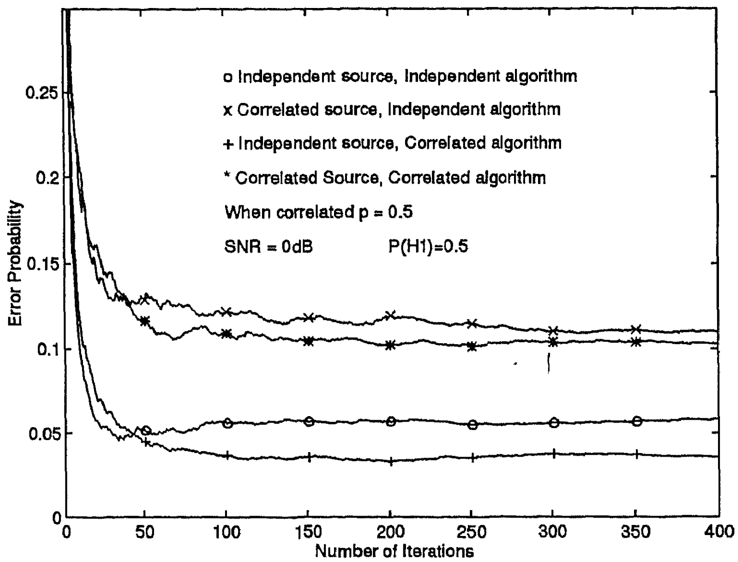

Figure 4.7 Performance comparison between the correlated and independent methods.

Table 4.1 The theoretical and estimated values of error probability (ρ=0.4).

N	3	5	7	9
Experimental Value	0.1130	0.0971	0.0923	0.0832
Theoretical Value	0.1107	0.0930	0.0840	0.0610

Table 4.2 The theoretical and estimated values of error probability (N=5).

ρ	-0.1	0.1	0.5	0.9
Experimental Value	0.0173	0.0500	0.1092	0.1457
Theoretical Value	0.0160	0.0473	0.1062	0.1533

Table 4.3 The estimated error probabilities for different algorithms and different sources.

	Independent Algorithm	Correlated Algorithm
Independent Source	0.0584	0.0358
Correlated Source	0.1100	0.1029

that the algorithm that considers correlated decisions includes more information in making its decision.

4.6 Summary

In this chapter, we have proposed an adaptive algorithm to solve the MAP-based optimal fusion problem when sensors are dependent from one another. The following main attributes of the algorithm can be concluded from the theoretical analysis and simulations.

1) It does not require *a priori* knowledge about the sensors and source, and thus is more practical.

2) It adapts the weights from time to time, and thus is suitable for a time-varying environment.

3) In some cases, it behaves as well as the optimal rule.

4) Its computational complexity is low, and thus implementable.

CHAPTER 5

PERFORMANCE COMPARISON OF FUSION RULES IN DISTRIBUTED DETECTION

In this chapter, the performance of logical AND and OR, majority and optimal fusion rules in both independent and correlated Gaussian noise is analyzed and compared in terms of their Receiver Operating Characteristics (ROCs). Various factors that affect the fusion performance are considered in the analysis. By varying the local decision thresholds, the ROCs under the influence of the number of sensors, signal-to-noise ratio (SNR), the deviation of local decision probabilities and correlation coefficient, are computed and plotted, respectively. Several interesting and key observations on the performance of fusion rules are drawn from the analysis.

5.1 Fusion Rules in Independent Noise

Consider the situation where there are N sensors. To avoid ambiguity that can happen in the MAJ rule, N is chosen to be odd. Each sensor receives an observation x_i and makes a decision u_i, $i = 1, 2, \cdots, N$. Note that

$$x_i = s + n_i, \tag{5.1}$$

where s$\in \{+1, -1\}$ is the signal component, and $n_i \in N(0, \sigma_i)$, denoting a zero-mean Gaussian random variable with a standard deviation of σ_i. H_0 and H_1 are used to denote the following two events

$$\begin{aligned} H_0 : &\quad \text{s} = \text{-1}; \\ H_1 : &\quad \text{s} = +1. \end{aligned}$$

Each local decision, u_i, is a binary random variable defined by

$$u_i = \begin{cases} -1 & \text{if } H_0 \text{ is declared at the } i\text{th sensor,} \\ +1 & \text{if } H_1 \text{ is declared at the } i\text{th sensor.} \end{cases}$$

Let t be the same threshold used by each local sensor in making its decision:

$$u_i = \begin{cases} +1 & \text{if } x_i \geq t, \\ -1 & \text{if } x_i < t. \end{cases}$$

These local decisions are sent to a fusion center that adopts one of the aforementioned fusion rules to make a final decision u_f. With the above assumptions, the following four quantities can be readily obtained.

$$\begin{aligned} P(u_i = +1|H_1) &= 1 - Q(\frac{t-1}{\sigma_i}), \quad P(u_i = -1|H_1) = Q(\tfrac{t-1}{\sigma_i}), \\ P(u_i = +1|H_0) &= 1 - Q(\frac{t+1}{\sigma_i}), \quad P(u_i = -1|H_0) = Q(\tfrac{t+1}{\sigma_i}), \end{aligned} \tag{5.2}$$

where $Q(\cdot)$ is the cumulative distribution function of a unit normal (Gaussian) random variable. These four terms are useful in deriving the probability of a detection P_D and a false alarm P_f of the aforementioned fusion rules in the independent case.

5.1.1 AND Rule

The "AND" fusion rule is defined by

$$u_f = \begin{cases} +1 & \text{if } N^+ = N, \\ -1 & \text{otherwise,} \end{cases}$$

where N^+ is the number of local decisions that are positive, i.e., $u_i = +1$. According to the above fusion rule, its probability of a detection is

$$\begin{aligned} P_D &= P(u_f = +1|H_1) \\ &= P(u_1 = +1, u_2 = +1, \cdots, u_N = +1|H_1) \\ &= P(u_1 = +1|H_1)P(u_2 = +1|H_1)\cdots P(u_N = +1|H_1) \\ &= [1 - Q(\frac{t-1}{\sigma_1})][1 - Q(\frac{t-1}{\sigma_2})]\cdots[1 - Q(\frac{t-1}{\sigma_N})] \\ &= \prod_{i=1}^{N}[1 - Q(\frac{t-1}{\sigma_i})]. \end{aligned} \tag{5.3}$$

The probability of a false alarm will be

$$P_f = P(u_f = +1|H_0)$$

$$\begin{aligned}
&= P(u_1 = +1, u_2 = +1, \cdots, u_N = +1|H_0) \\
&= P(u_1 = +1|H_0)P(u_2 = +1|H_0)\cdots P(u_N = +1|H_0) \\
&= [1 - Q(\frac{t+1}{\sigma_1})][1 - Q(\frac{t+1}{\sigma_2})]\cdots[1 - Q(\frac{t+1}{\sigma_N})] \\
&= \prod_{i=1}^{N}[1 - Q(\frac{t+1}{\sigma_i})]. \qquad (5.4)
\end{aligned}$$

For given σ_i and N, by varying t from $+\infty$ to $-\infty$, the ROC curve (P_D versus P_f) can be obtained.

5.1.2 OR Rule

The "OR" fusion rule is defined by

$$u_f = \begin{cases} +1 & \text{if } N^+ \geq 1, \\ -1 & \text{otherwise.} \end{cases}$$

Thus, the probability of a detection is

$$\begin{aligned}
P_D &= 1 - P(u_1 = -1, u_2 = -1, \cdots, u_N = -1|H_1) \\
&= 1 - \prod_{i=1}^{N} Q(\frac{t-1}{\sigma_i}). \qquad (5.5)
\end{aligned}$$

The probability of a false alarm is

$$\begin{aligned}
P_f &= 1 - P(u_1 = -1, u_2 = -1, \cdots, u_N = -1|H_0) \\
&= 1 - \prod_{i=1}^{N} Q(\frac{t+1}{\sigma_i}). \qquad (5.6)
\end{aligned}$$

Similarly, we can plot its ROC.

5.1.3 MAJ Rule

The MAJ rule is defined by

$$u_f = \begin{cases} +1 & \text{if } N^+ \geq \frac{N+1}{2}, \\ -1 & \text{otherwise.} \end{cases}$$

Thus, its P_D is

$$\begin{aligned} P_D &= P(\ N^+ \geq \tfrac{N+1}{2} | H_1) \\ &= \sum_{k=\frac{N+1}{2}}^{N} \sum_{i=1}^{C(N,k)} P_{i_1} P_{i_2} \cdots P_{i_k} (1 - P_{i_{k+1}})(1 - P_{i_{k+2}}) \cdots (1 - P_{i_N}), \end{aligned}$$

where $i_j \in \{1, 2, \cdots, N\}$ is the index of a sensor, $C(N,k) = \begin{pmatrix} N \\ k \end{pmatrix}$, $\{i_1, i_2, \cdots, i_k, \cdots, i_N\}$ is one out of the $C(N,k)$ possible combinations, and $P_{i_j} = P(u_{i_j} = +1 | H_1) = 1 - Q(\frac{t-1}{\sigma_{i_j}})$. Thus,

$$\begin{aligned} P_D &= \sum_{k=\frac{N+1}{2}}^{N} \sum_{i=1}^{C(N,k)} [1 - Q(\frac{t-1}{\sigma_{i_1}}][1 - Q(\frac{t-1}{\sigma_{i_2}}] \cdots [1 - Q(\frac{t-1}{\sigma_{i_k}}]Q(\frac{t-1}{\sigma_{i_{k+1}}})Q(\frac{t-1}{\sigma_{i_{k+2}}}) \cdots Q(\frac{t-1}{\sigma_{i_N}}) \\ &= \sum_{k=\frac{N+1}{2}}^{N} \sum_{i=1}^{C(N,k)} \prod_{j=1}^{k} [1 - Q(\frac{t-1}{\sigma_{i_j}})] \prod_{j=k+1}^{N} Q(\frac{t-1}{\sigma_{i_j}}). \end{aligned} \tag{5.7}$$

With similar reasoning, we have

$$\begin{aligned} P_f &= P(N^+ \geq \tfrac{N+1}{2} | H_0) \\ &= \sum_{k=\frac{N+1}{2}}^{N} \sum_{i=1}^{C(N,k)} \prod_{j=1}^{k} [1 - Q(\frac{t+1}{\sigma_{i_j}})] \prod_{j=k+1}^{N} Q(\frac{t+1}{\sigma_{i_j}}). \end{aligned} \tag{5.8}$$

5.1.4 OPT Rule

Let:

$$\Lambda(u_1, u_2, \cdots, u_N) = \frac{P(u_1, u_2, \cdots, u_N | H_1)}{P(u_1, u_2, \cdots, u_N | H_0)}. \tag{5.9}$$

The optimum fusion rule is:

$$u_f = \begin{cases} +1 & \text{if } \Lambda(u_1, u_2, \cdots, u_N) \geq \frac{P(H_0)}{P(H_1)} \ , \\ -1 & \text{if } \Lambda(u_1, u_2, \cdots, u_N) < \frac{P(H_0)}{P(H_1)} \ . \end{cases}$$

Denote $U = (u_1, u_2, \cdots, u_N)$ as the local decision vector, and define the set $U^+ = \{U : \Lambda(U) \geq \frac{P(H_0)}{P(H_1)}\}$, and U_i is a member of U^+. Let the number of $+1$ in U_i be M_i, and the number of -1 in U_i be N_i. We have

$$\begin{aligned} P_D &= P(U \in U^+|H_1) \\ &= \sum_{U_i \in U^+} \prod_{j=1}^{M_i} [1 - Q(\frac{t-1}{\sigma_{i_j}})] \prod_{j=M_{i+1}}^{N_i} Q(\frac{t-1}{\sigma_{i_j}}). \end{aligned} \tag{5.10}$$

Similarly,

$$\begin{aligned} P_f &= P(U \in U^+|H_0) \\ &= \sum_{U_i \in U^+} \prod_{j=1}^{M_i} [1 - Q(\frac{t+1}{\sigma_{i_j}})] \prod_{j=M_{i+1}}^{N_i} Q(\frac{t+1}{\sigma_{i_j}}). \end{aligned} \tag{5.11}$$

Without lose of generality, it is assumed that $P(H_1) = P(H_0)$ in the following analysis.

5.2 Performance Comparison in Independent Noise

When the local decisions are independent, factors that affect the detection performance include SNR, the number of sensors, and the deviation of local decision probabilities (the false-alarm and detection probabilities). In this section, the effect of each factor on the ROC is studied, respectively.

5.2.1 The Effect of the SNR

As specified in Section 2, the power of the signal is fixed to be 1. The SNR depends on only the power of noise (the variance of the noise σ_i^2). Figure 5.1 shows the plot of ROC for different fusion rules based on Eqs. (5.3)-(5.8), (5.10) and (5.11) with different SNRs, when N is 5 and the variances of all sensors are the same.

