AD/A-002 254

ASW INFORMATION PROCESSING AND OPTIMAL SURVEILLANCE IN A FALSE TARGET ENVIRONMENT

H. R. Richardson

Daniel H. Wagner Associates

Prepared for:

Office of Naval Research

31 October 1973

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: A003
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER: AD/A-002254
4. TITLE (and Subtitle): ASW INFORMATION PROCESSING AND OPTIMAL SURVEILLANCE IN A FALSE TARGET ENVIRONMENT
5. TYPE OF REPORT & PERIOD COVERED: Final
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): H. R. Richardson
8. CONTRACT OR GRANT NUMBER(s): Contract No. N00014-71-C-0309
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Daniel H. Wagner, Associates, Station Square One, Paoli, Pennsylvania 19301
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Department of the Navy, Arlington, Virginia 22217
12. REPORT DATE: October 31, 1973
13. NUMBER OF PAGES: 110
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE: N/A
16. DISTRIBUTION STATEMENT (of this Report): Approved for public release; distribution unlimited.
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):
18. SUPPLEMENTARY NOTES:
19. KEY WORDS (Continue on reverse side if necessary and identify by block number):
    - ASW search and surveillance
    - information theory
    - information processing
    - monte carlo methods
    - false targets
    - moving targets
20. ABSTRACT (Continue on reverse side if necessary and identify by block number):
    See abstract page.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

ASW INFORMATION PROCESSING AND OPTIMAL SURVEILLANCE IN A FALSE TARGET ENVIRONMENT

Report to

Naval Analysis Programs
Office of Naval Research

October 31, 1973

Under Contract No. N00014-71-C-0309

by

DANIEL H. WAGNER, ASSOCIATES
Paoli, Pennsylvania

By: Henry R. Richardson
Senior Vice President

Approved: Daniel H. Wagner
President

ia

ABSTRACT

Problems of information processing and optimal surveillance in a false target environment are investigated with ASW applications in view. The information processing procedures, among other things, make use of adaptive estimation techniques in order to identify uncertain system parameters. Procedures are presented for computing real-time estimates of the target location probability distribution in realistic tactical scenarios involving moving targets and false sensor responses. The procedures are applied to a variety of illustrative examples pertaining to the processing of responses obtained from a fixed sensor field in barrier and area surveillance scenarios.

The optimal allocation of ASW search resources in a false target environment is investigated in an exploratory analysis of an idealized surveillance situation. Several allocation policies are formulated including one based upon some concepts of information theory. This "maximum information gain" policy is shown by numerical examples to have very desirable characteristics. In order to further establish the relevance of the information-theoretic approach to the surveillance problem, the latter is formulated as a type of sequential statistical experimental design problem which has been studied extensively using information-theoretic concepts.

TABLE OF CONTENTS

PREFACE

This is a report to Naval Analysis Programs, Office of Naval Research (Code 431), under Contract No. N00014-71-C-0309. It presents methods of processing information from ASW sensors in the presence of false targets and for planning surveillance actions based on such processing. The methods are presented in ways that are suitable for real-time computer assistance to ASW surveillance operations and have in fact been motivated by actual applications of this nature. Related prior applications have also included computerized assistance to search and rescue operations by the Coast Guard.

We should like to express appreciation for the sponsorship and splendid cooperation that has been given to this work from Mr. J. Randolph Simpson of Naval Analysis Programs, ONR, and through him, Mr. Robert J. Miller, Director. We further acknowledge the contributions of our colleagues Dr. Thomas L. Corwin, who was responsible for Appendix B, and Mr. Brian D. Wenocur, who performed considerable programming and computation support.

SUMMARY

This report addresses problems pertaining to ASW information processing and optimal surveillance in a false target environment. The objective is to provide useful concepts and practical approaches for answering the question, "Where is the target?". Potential applications include continuous broad localization of the target through intermittent application of ASW search at selected times and places. No ASW action other than target location is considered.

The results on ASW information processing are given in Chapter II, and the results on optimal surveillance are given in Chapter III. Chapter I provides a brief introduction, and Appendices A and B support the material presented in Chapters II and III.

The first two sections of this summary discuss Chapters II and III. The third section discusses the appendices.

ASW Information Processing

Chapter II presents methods for processing ASW information in order to compute real-time estimates of the target location probability distribution. An illustrative ASW setting is used to demonstrate the potential applications of the processing concepts, and extensive numerical results are given. The methodology is discussed in computer programming oriented language in the final section of Chapter II and in greater generality in Appendix A.