In Figure 5.1, there are three groups of curves which correspond to three different SNRs. In each group, there are four curves corresponding to four different fusion rules. The probabilities of detection and false alarm using AND, OR and MAJ rules increase monotonically as the local decision threshold t decreases. The probabilities of detection and false alarm of the OPT rule, however, are not a simple

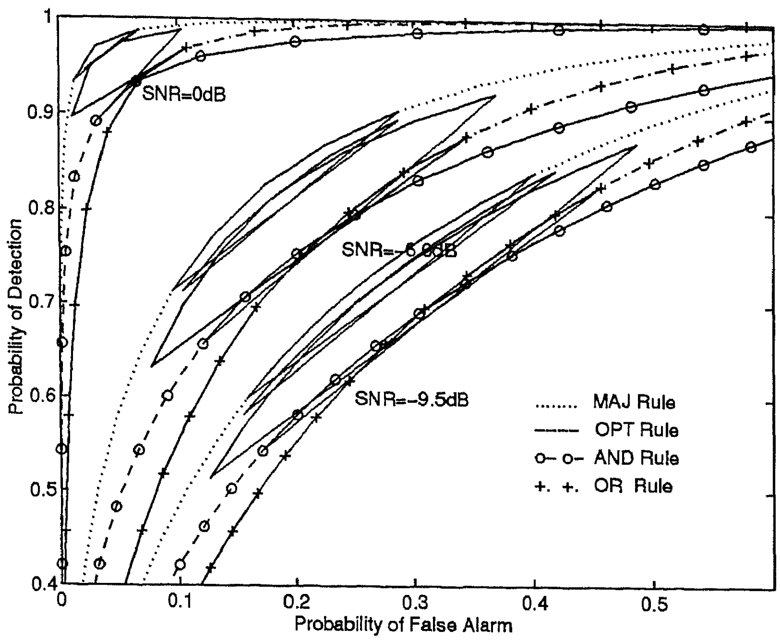

Figure 5.1 The ROC for the independent case with different SNR when N=5.

function of t. Its ROC [1] performance achieves the best at some given t (which corresponds to the optimal local threshold; here it is zero). Its performance deteriorates elsewhere. For some t, it yields the worst performance among the four fusion rules. Among the four rules, the MAJ rule always achieves the best performance. The AND rule has better performance than the OR rule in a low probability of detection region. The OR rule outperforms the AND rule in a high probability of detection region. This is consistent with the conclusion of [38]. The SNR affects the performance of all the fusion rules. Generally speaking, the larger the SNR, the better the performance achievable and the smaller the range of t in which the OPT rule achieves the same performance as the MAJ rule.

[1]Here, the OPT rule is optimum in the sense that it always minimizes the error probability for a given set of local thresholds. For the equal probable case, the error probability corresponds to one half of the sum of the probability of a miss and the probability of a false alarm. See Section 5.4.4 for explanation of the "zig-zag" behavior of the ROC curves for the OPT rule.

5.2.2 The Effect of the Number of Sensors

To illustrate the effect of the number of sensors on the ROC, the SNR is fixed at -6.0dB, while the number of sensors varies. Figure 5.2 shows the plot of ROCs for different N. In order to show the performance clearly, the ROCs corresponding to different N are shown in different diagrams. From Figure 5.2, the effect can be summarized below:

1. The larger the number, the better the performance.

2. The larger the number, the larger the difference in performance of different fusion rules.

3. The larger the number, the more sensitive the OPT rule is to the local threshold t.

5.2.3 The Effect of the Deviation of Local Decision Probabilities

The deviation of local decision probabilities is referred to as the dynamic range of detection and false alarm probabilities among different sensors. In the methods studied here, it can be reflected by the differences of noise variance among sensors. Let B=$(\sigma_1, \sigma_2, \cdots, \sigma_N)$ be the vector denoting the standard deviation of the additive Gaussian noise at the N local sensors. Define the mean M and the standard deviation D of B by

$$\begin{aligned} M &= \sum_{i=1}^{N} \sigma_i, \\ D &= \frac{1}{N}\sqrt{\sum_{i=1}^{N}(\sigma_i - M)^2}. \end{aligned} \tag{5.12}$$

D is used as a measure of deviation of local decision probabilities. Figure 5.3 shows the ROC plots for different B and D but the same M, when N=5.

As D increases, the performance of the MAJ rule gets better in the overall region. The OPT rule outperforms all the other rules only at some special range of t. The performance of AND and OR rules get better in some regions and worse

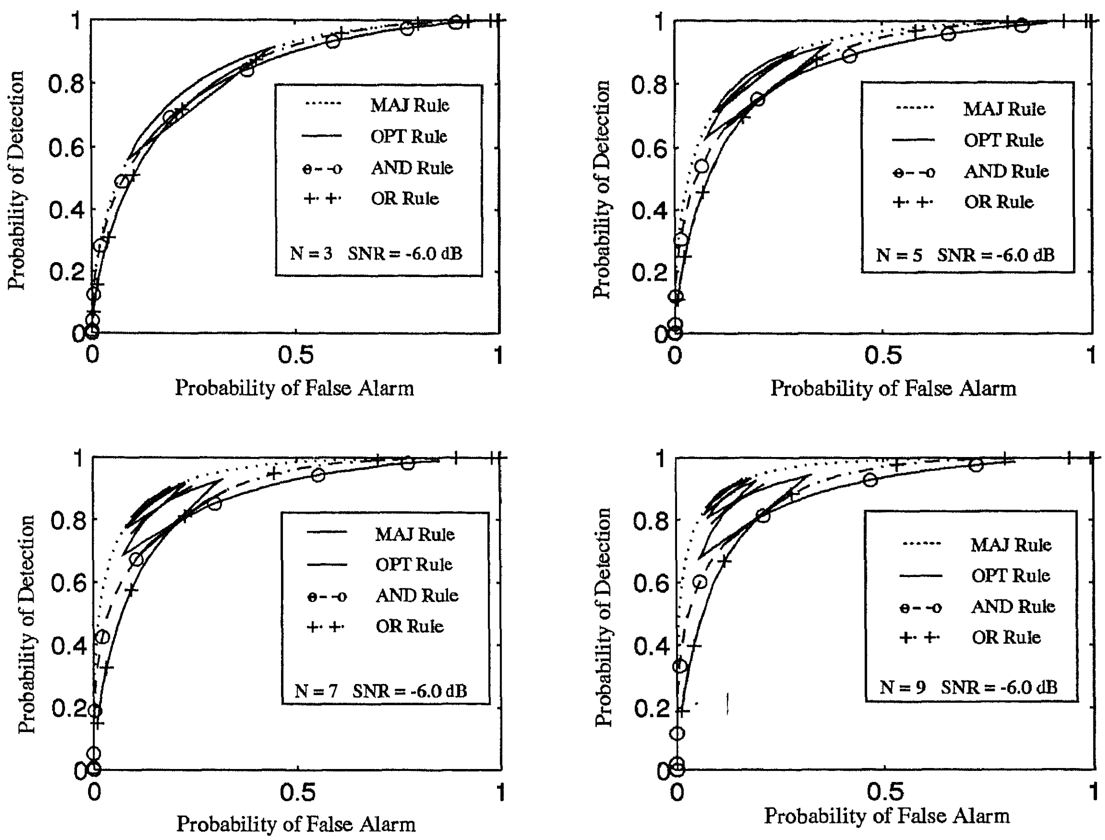

Figure 5.2 The ROC for the independent case with different N when SNR=-6.0dB.

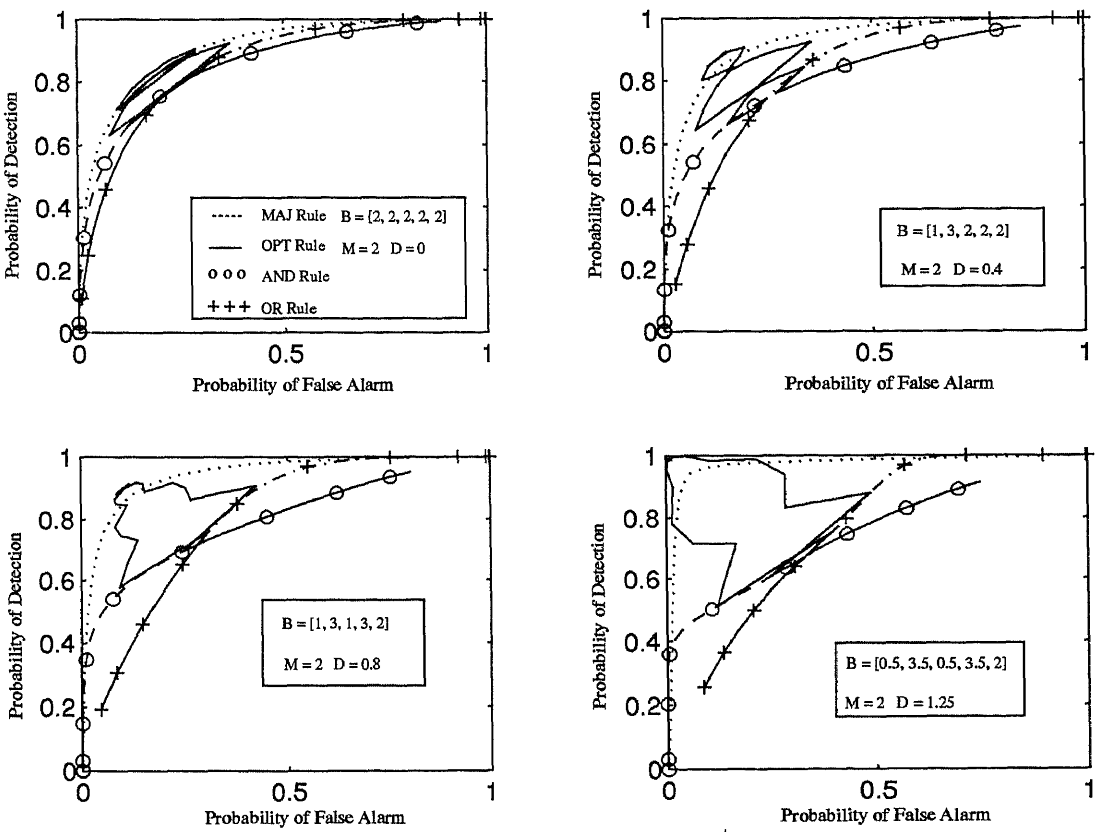

Figure 5.3 The ROC for the independent case with different D when N=5.

in others. In some regions, either the AND or the OR rule does better than the MAJ rule. The larger the D is, the larger the range of t in which OPT performs the best, and the greater the advantage of OPT over MAJ at optimal t is. Regardless of the SNR, D and N, the AND and OR rules have the same performance when t is optimal.

5.3 Fusion Rules in Correlated Noise

In the correlated case, besides the number of sensors, the SNR, and deviation of variance, the correlation coefficient plays a major role in affecting the performance of fusion. Since it is computationally too expensive to consider the general correlation case, we focus our attention on the case with equally correlated and identical sensors. In this case, the correlation coefficient ρ is the same for all sensors and each sensor has identical properties (i.e., identical noise variance σ^2 and a zero mean).

5.3.1 AND Rule

The probability of a detection becomes

$$\begin{aligned} P_D &= P(u_1 = +1, u_2 = +1, \cdots, u_N = +1 | H_1) \\ &= P(x_1 > t, x_2 > t, \cdots, x_N > t | H_1). \end{aligned}$$

Since the program available for computing probability in the correlated case [19] can only calculate the probability of $P(x_1 < t, x_2 < t, \cdots, x_N < t|H_1)$, P_D can be computed by the following equation [11].

$$P_D = \sum_{j=0}^{N} (-1)^j \binom{N}{j} P(x_1 < t, x_2 < t, \cdots, x_j < t|H_1), \tag{5.13}$$

where $P(x_1 < t, x_2 < t, \cdots, x_j < t|H_1)$ is the joint probability of j-dimensional, equally correlated, identical Gaussian-distributed random variables with a mean of $+1$, a variance of σ^2, a correlation coefficient of ρ, and an integral range $(-\infty, t)$.

Similarly,

$$\begin{aligned} P_f &= P(u_1 = +1, u_2 = +1, \cdots, u_N = +1|H_0) \\ &= P(x_1 > t, x_2 > t, \cdots, x_N > t|H_0) \\ &= \sum_{j=0}^{N} (-1)^j \binom{N}{j} P(x_1 < t, x_2 < t, \cdots, x_j < t|H_0), \end{aligned} \tag{5.14}$$

where $P(x_1 < t, x_2 < t, \cdots, x_j < t|H_0)$ is the same as $P(x_1 < t, x_2 < t, \cdots, x_j < t|H_1)$, except with a mean of -1.

5.3.2 OR Rule

The probability of a detection is

$$\begin{aligned} P_D &= 1 - P(u_1 = -1, u_2 = -1, \cdots, u_N = -1|H_1) \\ &= 1 - P(x_1 < t, x_2 < t, \cdots, x_N < t|H_1). \end{aligned} \tag{5.15}$$

The probability of a false alarm is

$$\begin{aligned} P_f &= 1 - P(u_1 = -1, u_2 = -1, \cdots, u_N = -1|H_0) \\ &= 1 - P(x_1 < t, x_2 < t, \cdots, x_N < t|H_0). \end{aligned} \tag{5.16}$$

5.3.3 MAJ Rule

$$\begin{aligned} P_D &= P(\ N^+ \geq \tfrac{N+1}{2}|H_1\ ) \\ &= \sum_{k=\frac{N+1}{2}}^{N} \binom{N}{k} P(x_1 < t, x_2 < t, \cdots, x_k < t, x_{k+1} > t, x_{k+2} > t, \cdots, x_N > t|H_1) \\ &= \sum_{k=\frac{N+1}{2}}^{N} \binom{N}{k} \sum_{j=0}^{k} (-1)^j \binom{k}{j} P(x_1 < t, x_2 < t, \cdots, x_{N-k+j} < t|H_1), \end{aligned} \tag{5.17}$$

$$P_f = \sum_{k=\frac{N+1}{2}}^{N} \binom{N}{k} \sum_{j=0}^{k} (-1)^j \binom{k}{j} P(x_1 < t, x_2 < t, \cdots, x_{N-k+j} < t|H_0). \tag{5.18}$$

5.3.4 OPT Rule

$$P_D = P\left(U \in U^+ | H_1\right).$$

Since every sensor is identical, and $\Lambda(U) = \frac{P(U|H_1)}{P(U|H_0)}$, $\Lambda(U)$ is only a function of k, where k is the number of sensors which yield $+1$, ranging from 0 to N. $\Lambda(U)$ can be computed numerically by the following equation.

$$\begin{aligned}\Lambda(U) &= \frac{P(\text{ k out of N sensors decide} + 1|H_1)}{P(\text{ k out of N sensors decide} + 1|H_0)} \\ &= \frac{\begin{pmatrix} N \\ k \end{pmatrix} \sum_{j=0}^{k}(-1)^j \begin{pmatrix} k \\ j \end{pmatrix} P(x_1 < t, x_2 < t, \cdots, x_{N-k+j} < t|H_1)}{\begin{pmatrix} N \\ k \end{pmatrix} \sum_{j=0}^{k}(-1)^j \begin{pmatrix} k \\ j \end{pmatrix} P(x_1 < t, x_2 < t, \cdots, x_{N-k+j} < t|H_0)} \qquad (5.19)\end{aligned}$$

Denote:

$$\begin{aligned} k^+ &= \left\{k : \Lambda(U) \geq \frac{P(H_0)}{P(H_1)}\right\}, \\ k^- &= \left\{k : \Lambda(U) < \frac{P(H_0)}{P(H_1)}\right\}. \end{aligned}$$

We have

$$P_D = \sum_{k \in k^+} \begin{pmatrix} N \\ k \end{pmatrix} \sum_{j=0}^{k}(-1)^j \begin{pmatrix} k \\ j \end{pmatrix} P(x_1 < t, x_2 < t, \cdots, x_{N-k+j} < t|H_1) \qquad (5.20)$$

$$P_f = \sum_{k \in k^+} \begin{pmatrix} N \\ k \end{pmatrix} \sum_{j=0}^{k}(-1)^j \begin{pmatrix} k \\ j \end{pmatrix} P(x_1 < t, x_2 < t, \cdots, x_{N-k+j} < t|H_0) \qquad (5.21)$$

For different n, σ and ρ, Eqs. (5.13)-(5.18), (5.20), and (5.21) can be computed numerically by using the multivariate normal integrals [19].

5.4 Performance Comparison in Correlated Noise

For the correlated noise, according to the assumptions made in Section 4, the factors that affect the performance of fusion are SNR, the number of sensors, and the correlation coefficient. Although the deviation of local decision probabilities also affects the performance, its discussion is beyond the scope of equally correlated and identical

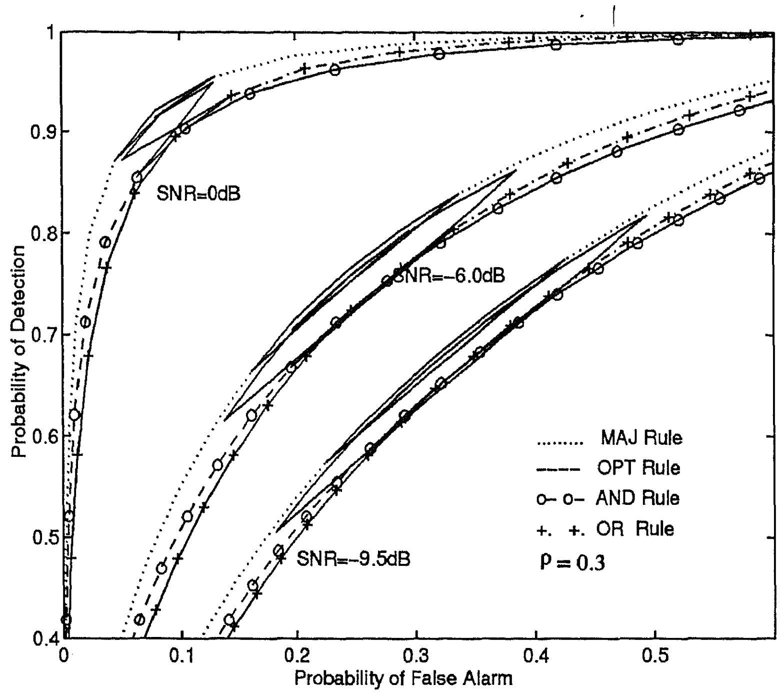

Figure 5.4 The ROC for the correlated case with different SNR when N=5.

sensors. The following discussion provides the analysis on the effect of SNR, number of sensors, and correlation coefficient on ROC.

5.4.1 The Effect of the SNR

Figure 5.4 shows the ROC for different SNR when ρ=0.3 and N=5. Comparing Figure 5.4 with Figure 5.1, it can be seen that the effect of SNR on ROC in the correlated case is similar to that in the independent case. The difference between them is that the performance advantage of one rule over another in the correlated case is less significant than that in the independent case.

5.4.2 The Effect of the Number of Sensors

Figure 5.5 shows the ROC for different N when ρ=0.3 and SNR=-6.0dB. The effect is similar to the independent case, but the differences are much less significant.

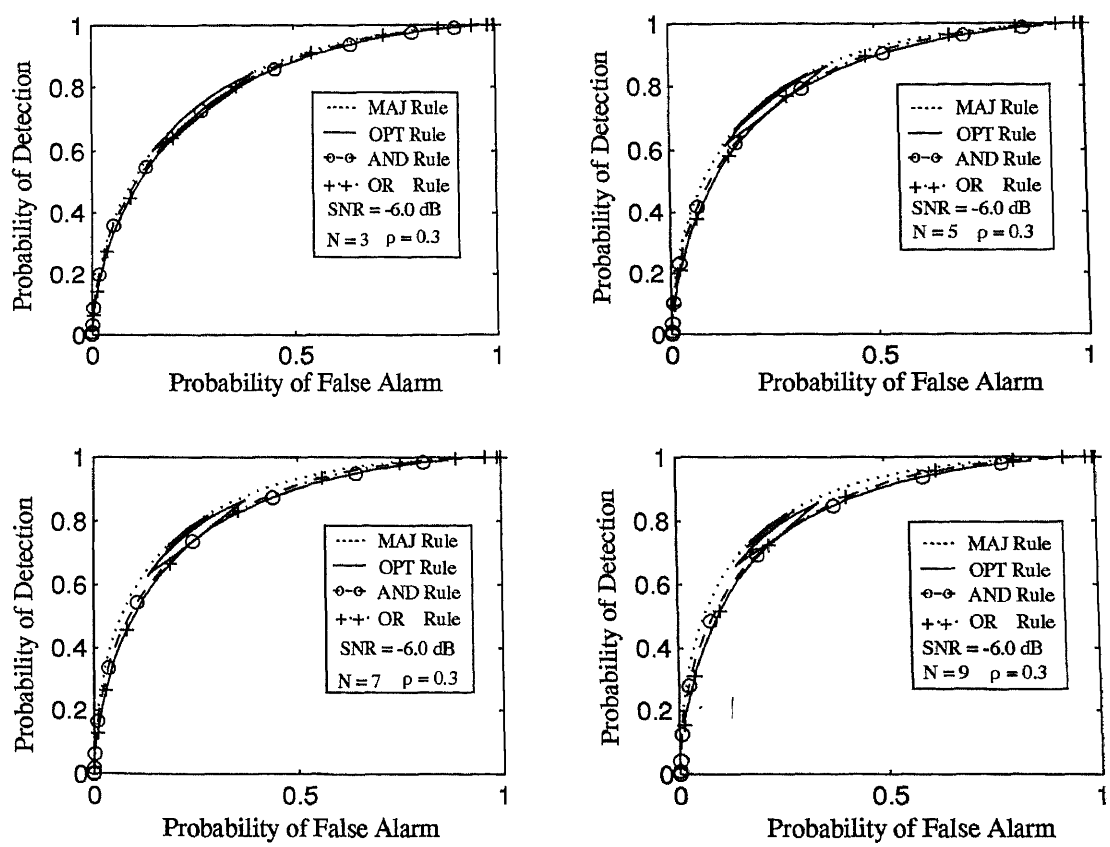

Figure 5.5 The ROC for the correlated case with different N when SNR=-6.0dB.

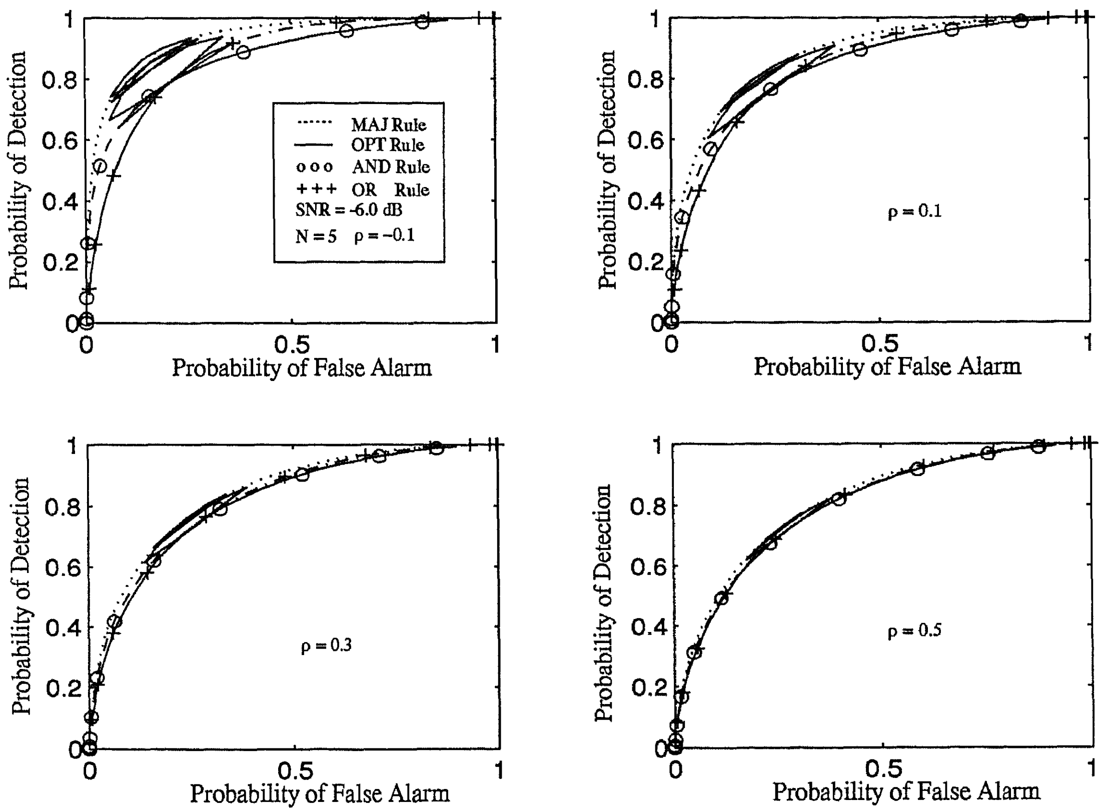

Figure 5.6 The ROC for the correlated case with different ρ when N=5.

5.4.3 The Effect of the Correlation Coefficient

The effect of the correlation coefficient is illustrated in Figure 5.6. The ROCs are divided into four groups. Each group corresponds to a different ρ, but the same N and SNR. Figure 5.6 shows that the larger the ρ, the worse the performance for all fusion rules. This agrees with what we have observed before [11]. In addition, as the correlation coefficient increases, the four fusion rules eventually come to have the same performance. When $\rho = 0.5$, they have almost the same performance.

5.4.4 The Effect of the Local Threshold

All the ROC curves are obtained by varying the local threshold t from +5 to -5. The behaviour of the AND, OR, and MAJ rules is quite predictable. The effect of the local threshold on the OPT rule is more complicated. For illustration, consider only three identical and independent sensors, in which case, the OPT rule [7] is

$$(\frac{P_D}{P_f})^i(\frac{1-P_D}{1-P_f})^{3-i} \underset{H_0}{\overset{H_1}{\gtrless}} 1, \tag{5.22}$$

where i=0,1,2,3 is the the number of local sensors that declare H_1. There exist eight local decision vectors: $U_1 = (1,1,1)$, $U_2 = (1,1,-1)$, $U_3 = (1,-1,1)$, $U_4 = (1,-1,-1)$, $U_5 = (-1,1,1)$, $U_6 = (-1,1,-1)$, $U_7 = (-1,-1,1)$ and $U_8 = (-1,-1,-1)$. When $t = 5$, the OPT rule first divides them into two classes: $D_1 = \{U_1, U_2, U_3, U_4, U_5, U_6, U_7\}$ and $D_0 = \{U_8\}$. Both P_D and P_f are calculated based on the components of D_1. When t decreases, both P_D and P_f increase but the sufficient statistics of Eq. (5.22) decrease. When t is small enough such that $(\frac{P_D}{P_f})(\frac{1-P_D}{1-P_f})^2 < 1$, the classification becomes$D_1 = \{U_1, U_2, U_3, U_5\}$ and $D_0 = \{U_4, U_6, U_7, U_8\}$, in which case, both P_D and P_f decrease abruptly. When t decreases again, P_D and P_f begin to increase. When t is so small that $(\frac{P_D}{P_f})^2(\frac{1-P_D}{1-P_f}) < 1$, another "drop" occurs, and the classification becomes $D_1 = \{U_1\}$

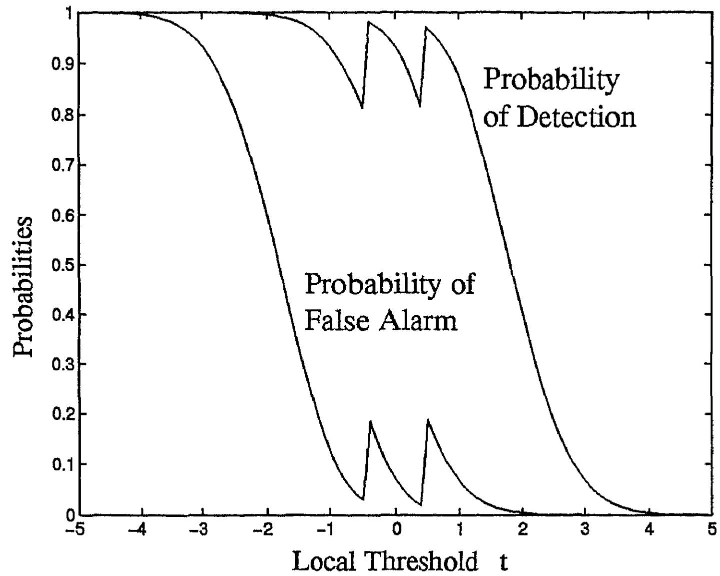

Figure 5.7 The effect of local threshold.

and $D_0 = \{U_2, U_3, U_4, U_5, U_6, U_7, U_8\}$. Since $(\frac{P_D}{P_f})^3 > 1$ for any t, any additional increase in t will not change the membership of classification further. Thus, P_D and P_f increase smoothly. The dynamic process is illustrated in Figure 5.7. Therefore, probabilities of a false alarm and a detection for the OPT rule are not monotonic functions of the local decision threshold. As a result, the ROC for the OPT rule is no longer concave downward as observed in Figures 5.1-5.6.