The target location probability distributions are computed by monte-carlo simulation and are expressed discretely in terms of grid cell probabilities. It is assumed for illustrative purposes in Chapter II that a fixed sensor field provides the source of real-time input to the processing system.

The term sensor response is used to indicate that a decision has been made that the sensor output contains a sufficient number of target-related cues so that the hypothesis that the target is present is preferred to the alternative hypothesis that the target is not present. A false response is a response generated by a non-target-related mechanism. The causes of false responses are dealt with from a decision-theoretic point of view in a predecessor report (reference [a]);

the present report adopts an operational point of view and focuses on overcoming the adverse effects of false targets.

The methods of Chapter II and Appendix A have been applied without false target considerations in Coast Guard search and rescue (SAR) cases (see reference [b]) and in certain ASW situations. In each instance, successful implementation of the methods has depended upon exploitation of the unique aspects of each application and construction of a mathematical model having a level of detail and realism consistent with the quality of the data and with the constraints imposed by computer memory size and computation speed. In view of this, the results in this report are intended to point the way rather than to give a comprehensive treatment which would cover all possible circumstances.

Among other things, Chapter II shows how to make use of adaptive estimation techniques (see reference [c]) in order to identify uncertain system parameters. In a sense, one begins with a family of models and the "correct" model is identified adaptively on the basis of observational information.

For example, the particular stochastic process underlying target motion is not assumed to be known. Rather, several possible processes ("scenarios") are postulated and each one is given an a priori probability ("credence"). The processing system revises the credences in accordance with the input sensor responses. Those scenarios which are most in agreement with the sensor responses eventually develop the highest credences.

Similarly, the single-sensor, single-glimpse probabilities of detection and false alarm are treated as unknown parameters. They are, however, related through a known ROC relationship. The probability of detection is initially assumed to be a random variable with a uniform distribution between known limits and the processing system adaptively revises this distribution in accordance with the sensor responses.

Table S-1 indicates illustrative results of the adaptive estimation procedures. In all cases, the scenarios for target motion are considered a priori to be equally likely. The true detection probability $P_D = .8$ is not known; it is assumed that P_D is a particular value of a random variable $\tilde{P}_D$ which is uniformly distributed in the interval from .5 to .9. The expected value of this prior probability distribution is .7. The estimated detection probabilities given in Table S-1 after incorporating sensor field responses are the expected values of the posterior distributions for $\tilde{P}_D$. The processing algorithms make use of the entire distribution for $\tilde{P}_D$, however, and not just the expected value.

Table S-1(a) pertains to a target patrolling station and is based on the results shown in Table II-1 of Chapter II. The correct scenario for target motion in this example is Scenario 2, and Table S-1(a) shows that the credence associated with this scenario rises to .95 as a result of processing all the sensor response information for four field glimpses. The estimated detection probability is .77, compared to the actual value of .8.

TABLE S-1

SUMMARY OF ADAPTIVE ESTIMATION RESULTS

Notes: (1) This table indicates illustrative results of adaptive estimation of target scenario and detection probability.

(2) The correct target scenario is circled, and in all cases the true single-sensor, single-glimpse detection probability is $P_D = .8$.

(a) Target Patrolling Station (see Table II-1)

	Scenario Credences			Estimated Single-Sensor, Single-Glimpse Detection Probability (true value is .8)
	1	(2)	3	
No Sensor Information Used	.33	.33	.34	.70
All Sensor Information Used (96 hours into mission--4 field glimpses)	.00	.95	.05	.77

(b) Target in Transit (see Table II-2)

	Scenario Credences					Estimated Single-Sensor, Single-Glimpse Detection Probability (true value is .8)
	(1)	(2)	3	4	5	
No Sensor Information Used	.2	.2	.2	.2	.2	.70
All Sensor Information Used (48 hours into mission--3 field glimpses)	.75	.18	.01	.03	.03	.73
All Sensor Information Used (96 hours into mission--5 field glimpses)	.33	.60	.00	.02	.05	.78

(c) Target Out of Grid Area (see Table II-3)

	Scenario Credences					Estimated Single-Sensor Single-Glimpse Detection Probability (true value is .8)
	1	2	3	4	(5)	
No Sensor Information Used	.2	.2	.2	.2	.2	.70
All Sensor Information Used (96 hours into mission--5 field glimpses)	.02	.06	.07	.08	.77	.83

Table S-1(b) pertains to a target in transit and is based on the results shown in Table III-2. In this illustration, the target is assumed to be a late Scenario 1 or an early Scenario 2. That is, the true target's position falls midway between the mean positions prescribed by Scenarios 1 and 2. Table S-1(b) shows that as a result of processing five field glimpses, the total credence associated with the two closest scenarios is .93 and the estimated detection probability is .78.