There is another unusual phenomenon about the ROC curves. Although the OPT rule achieves the best performance among all fusion rules for a given local threshold, the ROC curves of the OPT rule are sometimes below those of the other three fusion rules. Figure 5.8 shows the probabilities of a false alarm, a miss, detection, and error as a function of the local threshold for both the MAJ and OPT rules. In the sense of minimum error probability, the OPT rule is definitely optimum. As shown in Figure 5.8, for the same probability of a false alarm, different fusion rules use different local thresholds. This means that in the ROC plots (which are

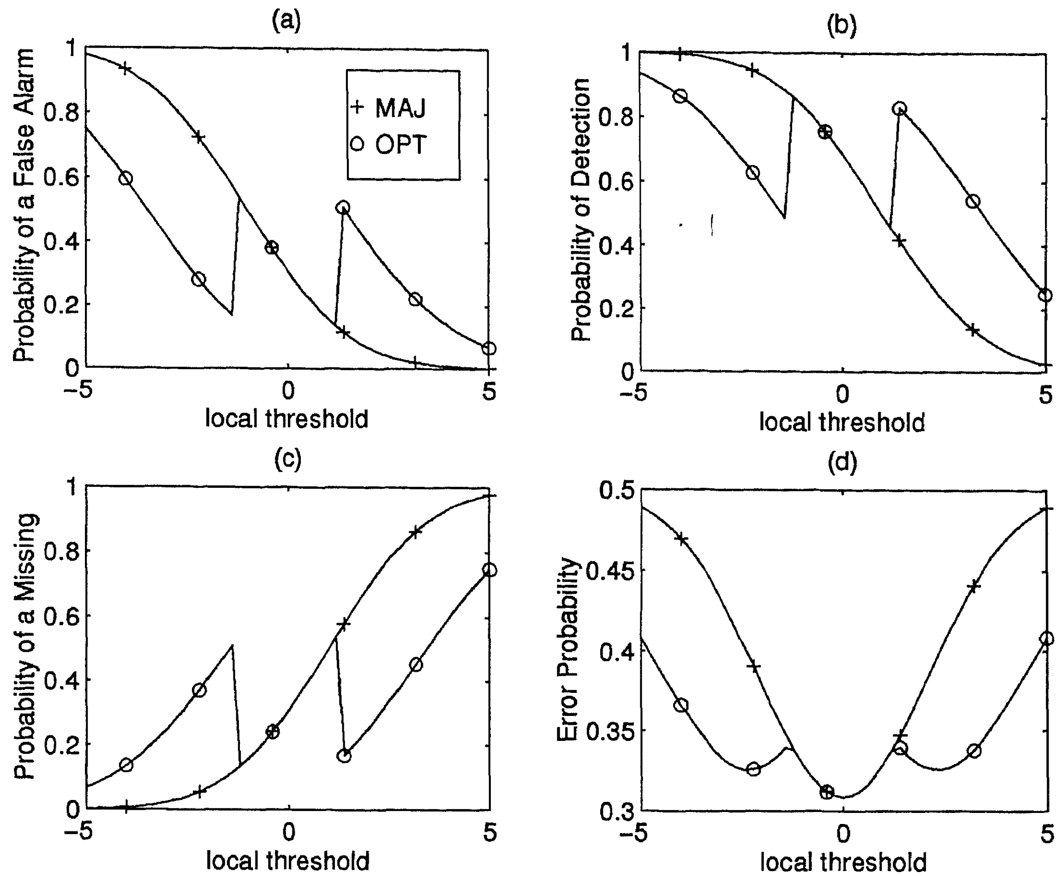

Figure 5.8 Various probabilities versus the local threshold.

used to compare the operating performance of different detection rules), the same probability of a false alarm corresponds to different operating points for different fusion rules, thus resulting in the unusual phenomenon as shown in Figure 5.9.

5.5 Summary

In this chapter, we have investigated the effects of different factors on the four different fusion rules. From the analysis, the following observations were made.

1. When all the local sensors have identical probabilities of detection and false alarm, the MAJ rule always performs the best regardless of the local threshold and other factors considered in this paper.

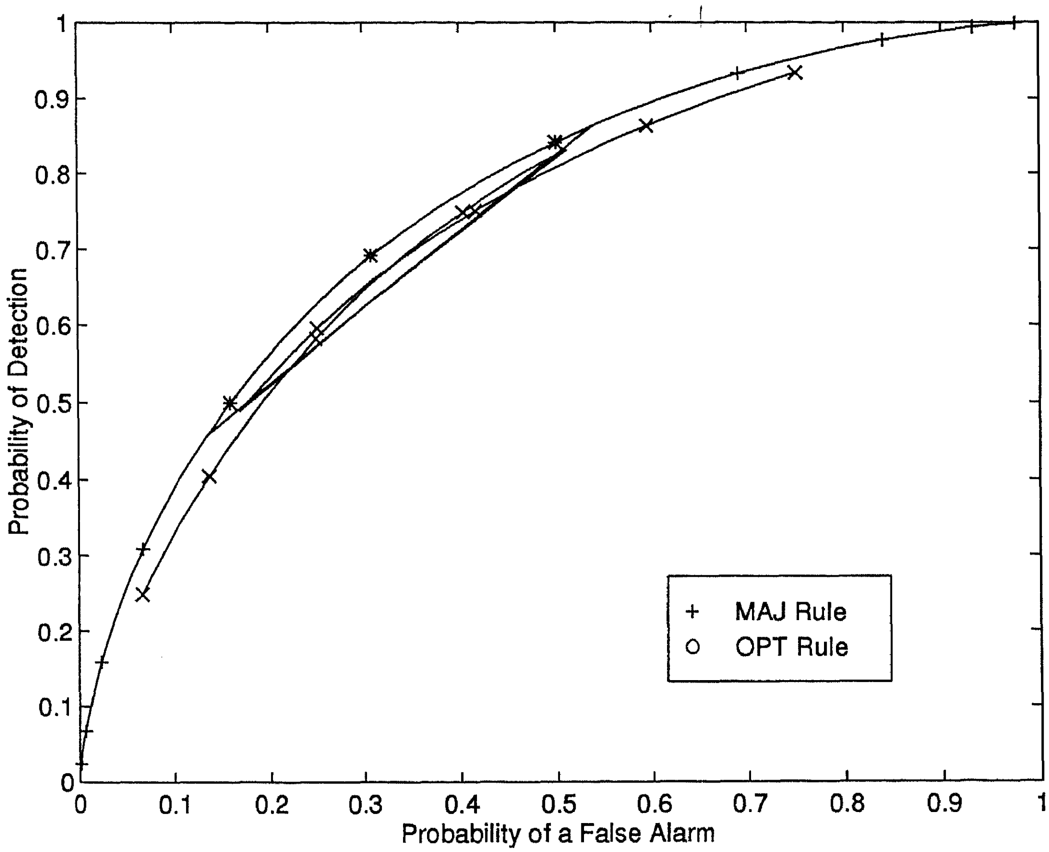

Figure 5.9 ROC comparsion between the MAJ and OPT rule.

2. Only when the deviation of local decision probabilities exists, can OPT, AND and OR rules outperform the MAJ rule. The larger the deviation, the better the performance that can be achieved by the OPT rule.

3. The larger the SNR, the number of sensors, and deviation of local decision probabilities, the more sensitive the OPT is to the local threshold. The smaller the correlation coefficient, the more sensitive the OPT is to the local threshold.

4. The effects of number of sensors and SNR on the performance in the independent case and in the correlated case are similar.

5. The larger the correlation coefficient, the more insignificant the performance difference among fusion rules becomes.

In order to fully exploit the performance advantage of the OPT rule, sophisticated algorithms are required to ensure that the local sensors are working at their optimal thresholds. This is a topic of our future research.

CHAPTER 6

APPLICATION OF DECISION FUSION TO MACRO DIVERSITY IN CELLULAR CDMA

In this chapter, a data/decision fusion technique is proposed to deal with the macro diversity problem. Instead of selecting the best base station, the user is always served by three base stations whenever it gets within their area of coverage. Every base station detects the desired transmission independently and conveys its detection results to a fusion center (or switching center) where the final detection result about that user's signal is formed by optimal data fusion [7]. Since the information from all base stations about the desired user is exploited, better performance than that with conventional selection diversity can be achieved.

6.1 Application of Fusion to Macroscopic Diversity

The cell geometry shown in Figure 6.1 is the same as in [12]. A simple sectored antenna is employed at each site with each antenna sector covering 120° azimuth. The detection is performed at each base station. The detection result is sent through a separate link to a fusion (or switching) center which, as symbolically shown in Figure 6.1, is shared by three base stations. The final detection is made at the fusion center by optimal fusion [7] based on the detected results from the three base stations covering the same area. Let $U = [u_1, u_2, u_3]$ be the vector of detected bits for the desired user. Here, $u_i \in \{1, -1\}$, $i = 1, 2, 3$, is the local decision made by the ith base station. Synchronization among the base stations is assumed, and thus, u_i for $i = 1, 2, 3$ corresponds to the same information bit transmitted. The final detection result at the fusion center for the same information bit, denoted by u_f, is a function of local decisions. The determination of u_f can be viewed as a two-hypothesis detection problem with individual local decisions being the observations, and the two hypotheses

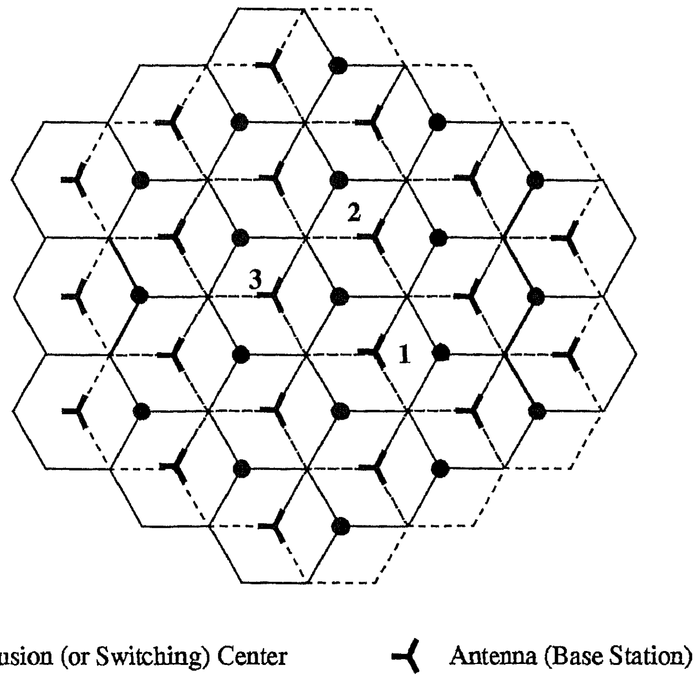

Figure 6.1 A fusion macroscopic diversity scheme.

H_1: The symbol +1 is transmitted,

H_0: The symbol −1 is transmitted.

When the minimum probability of error criterion is used, for the binary symmetric channel[1] (BSC) and equiprobable source bits, we have [13]

$$u_f = f(u_1, u_2, u_3) = \begin{cases} +1, & \text{if } \lambda = \sum_{i=1}^{3} a_i u_i > 0, \\ -1, & \text{otherwise,} \end{cases} \tag{6.1}$$

where the optimal weights are:

$$a_1 = \log \frac{1 - P_1}{P_1}, \quad a_2 = \log \frac{1 - P_2}{P_2}, \quad a_3 = \log \frac{1 - P_3}{P_3}, \tag{6.2}$$

and P_i, $i = 1, 2, 3$, is the error probability for each base station for the same user.

Proposition 1 Denote the BER for a user in each antenna site in a BSC for equiprobable sources as P_i, $i = 1, 2, 3$. Then the BER achieved at the fusion center,

[1] Implicit assumption is made that full interleaving is used.

P_f, for the same user is

$$P_f = \min\{P_m, P_1, P_2, P_3\}, \tag{6.3}$$

where $P_m = P_1P_2 + P_1P_3 + P_2P_3 - 2P_1P_2P_3$. □

Proof The user index is omitted for notational simplicity. Thus, from Eq. (7.1),

$$\lambda = \sum_{i=1}^{3} (a_i)\, u_i = \sum_{i=1}^{3} \log\left(\frac{1-P_i}{P_i}\right) u_i. \tag{6.4}$$

According to the optimal fusion rule, the bit error probability at the fusion center is

$$P_b = P(\lambda > 0|H_0)P(H_0) + P(\lambda < 0|H_1)P(H_1) = \frac{1}{2}[P(\lambda > 0|H_0) + P(\lambda < 0|H_1)], \tag{6.5}$$

where H_0 and H_1 represent the events that symbols -1 and $+1$ are transmitted, respectively. Without loss of generality, let $P_1 < P_2 < P_3$. Thus, $a_1 > a_2 > a_3$, implying that

$$\lambda \begin{cases} > 0 & \text{if } [u_1, u_2, u_3]^t = [1,1,1]^t,\ [1,1,-1]^t \text{ or } [1,-1,1]^t \\ < 0 & \text{if } [u_1, u_2, u_3]^t = [-1,1,-1]^t,\ [-1,-1,1]^t \text{ or } [-1,-1,-1]^t \\ \text{either} & \text{if } [u_1, u_2, u_3]^t = [1,-1,-1]^t \text{ or } [-1,1,1]^t. \end{cases} \tag{6.6}$$

The case referred to above as "either" can further be split into the following two cases:

		$[u_1, u_2, u_3]^t$	
		$[1,-1,-1]^t$	$[-1,1,1]^t$
λ	case I	> 0	< 0
	case II	< 0	> 0

Case I implies that

$$\log(\frac{1-P_2}{P_2}) + \log(\frac{1-P_3}{P_3}) < \log(\frac{1-P_1}{P_1}), \tag{6.7}$$

and thus

$$P_1 < P_1P_2 + P_1P_3 + P_2P_3 - 2P_1P_2P_3 = P_m. \tag{6.8}$$

Hence,

$$\begin{aligned} P(\lambda > 0|H_0) &= P_1P_2P_3 + P_1P_2(1-P_3) + P_1(1-P_2)P_3 + P_1(1-P_2)(1-P_3) \\ &= P_1P_2 + P_1(1-P_2) = P_1. \end{aligned} \tag{6.9}$$

Likewise, by the symmetry of error probability for each channel,

$$P(\lambda < 0|H_1) = P_1.$$

Therefore, $P_b = \frac{1}{2}\left(P(\lambda > 0|H_0) + P(\lambda < 0|H_1)\right) = P_1 \leq P_m$.

Similarly, case II implies that

$$a_2 + a_3 > a_1 \quad \Longrightarrow \quad P_1 > P_m. \tag{6.10}$$

Also,

$$\begin{aligned} P(\lambda > 0|H_0) &= P_1P_2P_3 + P_1P_2(1-P_3) + P_1(1-P_2)P_3 + (1-P_1)P_2P_3 \\ &= P_1P_2 + P_1P_3 + P_2P_3 - 2P_1P_2P_3 = P_m \\ P(\lambda < 0|H_1) &= P(\lambda > 0|H_0) = P_m. \end{aligned} \tag{6.11}$$

Therefore, $P_f = P_m < P_1$.

Combining the two cases, we have:

$$P_f = \min\{P_1, P_m\}. \tag{6.12}$$

In conclusion, $P_f = \min\{P_m, P_1, P_2, P_3\}$.

□

Proposition 1 shows that the instantaneous BER at the fusion center is less than or equal to the minimum instantaneous BER of each receiver of the three base stations.

6.2 Adaptive Algorithm Analysis

When the error probabilities, P_i, at each antenna are unknown and time-varying, the following adaptive algorithm is introduced to perform the fusion operation [4]:

$$a_i \approx \begin{cases} \log \frac{m_{1i}}{n_{1i}}, & \text{if} \quad u_i = +1, \\ \log \frac{m_{0i}}{n_{0i}}, & \text{if} \quad u_i = -1, \end{cases} \tag{6.13}$$

where,

m_{1i} is the number of the occurrences of $u_i = +1$ and $u_f = +1$,

m_{0i} is the number of the occurrences of $u_i = -1$ and $u_f = -1$,

n_{1i} is the number of the occurrences of $u_i = +1$ and $u_f = -1$, and

n_{0i} is the number of the occurrences of $u_i = -1$ and $u_f = +1$.

The estimated weights obtained by the adaptive fusion algorithm [13] are:

$$b_1 = \log \frac{1-q_1}{q_1}, \quad b_2 = \log \frac{1-q_2}{q_2}, \quad b_3 = \log \frac{1-q_3}{q_3},$$

where q_i is the estimation of P_i. The errors between the estimated and optimal weights are:

$$\varepsilon_1 = b_1 - a_1, \quad \varepsilon_2 = b_2 - a_2, \quad \varepsilon_3 = b_3 - a_3. \tag{6.14}$$

According to [4], the minimal weights errors that can be achieved are:

$$\begin{aligned}
\varepsilon_1 &= \log \frac{1 + \frac{P_1P_2P_3}{(1-P_1)(1-P_2)(1-P_3)}}{1 + \frac{(1-P_1)P_2P_3}{P_1(1-P_2)(1-P_3)}}, \\
\varepsilon_2 &= \log \frac{1 + \frac{P_1P_2P_3}{(1-P_1)(1-P_2)(1-P_3)}}{1 + \frac{(1-P_2)P_1P_3}{P_2(1-P_1)(1-P_3)}}, \\
\varepsilon_3 &= \log \frac{1 + \frac{P_1P_2P_3}{(1-P_1)(1-P_2)(1-P_3)}}{1 + \frac{(1-P_3)P_1P_2}{P_3(1-P_1)(1-P_2)}}.
\end{aligned} \tag{6.15}$$

The following propositions leading to the derivation of the minimum error probability show the relationship between the optimal and estimated weights.

Proposition 2 If $P_1 < 0.5$, $P_2 < 0.5$, and $P_3 < 0.5$, then $a_1 > a_2 > a_3 \Longrightarrow b_1 > b_2 > b_3$.

The proof is given in Appendix A.

Proposition 3 If $P_1 < 0.5$, $P_2 < 0.5$, and $P_3 < 0.5$, then $b_1 < b_2 + b_3, b_2 < b_1 + b_3$, and $b_3 < b_1 + b_2$.

The proof is given in Appendix B.

Proposition 4 The minimum error probability using the adaptive fusion algorithm is

$$P_m = P_1P_2 + P_1P_3 + P_2P_3 - 2P_1P_2P_3.$$

Proof:

From Proposition 3, $b_2 + b_3 > b_1$, $b_1 + b_3 > b_2$, and $b_1 + b_2 > b_3$, implies that the vectors which make $\lambda > 0$ are [1,1,1], [1, 1 -1], [1, -1, 1], and [-1, 1,1]. Thus, the error probability when the adaptive fusion scheme is used and u_i, $i = 1, 2, 3$ is independent will be :

$$\begin{aligned}
P(\lambda > 0|H_0) &= P(u_1 = +1, u_2 = +1, u_3 = +1|H_0) + \\
&\quad P(u_1 = -1, u_2 = +1, u_3 = +1|H_0) + \\
&\quad P(u_1 = +1, u_2 = -1, u_3 = +1|H_0) + \\
&\quad P(u_1 = +1, u_2 = +1, u_3 = -1|H_0) \\
&= P(u_1 = +1|H_0)P(u_2 = +1|H_0)P(u_3 = +1|H_0) + \\
&\quad P(u_1 = -1|H_0)P(u_2 = +1|H_0)P(u_3 = +1|H_0) + \\
&\quad P(u_1 = +1|H_0)P(u_2 = -1|H_0)P(u_3 = +1|H_0) + \\
&\quad P(u_1 = +1|H_0)P(u_2 = +1|H_0)P(u_3 = -1|H_0) \\
&= P_1P_2P_3 + (1 - P_1)P_2P_3 + P_1(1 - P_2)P_3 + P_1P_2(1 - P_3) \\
&= P_1P_2 + P_1P_3 + P_2P_3 - 2P_1P_2P_3 \\
&= P_m.
\end{aligned} \tag{6.16}$$

6.3 A Realization of the Optimal Fusion Scheme for Macro Diversity

In current macroscopic diversity systems, when selection diversity is used to form the final decision about the information transmitted, the achievable bit error rate (BER) of the final decision, P_f, is

$$P_f = \min\{P_1, P_2, P_3\} \tag{6.17}$$

where P_i, $i = 1, 2, 3$ as defined previously, is the BER of the ith base station for the same information bit. When the optimal fusion scheme (Eq. (7.1)) is used for macroscopic diversity combination, it has been proved [13] that

$$P_f = \min\{P_1, P_2, P_3, P_m\}, \tag{6.18}$$

where $P_m = P_1P_2 + P_1P_3 + P_2P_3 - 2P_1P_2P_3$ as in Proposition 3. It has been shown that the fusion scheme has better performance than selection diversity when only shadowing is considered [12]. Whenever P_m is less than $\min\{P_1, P_2, P_3\}$, especially when differences between P_1, P_2 and P_3 are small, P_f in Eq. (6.18) is less than the P_f in Eq. (6.17). The drawback of the fusion method is its higher complexity compared to the selection diversity. All of the u_1, u_2 and u_3 have to be transmitted to a switching or fusion center where the optimal combination based on P_1, P_2 and P_3 is performed according to Eq. (7.1). In this paper, based on the analysis resulting in Eq. (6.18), we propose a simplified realization of the optimal fusion scheme that has lower complexity. In this realization, the final detection of a transmitted information bit, u_f, will be an element selected from the binary data set $D = \{u_1, u_2, u_3, \mathrm{Maj}(u_1, u_2, u_3)\}$ with the smallest error probability, where Maj(.) stands for the majority operator defined by

$$\mathrm{Maj}(u_1, u_2, u_3) = \begin{cases} +1 & \text{if } u_1 + u_2 + u_3 > 0, \\ -1 & \text{if } u_1 + u_2 + u_3 < 0. \end{cases}$$

When u_i's are mutually independent with respective BERs P_i for $i = 1, 2, 3$, the BER for this majority operator is :

$$P(\mathrm{Maj}(u_1, u_2, u_3) = +1|H_0) \;=\; P(u_1 = +1, u_2 = +1, u_3 = +1|H_0) +$$

$$
\begin{aligned}
& P(u_1=-1,u_2=+1,u_3=+1|H_0)+ \\
& P(u_1=+1,u_2=-1,u_3=+1|H_0)+ \\
& P(u_1=+1,u_2=+1,u_3=-1|H_0) \\
= \; & P(u_1=+1|H_0)P(u_2=+1|H_0)P(u_3=+1|H_0)+ \\
& P(u_1=-1|H_0)P(u_2=+1|H_0)P(u_3=+1|H_0)+ \\
& P(u_1=+1|H_0)P(u_2=-1|H_0)P(u_3=+1|H_0)+ \\
& P(u_1=+1|H_0)P(u_2=+1|H_0)P(u_3=-1|H_0) \\
= \; & P_1P_2P_3+(1-P_1)P_2P_3+ \\
& P_1(1-P_2)P_3+P_1P_2(1-P_3) \\
= \; & P_1P_2+P_1P_3+P_2P_3-2P_1P_2P_3 \\
= \; & P_m.
\end{aligned} \tag{6.19}
$$

The above equation implies that the majority operator yields a BER P_m. Therefore, the above realization implements the optimal fusion rule. The realization is much easier than the direct realization according to Eq. (7.1), because only selection and the majority operator are required. The majority operator for macroscopic diversity has been proposed in [8] and [42]. Another advantage of this realization is that the entire $U = \{u_1, u_2, u_3\}$ does not always have to be sent to the switching center. Only when $P_m < \min\{P_1, P_2, P_3\}$, all three elements of U are required at the switching center for the majority operation. Otherwise, only the element u_i with the smallest BER is transmitted to the switching center.

6.4 Performance Comparison

In [12], we compared the performance of using the fusion scheme with the selection diversity when only shadowing distortion is considered. Here, both the shadowing and fading effects are taken into consideration. The flat Rayleigh fading and shadowing effect modeled by log-normal distributed are assumed. The error probability is considered as the performance index. In addition, as in a practical system, we assume that the maximal ratio combiner is used for the microscopic diversity to combat the fading distortion. According to [36], the instantaneous received power at the output of the L-branch micro combiner is a chi-square distributed random variable with 2L degrees of freedom. The conditional error probability for the fixed local mean received power will be the instantaneous bit error probability averaged over fading channel statistics, which can be written as follows

$$P_i = \left[\frac{1}{2}(1-\mu_i)\right]^L \sum_{k=0}^{L-1} \binom{L-1+k}{k} \left[\frac{1}{2}(1+\mu_i)\right]^k, \tag{6.20}$$

where, by definition,

$$\mu_i = \sqrt{\frac{\gamma_i}{1+\gamma_i}},$$

and γ_i is the averaged signal to noise ratio (SNR) over fading statistics for the ith base station (local mean of SNR). When the shadowing effect is considered, the local mean of the received power is a log-normal distributed random variable. When the power spectrum density of thermal noise and interference are assumed to be a constant, the local mean of the SNR is also log-normal distributed. Thus the area-mean BER, when *no macroscopic diversity* is employed, equals to

$$P_f = \int P(\gamma)f(\gamma)d\gamma, \tag{6.21}$$

where $f(\gamma)$ is the probability density function of γ, the local mean of the SNR, at a base station. According to the previous discussion, $f(\gamma_1)$ is a log-normal function with a mean (determined by the distance between the mobile user and the base station, and the propagation environment), and a variance (determined by the power control scheme). When the macroscopic selection diversity is used, the area-mean BER is

$$P_f = \int\int\int \min(P_1, P_2, P_3) f(\gamma_1, \gamma_2, \gamma_3) d\gamma_1 d\gamma_2 d\gamma_3, \tag{6.22}$$

where $f(\gamma_1, \gamma_2, \gamma_3)$ is the joint probability density function. When the fusion based macroscopic diversity is implemented for three base stations, the area-mean BER is

$$P_f = \int\int\int \min(P_1, P_2, P_3, P_m) f(\gamma_1, \gamma_2, \gamma_3) d\gamma_1 d\gamma_2 d\gamma_3. \tag{6.23}$$

Since base stations are far away from each other, the random variables γ_1, γ_2 and γ_3 can be regarded as independent variables. When a mobile user is equidistant from the three base stations, the local mean SNR have the same statistical parameters, and thus

$$f(\gamma_1, \gamma_2, \gamma_3) = f(\gamma_1) f(\gamma_2) f(\gamma_3).$$

Figure 6.2 shows the curve of the area-mean BER versus SNR for the nonmacro diversity, selection macro diversity, and fusion based macro diversity obtained, by numerically calculating Eqs (6.21)-(6.23). The numerical results are derived for L=3, and the standard deviation of the local-mean SNR of 1.5 dB. From this Figure, it is shown that significant improvement can be achieved by the fusion based macro diversity even in the presence of both fading and shadowing.