Table S-1(c) pertains to a target which is out of the grid area, that is, during the period of observation considered there has been no transit by the target through the area of interest. The true scenario in this illustration is Scenario 5 and as a result of processing five field glimpses, the credence associated with this scenario rises to .77. The estimated detection probability is .83.

Figure S-1 shows selected probability distributions for the cases considered in Table S-1(a) and S-1(b). It should be noted that only probability which falls within the grid is shown and thus the numbers need not add to one. The probability distributions on the left are based upon the a priori scenario and make use of no sensor information. The probability distributions on the right are based upon use of all sensor information available.

It is evident from Figure S-1 that processing of sensor response information by the methods described in Chapter II and Appendix A results in considerable concentration of the target location probability distribution.

Optimal Surveillance

Chapter III is addressed to optimal surveillance in a false target environment. An exploratory analysis is presented for the purpose of gauging the effectiveness of a surveillance policy based upon maximization of the expected information gain in the target location probability distribution. Here, the term information is used in the technical sense of communications theory (see, for example, reference [d]).

The concepts presented in this chapter are expressed in terms of rather idealized assumptions and further development is required before application can be made to large-scale practical problems. The objective of this chapter is to demonstrate through examples that the concepts of information theory are relevant to certain kinds of search and surveillance problems, particularly when false targets are considered.

It is assumed that the performance of an ASW search system is idealized in terms of a $J \times J$ response array (J is the number of search cells),

$$R = \Big(R(i,j)\Big),$$

FIGURE S-1

THE INFLUENCE OF SENSOR RESPONSES ON THE TARGET LOCATION PROBABILITY DISTRIBUTION

Note: Only probability inside grid is shown and thus numbers need not add to one.

(a) Target Patrolling Station (See Figures II-5 and II-6)

No Sensor Information Used
(96 hours into mission--0 field glimpses)

	1	2	3	4	5	6
A						
B		.02	.05	.02		
C	.01	.13	.15	.05	.04	
D		.04	.09	.18	.08	.01
E		.01	.01	.05	.03	
F						

Target

All Sensor Information Used
(96 hours into mission--4 field glimpses)

	1	2	3	4	5	6
A						
B						
C				.01		
D				.49	.48	
E						
F						

Target

(b) Target In Transit (See Figures II-9 and II-10)

No Sensor Information Used
(48 hours into mission--0 field glimpses)

	1	2	3	4	5	6
A		.02	.05	.03	.02	
B		.01	.03	.02		
C		.02	.05	.04	.02	
D		.01	.03	.03	.01	
E		.02	.03	.07	.02	
F				.02		

Target

All Sensor Information Used
(48 hours into mission--3 field glimpses)

	1	2	3	4	5	6
A				.01		
B						
C			.01			
D			.36	.36		
E			.12			
F		.02	.03			

Target

where $R(i, j)$ is the probability that an increment of search effort applied to the j^{th} cell will result in a response given that the target is located in the i^{th} cell. The desired modification of the target location probability distribution is accomplished by the sequential application of search in selected cells. The surveillance is carried out in stages, and at the end of each stage one is required to estimate which cell contains the target. For all policies examined, the selection rule at the end of a stage is to pick the cell having the highest target location probability distribution based upon evaluation of the search results. The cell searched during a stage, however, may or may not be the highest probability cell depending upon the policy.

The measures of effectiveness are the probabilities $S(k)$ of correctly selecting the cell containing the target at the end of the k^{th} stage for $k = 1, 2, \ldots$. A surveillance policy which maximizes $S(k)$ for some particular k is referred to as a k-optimal surveillance policy and a surveillance policy which maximizes $S(k)$ for all $k \geq 1$ is referred to as a uniformly optimal surveillance policy. Within the framework of our analysis, k-optimal policies are guaranteed to exist since the set of all possible policies is finite; however, existence of uniformly optimal surveillance policies is not guaranteed.