6.5 Summary

The performance of the adaptive fusion algorithm for macroscopic diversity has been analyzed. The minimum error probability that can be achieved by the adaptive fusion method equals to that by the majority rule. A less complex realization of the optimal fusion scheme is also proposed. The realization is equivalent to the combination of the conventional macro selection diversity and a majority operator,

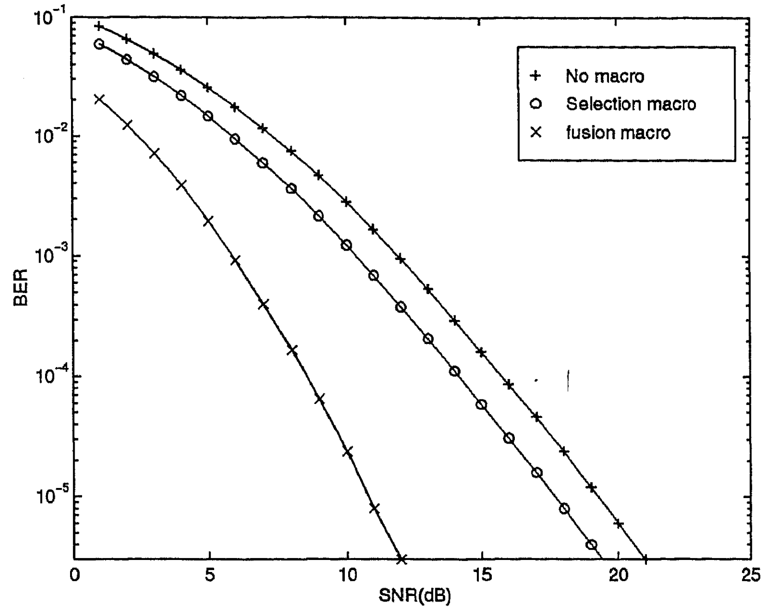

Figure 6.2 Performance of different macro diversity schemes.

and is demonstrated to outperform the conventional macroscopic selection diversity when both fading and shadowing are involved.

CHAPTER 7

DECISION FUSION FOR HANDOFF IN CELLULAR CDMA

Handoff, an essential component of cellular networks, provides uninterrupted communication and maintains call quality while a mobile user is in the transition from one cell coverage area to another. Generally, there are two approaches for implementing this network function: hard handoff and soft handoff [46], [45]. Hard handoff is the technique that abruptly transfers the services from one base station to another base station which provides better service quality. Because of the mobility of users and randomness of the received signal power levels, sufficient overlap in coverage area between adjacent cells has to be established, which requires extra power in order to maintain service quality. Under soft handoff, a mobile user is supported by more than one base station simultaneously, whenever such a user enters the boundary region among cells. By always choosing the base station that receives the strongest signal from the desired user, the switching center imposes a lower power requirement from the mobile user compared to the hard handoff [46], [45]. The fluctuation of the received signal power due to fading and shadowing is considered as added difficulty both to hard handoff and soft handoff. Both techniques require the determination of what is actually the best station for a particular user. Sophisticated methods such as adaptive averaging [26], which make use of signal strength variation, have been proposed to find such a station.

Recently, decision fusion has been used to improve the performance of CDMA [9], [10]. In [9, 10], a distributed detection combining rule is proposed to incorporate local decisions when the bit error rate at each branch is assumed the same. In previous chapter, an decision fusion scheme is applied to CDMA macroscopic diversity combining, whereby each base station is connected with three widely spaced antenna sectors (and their receivers) that separately detect the received signal. The detection results are conveyed to the base station where the final detection result is made by optimal fusion.

In this chapter, a data/decision fusion technique is proposed to deal with the handoff problem. Instead of selecting the best base station, the user is always served by three base stations whenever it gets within their area of coverage. Every base station detects the desired transmission independently and conveys its detection results to a fusion center (or switching center) where the final detection result about that user's signal is formed by optimal data fusion [7]. Since the information from all base stations about the desired user is exploited, better performance than that with soft handoff can be achieved. An additional advantage of the proposed approach is that it reduces the number of handoffs.

7.1 Data Fusion for Handoff

The geometrical arrangement of antennae for the approach described in this chapter is the same as in Chapter 6. A simple sectored antenna is employed at each site with each antenna sector covering 120° azimuth. The detection is performed at each antenna sector with its associated receiver. The detection result is sent through a separate link to a fusion center which, as symbolically shown in Figure 6.1, is shared by three simple sectored antennae. The final detection is made at the fusion center by optimal fusion [7] based on the detected results from three separate antenna sectors covering the same area. Let $U = [u_1, u_2, u_3]$ be the vector of detected bits for the desired user. Here, $u_i \in \{1, -1\}$, $i = 1, 2, 3$, is the local decision made by the ith antenna sector. Assume synchronization has been achieved among antennae, so that u_i for $i = 1, 2, 3$ correspond to the same information bit transmitted. The final detection result at the fusion center for the same information bit is denoted by u_f, which is a function of local decisions. The determination of u_f can be viewed as a two-hypothesis detection problem with individual local decisions being the observations.

When the minimum probability of error criterion is adopted for equal probable source and BSC channel, we have

$$u_f = f(u_1, u_2, u_3) = \begin{cases} +1, & \text{if} \quad \lambda = \sum_{i=1}^{3} a_i u_i > 0, \\ -1, & \text{otherwise,} \end{cases} \tag{7.1}$$

where

$$a_i = \log \frac{1 - P_i}{P_i} \tag{7.2}$$

7.2 Handoff Performance Analysis of Distributed Detection

From previous section, it can be seen that the detected result at the fusion center is based on an optimal combination of information from three separate channels. Under soft handoff, however, the detection on the reverse link is performed on the strongest signal at a time, on a frame by frame basis [45]. Although the signal received by three base stations from a mobile user located at the boundary of three cells has a different power level, the difference between them is often not very large. Selecting only the best base station loses the useful information about the desired user at "inferior" base stations.[1] Therefore, better performance is expected from the decision fusion method.

Due to shadowing and intercell interference, the signal power, and thus the signal-to-interference plus noise ratio (SINR) can, following [46] and [47], be treated as random variables.[2] Furthermore, in analyzing the system capacity, the effect on the SINR from intercell interference can be approximated by a constant, and therefore the randomness of the SINR can be solely attributed to shadowing. The rate of change of the signal power caused by shadowing can be such that the power

[1]The difference between decision fusion and soft handoff is similar to the difference between maximal ratio combining and selection diversity in a diversity combination system [27].

[2]The reason for not including the fading effect is that each base station is usually equipped with a diversity receiver (e.g., RAKE receiver), which mitigates the fading due to multipath.

level is almost constant during intervals ranging from seconds to hours, making the deviation of the signal power quite slow relative to the bit rate (e.g., 9600 bit/sec). Within such intervals, the SINR can thus be assumed as fixed. Therefore, according to analysis in Chapter 6, the bit error rate (BER) at the base stations and fusion center exhibits the following important property:

$$P_f = \min\{P_m, P_1, P_2, P_3\}, \tag{7.3}$$

where $P_m = P_1P_2 + P_1P_3 + P_2P_3 - 2P_1P_2P_3$. Eq. (7.3) shows that the instantaneous BER at the fusion center is less than or equal to the minimum instantaneous BER of each receiver of the three base stations. The BERs at both the base stations and the fusion center, which are random variables because of the shadowing effect, cannot be approximated as constants over a large time scale. The outage probability is introduced to measure the performance of a detection scheme, and is defined as the probability of the BER exceeding a certain threshold. The next proposition shows the relationship of the outage probability in the base station and the fusion center.

Proposition 7.1 If R_1, R_2 and R_3 are denoted as the outage probabilities of a user at antenna sites 1, 2, and 3, respectively, the upper bound on the outage probability at the fusion center, R_f, is

$$R_f < R_1R_2R_3. \tag{7.4}$$

□

Proof According to fundamental statistics [34], let $r > 0$ denote the protection margin. By the definition of outage probability and Proposition 1, we have

$$\begin{aligned} R_f &= P(P_f > r) = P(\min(P_1, P_2, P_3, P_m) > r) \\ &= P(P_1 > r, P_2 > r, P_3 > r, P_m > r) \\ &< P(P_1 > r, P_2 > r, P_3 > r, P_m > 0). \end{aligned}$$

Since

$$P(P_1 > r, P_2 > r, P_3 > r, P_m > 0) = P(P_1 > r, P_2 > r, P_3 > r),$$

$$R_f < P(P_1 > r, P_2 > r, P_3 > r) = P(P_1 > r)P(P_2 > r)P(P_3 > r) = R_1 R_2 R_3.$$

□

Proposition 7.1 shows that the outage probability at the fusion center is less than the product of the outage probability at the three antenna sectors. Here, R_i, $i = 1,2,3$, correspond to the outage probability for hard handoff at each individual antenna sector.

7.3 Handoff Performance Comparison

The performance advantage of decision fusion over soft handoff can be analyzed in three different aspects. Without loss of generality, suppose that the user under consideration is in the coverage of site 1.

7.3.1 Reverse Link Capacity

It has been demonstrated both theoretically and experimentally that the signal power received at a base station is a random variable. When only shadowing is considered, the received signal power W_i is usually modeled as the product of distance attenuation and a log-normal random variable [45]:

$$W_i = S r_i^{-\delta} 10^{\frac{\zeta_i}{10}}, \quad i = 1,2,3, \tag{7.5}$$

where i refers to the index of the base station, and $W, S, r,$ and δ are the received power, the power transmitted by the mobile user, the distance between the base station and the user, and distance attenuation exponent, respectively. In most of the analyses, δ is usually set to 4. ζ_i is a normally distributed random variable with a zero mean and a standard deviation σ_i ranging from 2.5 dB for a power controlled user to 8 dB for a non-power controlled user.

Table 7.1 The relative other-cell interference factor.

The Standard Deviation σ	Hard Handoff	Soft Handoff with Two Base Stations	Soft Handoff with Three Base Stations
2	0.48	0.43	0.43
4	0.67	0.47	0.45
6	1.13	0.56	0.49
8	2.38	0.77	0.57

In the proposed scheme, even when the desired user is deep in the coverage of antenna 1, base stations 2 and 3 also receive its signal and perform detection. Under this scenario, however, because of large σ_2 and σ_3, the performance on the reverse link mostly relies on base station 1. When the mobile is approaching the boundary, σ_1 is increasing, and σ_2 and/or σ_3 are decreasing. When the mobile is exactly on the boundary between base stations 1 and 2, it can be assumed that σ_1 is equal to σ_2. If the mobile is at the boundary between three base stations, the three σ's can be assumed equal. The outage probability for the desired user at base station i can be well approximated by [45, 47]

$$R_i = Q\left[\frac{K_0' - \rho(\lambda/\mu)(1+f)e^{(\beta\sigma_i)^2/2}}{\sqrt{\rho(\lambda/\mu)(1+f)e^{(\beta\sigma_i)^2}}}\right], \tag{7.6}$$

where K_0' is a parameter determined by the processing gain and the required signal-to-interference plus noise ratio. To allow direct comparison with the results in [45], in this paper K_0' is set to 230.4. In Eq. (7.6), $\rho \approx 0.4$ is the voice activity factor, λ/μ is the normalized average user occupancy, and $\beta = \log_e(10)/10$ is a constant. The relative other-cell interference factor, f, which is calculated by Viterbi [45], is listed in Table (7.1), for different values of σ, when distance attenuation exponent, δ, is equal to 4.

According to Table 7.1 and the approximation leading to Eq. (7.6), the outage probability for hard handoff is the same as in Eq. (7.6) for σ and f, given in Table 7.1.

Table 7.2 The ratio of the relative other-cell interference factor between soft and hard handoff.

Standard deviation	2	4	6	8
ϕ (two base stations)	0.9662	0.8802	0.7324	0.5237
ϕ (three base stations)	0.9662	0.8683	0.6995	0.4645

The reverse link capacity for soft handoff was investigated in [46]. It has been shown that soft handoff reduces the relative other-cell interference. Here a modified version of Eq. (7.6) is used to calculate the outage probability for soft handoff by including a reducing factor, ϕ.

$$R_f = Q\left[\frac{K_0' - \phi\rho(\lambda/\mu)(1+f)e^{(\beta\sigma)^2/2}}{\sqrt{\rho(\lambda/\mu)(1+f)e^{(\beta\sigma)^2}}}\right], \tag{7.7}$$

where ϕ is a ratio of the relative other-cell interference factor between soft and hard handoff, obtained from Table 7.1, and listed in Table 7.2.

The exact expression of outage probability for decision fusion is difficult, if not impossible, to obtain. According to Proposition 7.1, however, it is upper bounded by

$$R_f < \left\{Q\left[\frac{K_0' - \rho(\lambda/\mu)(1+f)e^{(\beta\sigma)^2/2}}{\sqrt{\rho(\lambda/\mu)(1+f)e^{(\beta\sigma)^2}}}\right]\right\}^j, \tag{7.8}$$

where j is either 2 or 3, depending on the position of the mobile user. When the mobile user is at the boundary between two base stations and far away from the third base station, the third base station contributes little information about the mobile user because of distance attenuation. The value of j is then chosen as 2 in Eq. (7.8). When the mobile is right at the boundary among three base stations, j is set to 3. Figure 7.1 plots the outage probabilities versus system capacity ($\rho(\lambda/\mu)(1+f)$) for hard handoff, soft handoff and data fusion when the desired user is at the boundary between base stations 1 and 2 (here $j = 2$). Figure 7.2 shows the outage probability when the user is at the common boundary among the three base stations.

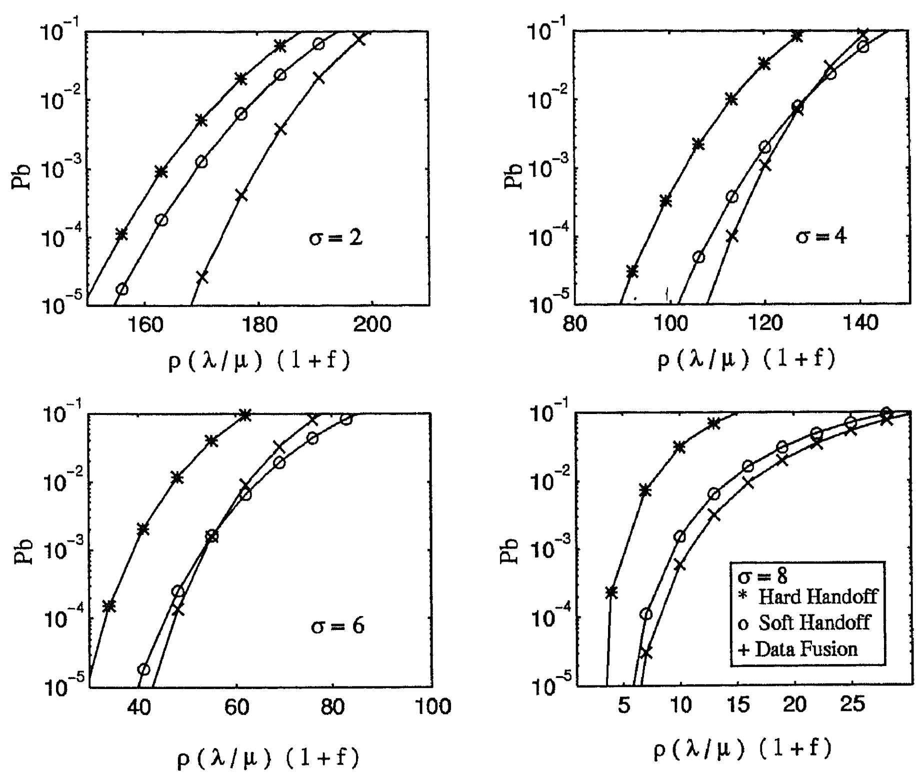

Figure 7.1 System performance when the user is at the boundary of 2 base stations.

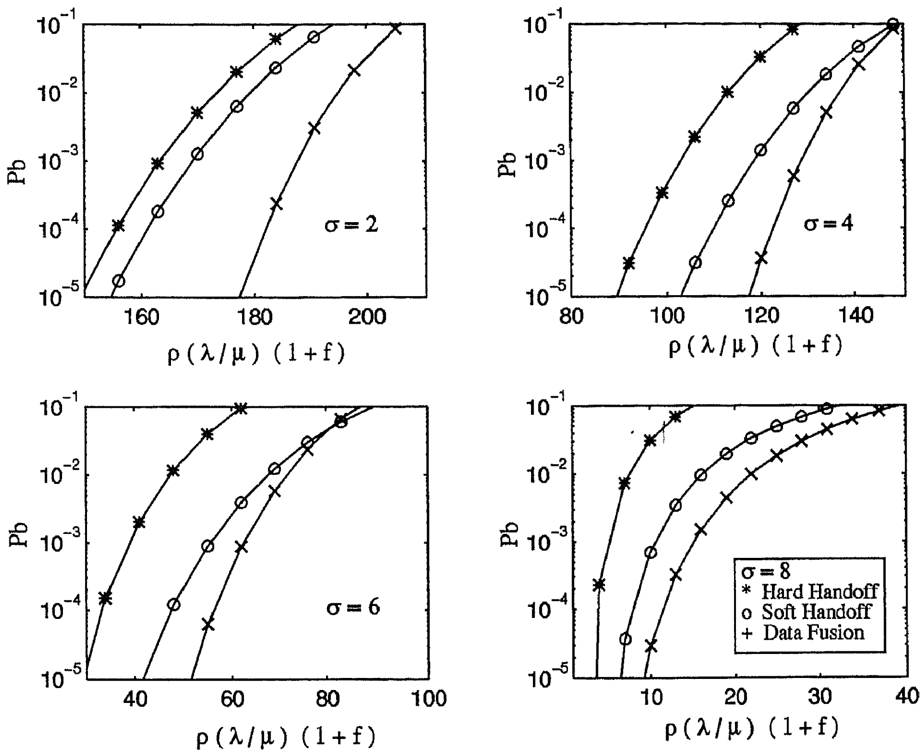

Figure 7.2 System performance when the user is at the boundary of 3 base stations.

7.3.2 Cell Coverage

From Figures 7.1 and 7.2, it can be seen that handoff by decision fusion outperforms the soft handoff in most situations. For some situations, however, both schemes have comparable performance. The reason that the advantage of decision fusion over soft handoff is not so obvious in some situations lies in the fact that both performance calculations introduce approximation, especially for the fusion scheme. In particular, the effect of R_m is omitted, which has a significant effect on performance when the mobile user is at the boundary area, since P_m is much smaller when P_1, P_2 and P_3 are similar.

The handoff performance can also be analyzed in another way by simplifying the outage probability at each base station, as in [46] and [45]:

$$R_1 = \frac{1}{\sqrt{2\pi}\sigma} \int_{\gamma}^{\infty} e^{-\zeta^2/2\sigma^2} d\zeta = Q(\frac{\gamma - 10\delta \log r_1}{\sigma}) \tag{7.9}$$

and

$$R_2 = Q(\frac{\gamma - 10\delta \log r_2}{\sigma}), \qquad R_3 = Q(\frac{\gamma - 10\delta \log r_3}{\sigma}), \tag{7.10}$$

where σ, as defined before, is the standard deviation of shadowing loss, γ is the power margin to achieve the required outage probability, and r_1, r_2 and r_3 are the distances between the user and base stations 1, 2 and 3, respectively. Using Proposition 7.1 again, an upper bound (worst case) for outage probability by decision fusion can be established. Thus,

$$R_f \le Q(\frac{\gamma - 10\delta \log r_1}{\sigma}) Q(\frac{\gamma - 10\delta \log r_2}{\sigma}) Q(\frac{\gamma - 10\delta \log r_3}{\sigma}). \tag{7.11}$$

The outage probability for soft handoff is given by [46]:

$$R_f = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} Q\left(\frac{\gamma - 10\delta \log r_1 + a\sigma x}{b\sigma}\right)$$
$$Q\left(\frac{\gamma - 10\delta \log r_2 + a\sigma x}{b\sigma}\right) Q\left(\frac{\gamma - 10\delta \log r_3 + a\sigma x}{b\sigma}\right) dx, \tag{7.12}$$

Table 7.3 The required power margin by three handoff methods.

Relative Distance		Hard Handoff	Soft Handoff	Decision
r_1	r_2 and r_3	γ_{hard} (dB)	γ_{soft} (dB)	Fusion γ_{dis} (dB)
1	1	10.25	4.23	0.72
1.05	0.98	11.10	4.19	0.70
1.10	0.95	11.91	4.12	0.66
1.15	0.93	12.68	4.01	0.60
1.20	0.92	13.42	3.89	0.54
1.25	0.90	14.13	3.76	0.47

Table 7.4 The increase in coverage.

Relative Distance		Soft Handoff	Decision Fusion
r_1	r_2 and r_3	Coverage Area	Coverage Area
1	1	2.00	3.00
1.05	0.98	2.22	3.31
1.10	0.95	2.45	3.65
1.15	0.93	2.71	4.02
1.20	0.92	3.00	4.41
1.25	0.90	3.30	4.82

where $a = b = 1/\sqrt{2}$, as in [46]. For the purpose of comparison, the outage probability for hard handoff is also given as

$$R_f = Q(\frac{\gamma - 10\delta \log r_1}{\sigma}), \tag{7.13}$$

where r_1 is set to greater than 1 to reduce the "ping-pong" effect.

Table 7.3 lists the required power margin by three handoff methods for different relative distances when σ=8dB. The corresponding increase in coverage is listed in Table 7.4.

7.3.3 The Number of Handoffs

In addition to the advantage of increased coverage and decreased outage probability, the proposed decision fusion technique also reduces the number of handoffs compared to hard and soft handoff. From Figure 6.1, it can be seen that each user is always served by three base stations simultaneously, although sometimes, depending on the user's position, the contributions of some base stations are insignificant. In fact, by using decision fusion, a new geometrical structure is formed, which is denoted by the dashed line in Figure 6.1. When a user is within the dashed line boundary, no transition has to be done, even though the mobile may be crossing the boundary between different base stations. Only when the mobile moves from one dashed hexagon to another, the service is transferred to another group of base stations. Thus, the number of handoffs is reduced, as pointed out by Lee [31], by one half.

7.4 Summary

In this chapter, a decision fusion method to address the handoff problem in cellular CDMA by using optimal data fusion was proposed and analyzed. The performance of decision fusion was analyzed in terms of reverse link system capacity, cell coverage, and the number of handoffs. It was demonstrated by numerical examples that decision decision fusion approach generally outperforms both soft and hard handoff methods.

CHAPTER 8
CONCLUSIONS

In a real-world environment, probability mass functions of observations at local detectors may not be known and the performance of the local detectors may not be stationary. Under such circumstances, it is desirable to have a system which can adapt itself during the decision making process. This dissertation proposes such an adaptive system for both equal probable and uneuqal probable sources when the local decisions are independent as well as correlated. The major advantage of the system is that *a priori* knowledge of the probability mass functions of the observations is not required. The system can acquire the knowledge about the reliability of the local detectors by itself — *it can learn by doing.* A reinforcement learning rule is proposed and adopted, and its convergence is analytically proven. The simulation results conform to our theoretical analysis. The following main attributes of the adaptive fusion algorithm can be concluded from the theoretical analysis and simulations.

1) It does not require *a priori* knowledge about the sensors and source, and thus is more practical.

2) It adapts the weights from time to time, and thus is suitable for a time-varying environment.

3) In most cases, it behaves as well as the optimal rule.

4) Its computational complexity is low, and thus implementable.

In addition, we have compared the performance of following four practical fusion rules: AND, (referred to as AND in this dissertation), OR (OR), majority (MAJ) and Chair's optimal rule (OPT). We have investigated the effects of different factors on the four different fusion rules. From the analysis, the following observations were made.

1. When all the local sensors have identical probabilities of detection and false alarm, the MAJ rule always performs better or as good as the other rules regardless of the local threshold and other factors considered in this dissertation.

2. Only when the deviation of local decision probabilities exists, can OPT, AND and OR rules outperform the MAJ rule. The larger the deviation, the better the performance that can be achieved by the OPT rule.

3. The larger the SNR, the number of sensors, and deviation of local decision probabilities, the more sensitive the OPT is to the local threshold. The smaller the correlation coefficient, the more sensitive the OPT is to the local threshold.

4. The effects of the number of sensors and SNR on the performance in the independent case and in the correlated case are similar.

5. The larger the correlation coefficient, the more insignificant the performance difference among fusion rules becomes.

In order to fully exploit the performance advantage of the OPT rule, sophisticated algorithms are required to ensure that local sensors are working at their optimal thresholds.

Finially, the data fusion scheme is applied to cellular Code Division Multiple Access (CDMA) to improve the performance of macroscopic diversity. Theoretical analysis and computer simulations have shown more reliable detection is obtained at the switching center by fusing the detected results from base stations. This method exploits the spatial diversity that already exists in the current system without increasing the amount of data transmission between base stations and the switching center. A simple realization of the fusion scheme in which the combination can be replaced by selection and majority vote, has also been proposed in this dissertation.

APPENDIX A

Proof of Proposition 2 in Chapter 6

$a_1 > a_2 > a_3$ implies $P_3 > P_2 > P_1$. From the condition

$$1 > 2P_3,$$

we have

$$P_2 - P_1 > 2P_3(P_2 - P_1).$$

Subtracting $P_1P_2P_3$ from both sides and rearranging the terms, we have

$$P_2(1 - P_3) + (1 - P_2)P_1P_3 > P_1(1 - P_3) + (1 - P_1)P_2P_3.$$

Dividing both sides by $(1 - P_3) > 0$,

$$P_2 + \frac{(1 - P_2)P_1P_3}{1 - P_3} > P_1 + \frac{(1 - P_1)P_2P_3}{1 - P_3}.$$

Subtracting P_1P_2 from both sides and factoring out $(1 - P_1)P_2$ from the left side and $(1 - P_2)P_1$ from the right side, the above inequality becomes

$$(1 - P_1)P_2[1 + \frac{(1 - P_2)P_1P_3}{P_2(1 - P_1)(1 - P_3)}] > P_1(1 - P_2)[1 + \frac{(1 - P_1)P_2P_3}{P_1(1 - P_2)(1 - P_3)}]. \quad (A.1)$$

Since

$$P_1P_2[1 + \frac{(1 - P_2)P_1P_3}{P_2(1 - P_1)(1 - P_3)}][1 + \frac{(1 - P_1)P_2P_3}{P_1(1 - P_2)(1 - P_3)}] > 0, \quad (A.2)$$

dividing Eq. (A.1) by Eq. (A.2),

$$\frac{1 - P_1}{P_1\left[1 + \frac{(1-P_1)P_2P_3}{P_1(1-P_2)(1-P_3)}\right]} > \frac{1 - P_2}{P_2\left[1 + \frac{(1-P_2)P_1P_3}{P_2(1-P_1)(1-P_3)}\right]}.$$

Multiplying both sides by $1 + \frac{P_1P_2P_3}{(1-P_1)(1-P_2)(1-P_3)} > 0$,

$$\frac{1 - P_1}{P_1} \cdot \frac{1 + \frac{P_1P_2P_3}{(1-P_1)(1-P_2)(1-P_3)}}{1 + \frac{(1-P_1)P_2P_3}{P_1(1-P_2)(1-P_3)}} > \frac{1 - P_2}{P_2} \cdot \frac{1 + \frac{P_1P_2P_3}{(1-P_1)(1-P_2)(1-P_3)}}{1 + \frac{(1-P_2)P_1P_3}{P_2(1-P_1)(1-P_3)}}.$$

Taking the logarithm on both sides,

$$\log\frac{1-P_1}{P_1}+\log\frac{1+\frac{P_1P_2P_3}{(1-P_1)(1-P_2)(1-P_3)}}{1+\frac{(1-P_1)P_2P_3}{P_1(1-P_2)(1-P_3)}}>\log\frac{1-P_2}{P_2}+\log\frac{1+\frac{P_1P_2P_3}{(1-P_1)(1-P_2)(1-P_3)}}{1+\frac{(1-P_2)P_1P_3}{P_2(1-P_1)(1-P_3)}}.$$

According to Eqs. (6.2)-(6.15), the above inequality implies that

$$b_1=a_1+\varepsilon_1>b_2=a_2+\varepsilon_2.$$

$b_2>b_3$ can be similarly proved. Therefore, when $P_1<0.5$, $P_2<0.5$, and $P_3<0.5$, then $a_1>a_2>a_3\Longrightarrow b_1>b_2>b_3$.

APPENDIX B

Proof of Proposition 3 in Chapter 6

According to the condition

$$P_3 < \frac{1}{2},$$

we have

$$P_3^2 < (1 - P_3)^2.$$

Multiplying both sides by $1 - 2P_2$,

$$P_3^2(1 - 2P_2) < (1 - P_3)^2(1 - 2P_2),$$

and

$$P_3^2[(1 - P_2)^2 - P_2^2] < (1 - P_3)^2[(1 - P_2)^2 - P_2^2].$$

Multiplying both sides by $P_1(1 - P_1)$ and rearranging the terms,

$$(1-P_2)^2(1-P_1)P_1P_3^2+(1-P_1)(1-P_3)^2P_1P_2^2 < P_1(1-P_1)(1-P_2)^2(1-P_3)^2+(1-P_1)P_1P_2^2P_3^2.$$

After further manipulation,

$$[(1 - P_1)(1 - P_3)P_2 + (1 - P_2)P_1P_3][(1 - P_1)(1 - P_2)P_3 + (1 - P_3)P_1P_2] <$$
$$[P_1(1 - P_2)(1 - P_3) + (1 - P_1)P_2P_3][(1 - P_1)(1 - P_2)(1 - P_3) + P_1P_2P_3].$$

Dividing both sides of the above inequality by $(1 - P_1)(1 - P_2)(1 - P_3)$,

$$\left[(1 - P_1)P_2 + \frac{(1 - P_2)P_1P_3}{1 - P_3}\right]\left[P_3 + \frac{(1 - P_3)P_1P_2}{(1 - P_1)(1 - P_2)}\right] <$$
$$[P_1(1 - P_2)(1 - P_3) + (1 - P_1)P_2P_3]\left[1 + \frac{P_1P_2P_3}{(1 - P_1)(1 - P_2)(1 - P_3)}\right].$$

Factoring out $(1 - P_1)P_2P_3$ from the left side, and $P_1(1 - P_2)(1 - P_3)$ from the right side,

$$(1 - P_1)P_2P_3[1 + \frac{(1 - P_2)P_1P_3}{P_2(1 - P_1)(1 - P_3)}][1 + \frac{(1 - P_3)P_1P_2}{P_3(1 - P_1)(1 - P_2)}] <$$
$$P_1(1 - P_2)(1 - P_3)[1 + \frac{(1 - P_1)P_2P_3}{P_1(1 - P_2)(1 - P_3)}][1 + \frac{P_1P_2P_3}{(1 - P_1)(1 - P_2)(1 - P_3)}]$$

Dividing both sides of the above inequality by the following factor,

$$P_2P_3[1+\frac{(1-P_2)P_1P_3}{P_2(1-P_1)(1-P_3)}][1+\frac{(1-P_3)P_1P_2}{P_3(1-P_1)(1-P_2)}]P_1[1+\frac{(1-P_1)P_2P_3}{P_1(1-P_2)(1-P_3)}],$$

we get

$$\frac{(1-P_1)}{P_1[1+\frac{(1-P_1)P_2P_3}{P_1(1-P_2)(1-P_3)}]} < \frac{(1-P_2)(1-P_3)[1+\frac{P_1P_2P_3}{(1-P_1)(1-P_2)(1-P_3)}]}{P_2P_3[1+\frac{(1-P_2)P_1P_3}{P_2(1-P_1)(1-P_3)}][1+\frac{(1-P_3)P_1P_2}{P_3(1-P_1)(1-P_2)}]}.$$

Multiplying both sides by $1+\frac{P_1P_2P_3}{(1-P_1)(1-P_2)(1-P_3)}$ and taking the logarithm operation,

$$\log\left(\frac{1-P_1}{P_1}\cdot\frac{1+\frac{P_1P_2P_3}{(1-P_1)(1-P_2)(1-P_3)}}{1+\frac{(1-P_1)P_2P_3}{P_1(1-P_2)(1-P_3)}}\right) <$$

$$\log\left(\frac{1-P_2}{P_2}\cdot\frac{1+\frac{P_1P_2P_3}{(1-P_1)(1-P_2)(1-P_3)}}{1+\frac{(1-P_2)P_1P_3}{P_2(1-P_1)(1-P_3)}}\cdot\frac{1-P_3}{P_3}\cdot\frac{1+\frac{P_1P_2P_3}{(1-P_1)(1-P_2)(1-P_3)}}{1+\frac{(1-P_3)P_1P_2}{P_3(1-P_1)(1-P_2)}}\right).$$

Again according to Eqs. (6.2)-(6.15), the above inequality implies that

$$b_1 < b_2 + b_3.$$

Similarly, we can prove that $b_2 < b_1 + b_3$, and $b_3 < b_1 + b_2$.

REFERENCES

1. V. Aalo and R. Viswanathan, "On distributed detection with correlated sensors: Two examples," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 25, no. 3, pp. 414–421, May 1989.

2. A. A. Abu-Dayya and N. C. Beaulieu, "Micro- and macrodiversity mdpsk on shadowed frequency-selective channels," *IEEE Trans. on Comm.*, vol. 43, no. 8, pp. 2334–2343, Aug. 1995.

3. N. Ansari, J. Chen, and Y. Zhang, "Distributed detection by reinforcement learning for un-equiprobable sources," *Proc. of CISS'95*, pp. 612–617, March 1995. Baltimore, MD.

4. N. Ansari, E. Hou, B. Zhu, and J. Chen, "An adaptive fusion model for distributed detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 524-531, Apr. 1996.

5. B. Bornkamp, "Uplink and downlink performance analysis of a cellular radio system using corner base stations and directional antennas," *Int. Conf. VT*, pp. 375–379, July 1995. Chicago, IL.

6. D. G. Brennan, "Linear diversity combining techniques," *Proc. of IRE*, vol. 47, pp. 1075–1102, June 1959.

7. Z. Chair and P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 22, no. 1, pp. 98–101, Jan. 1986.

8. L. F. Chang and J. C. I. Chuang, "Diversity selection using coding in a portable radio communications channel with frequency-selective fading," *IEEE JSAC*, vol. 7, no. 1, pp. 89–98, Jan. 1989.

9. Y. Chau and A. Laih, "Diversity for noncoherent DPSK direct-sequence CDMA over a shadowed Rician-fading satellite channel," *IEE Proc.–Communications*, vol. 142, no. 6, pp. 386–392, Dec. 1995.

10. Y. Chau and J. Sun, "Diversity with distributed decision combining for direct-sequence CDMA in a shadowed Rician-fading land-mobile satellite channel," *IEEE Trans. Vehicular Tech.*, vol. 45, no. 2, pp. 237–247, May 1996.

11. J. Chen and N. Ansari, "Adaptive fusion of correlated local decisions," *Proc. of CISS'96*, pp. 1071–1076, March 1996. Princeton, NJ.

12. J. Chen, N. Ansari, and Z. Siveski, "Adaptive data fusion for handoffs in cellular CDMA," submitted to *IEEE Trans. on Vehicular Tech.*

13. J. Chen, N. Ansari, and Z. Siveski, "Distributed detection for cellular CDMA," *IEE Electronics Letters*, vol. 32, no. 3, pp. 169–171, Feb. 1996.

14. J. Chen, N. Ansari, and Z. Siveski, "Performance analysis and realization of decision fusion for macroscopic diversity in cellular wireless systems," submitted to Int. Conf. on Comm. 1997. Montreal, Canada.

15. D. C. Cox, "Universal digital portable radio communications," *Proceedings of the IEEE*, vol. 75, no. 4, pp. 436–65, April 1987.

16. B. V. Dasarathy, *Decision Fusion*, IEEE Computer Society Press, Los Alamitos, CA, 1994.

17. K. Demirbas, "Maximum a *posteriori* approach to object recognition with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 3, pp. 309–313, May 1988.

18. E. Drakopoulos and C. Lee, "Optimum multi-sensor fusion of correlated local decision," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 27, no. 4, pp. 593–605, July 1991.

19. Z. Drezner, "Computation of the multivariate normal integral," *ACM Trans. on Mathematical Software*, vol. 18, no. 4, pp. 470–480, Dec. 1992.

20. W. Feller, *An Introduction to Probability Theory and Its Applications, 3rd edition, vol. 2*, Wiley, New York, NY, 1968.

21. S. S. Gupta, "Probability integrals of multivariable normal and multivariate t," *Annals of Mathematical Statistics*, vol. 34, no. 9, pp. 792–828, Sept. 1963.

22. M. H. Hassoun, *Fundamentals of Artificial Neural Networks*, The MIT Press, Cambridge, MA, 1995.

23. S. Haykin, *Adaptive Filter Theory, 2nd Edition*, Prentice Hall, Englewood Cliffs, NJ, 1991.

24. R. Hecht-Nielsen, *Neurocomputing*, Addison Wesley, Reading, MA, 1990.

25. J. A. Hertz, *Introduction to The Theory of Neural Computation*, Addison Wesley, Reading, MA, 1990.

26. J. Holtzman and A. Sampath, "Adaptive averaging methodology for handoffs in cellular systems," *IEEE Trans. Vehicular Tech.*, vol. 44, no. 1, pp. 59–66, Feb. 1995.

27. W. C. Jakes, *Microwave Mobile Communication*, J. Wiley & Sons, 1974, New York, NY, reprinted by IEEE Press, 1994.

28. M. Kam, Q. Zhu, and W. Gray, "Optimal data fusion of correlated local decisions in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 3, pp. 916–920, July 1992.

29. S. Kandala, E. Sousa, and S. Pasupathy, "Multi-user multi-sensor detectors for CDMA networks," *IEEE Trans Commun.*, vol. 43, no. 2/3/4, pp. 946–957, Feb./Mar./Apr. 1995.

30. B. Kosko, *Neural Networks and Fuzzy Systems*, Prentice Hall, Englewood Cliffs, NJ, 1992.

31. W. Lee, "Smaller cells for greater performance," *IEEE Comm. Magazine*, pp. 19–23, Nov. 1991.

32. J. M. G. Linnartz, "Site diversity in land-mobile cellular telephony network with discontinuous voice transmission," *Europian Trans. on Telecomm.*, vol. 2, no. 5, pp. 471–479, Sept.-Oct. 1991.

33. M. Minsky and S. Paperk, *Perceptron — Expanded Edition*, MIT Press, Cambridge, MA., 1988.

34. A. M. Mood, *Introduction to the Theory of Statistics*, McGraw-Hill, Inc., New York, NY, 1974.

35. W. Papen, "Uplink performance of a new macro-diversity cellular mobile radio architecture," *PIMRC*, pp. 1118–1122, 1995.

36. J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, 1995.

37. A. R. Reibman and L. W. Nolte, "Optimal detection and performance of distributed sensor systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 23, no. 1, pp. 24–30, Jan. 1987.

38. A. Reibman and L. Nolte, "Design and performance comparison of distributed detection networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 23, no. 6, pp. 789–797, Nov. 1987.

39. W. S. S. Zurbes, W. Papen and, "A new architecture for mobile radio with macroscopic diversity and overlapping cells," *PIMRC*, pp. 640–644, 1994.

40. F. Sadjadi, "Hypotheses testing in a distributed environment," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 22, no. 2, pp. 134–137, Mar. 1986.

41. R. Srinivasan, "Distributed radar detection theory," *IEE Proceedings part F*, vol. 133, no. 1, pp. pp 55–60, Feb. 1986.

42. G. Stuber and C. Kchao, "Analysis of a multiple-cell direct-sequence CDMA cellular mobile radio system," *IEEE Journal on Selected Areas in Communication*, vol. 10, no. 4, pp. 669–679, May 1992.

43. R. Tenney and J. N.R. Sandell, "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 17, no. 4, pp. 501–510, July 1981.

44. S. Thomopoulos, R. Viswanathan, and D. Bougooulias, "Optimal decision fusion in multiple sensor systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 23, no. 5, pp. 644–653, Sept. 1987.

45. A. Viterbi, *CDMA: Principles of Spread Spectrum Communication*, Addison-Wesley Publishing Co., Reading, MA, 1995.

46. A. Viterbi, A. Viterbi, K. Gilhousen, and E. Zehavi, "Soft handoff extends CDMA cell coverage and increases reverse link capacity," *IEEE JSAC*, vol. 12, no. 8, pp. 1281–1288, Oct. 1994.

47. A. Viterbi and A. Viterbi, "Erlang capacity of a power controlled CDMA system," *IEEE JSAC*, vol. 11, no. 6, pp. 892–900, June 1993.

48. E. Waltz and J. Llinas, *Multisensor Data Fusion*, Artech House, Norwood, MA, 1990.

49. P. B. Wong and D. C. Cox, "Low-complexity co-channel interference cancellation and macroscopic diversity for high capacity PCS," *Int. Conf. on Comm*, pp. 852–857, 1995.

50. J. C. W. Y. S. Yeh and S. C. Schwartz, "Outage probability in mobile telephony with directive antennas and macrodiversity," *IEEE Trans. on Vehicular Tech.*, vol. 33, no. 3, pp. 123–127, Aug. 1984.

51. J. M. Zwada, *Introduction to Artificial Neural Systems*, West Publishing Company, St. Paul, MN, 1992.