When target motion is considered, it is assumed for illustration to be Markovian. This is not essential, however, and target motion could equally as well be described by the non-Markovian processes considered in Chapter II and Appendix A. The $J \times J$ transition matrix D for the Markov process is assumed for illustration to be given by, for some $0 \leq \delta \leq 1$,

$$D = \begin{pmatrix} 1-\frac{(J-1)}{J}\delta, & \frac{\delta}{J}, & \ldots, & \frac{\delta}{J} \\ \frac{\delta}{J}, & 1-\frac{(J-1)}{J}\delta, & \frac{\delta}{J}, \ldots, & \frac{\delta}{J} \\ \vdots & & & \vdots \\ \frac{\delta}{J}, & \ldots, & \frac{\delta}{J}, & 1-\frac{(J-1)}{J}\delta \end{pmatrix}$$

This transition matrix depends upon a single parameter δ, here referred to as the dispersion constant. The initial distribution for the process is denoted d.

The problem of finding an optimal surveillance policy can be formulated in terms of a stochastic control problem, and this is discussed briefly in Chapter III.

Visualizing the problem in this way, it seems apparent that k-optimal plans may be found by dynamic programming, but we do not develop these solutions in this report. Our interest is in the entire time behavior of the success function S rather than the value of the function at some fixed stage.

Four surveillance policies are examined in Chapter III using a variety of assumptions about the false target environment and about the prior target location probability distribution and target motion characteristics. To do this, let (for a given stage) $P_B(j)$ be the before-search probability that the target is located in the j^{th} cell for $1 \leq j \leq J$. Let $p_A(r, i, j)$ be the conditional after-search probability that the target is located in the i^{th} cell given that the j^{th} cell was searched and result r was obtained. Here, $r = 1$ indicates a target-like response and $r = 0$ indicates a non-target-like response. The four policies examined are as follows:

<u>I. The optimal single-stage look-ahead policy.</u> The optimal single-stage look-ahead policy is to search in the cell which, based upon the estimated vector P_B, maximizes the probability of correctly selecting the target cell at the end of a single stage. This is a generalization of the optimal whereabouts plan formulated in reference [e] for searches without false responses. If

$$B(j) = \max\{P_B(i) R(i, j) : 1 \leq i \leq J\} + \max\{P_B(i) [1 - R(i, j)] : 1 \leq i \leq J\},$$

then it is shown in Chapter III that the optimal single-stage look-ahead policy is to search in cell j^* for which

$$B(j^*) \geq B(j) \quad \text{for } 1 \leq j \leq J.$$

<u>II. The maximum information-gain policy.</u> The maximum information-gain policy is to search in the cell which maximizes the expected information content (or, equivalently, minimizes the expected entropy) of the posterior after-search target location probability distribution.

For any discrete probability distribution P over J cells, the entropy H(P) is defined by

$$H(P) = - \sum_{j=1}^{J} P(j) \ln P(j).$$

The expected entropy U(j) of the posterior target location probability distribution given search in cell j is shown in Chapter III to be given by

$$U(j) = -\sum_{i=1}^{J} p_B(i) \{R(i,j) \ln p_A(1,i,j) + (1-R(i,j)) \ln p_A(0,i,j)\}.$$

The maximum information-gain policy is to search in any cell j* for which

$$U(j^*) \leq U(j) \quad \text{for } 1 \leq j \leq J.$$

<u>III. The highest probability cell policy</u>. The highest probability cell policy is to search in the cell with the highest probability, that is, to search in any cell j* for which

$$p_B(j^*) \geq p_B(j) \quad \text{for } 1 \leq j \leq J.$$

Once p_B is determined from the search results of the previous stage, this plan does not make further use of the response matrix R.

<u>IV. The uniform surveillance policy</u>. The uniform surveillance policy is to search systematically through all search cells in a fixed rotation, that is, one searches the J cells in order and then repeats as often as required. This plan does not make use of the target location probability distribution nor of the response matrix.

Figure S-2 illustrates the behavior of the above-mentioned surveillance policies in one of the cases (Case I(a)) considered in Chapter III. The target is assumed to be stationary with a uniform prior distribution, i.e., d(1) = .33, d(2) = .33, and d(3) = .34. For all cells, if the target is in the cell searched, then the probability of response is .8. The probability of false response is .7 in the first cell and the probability of false response is .1 in the second and third cells. This means that very little information is gained by a search in the first cell since the probabilities of correct response and false response are nearly equal.

FIGURE S-2

COMPARISON OF SURVEILLANCE POLICIES

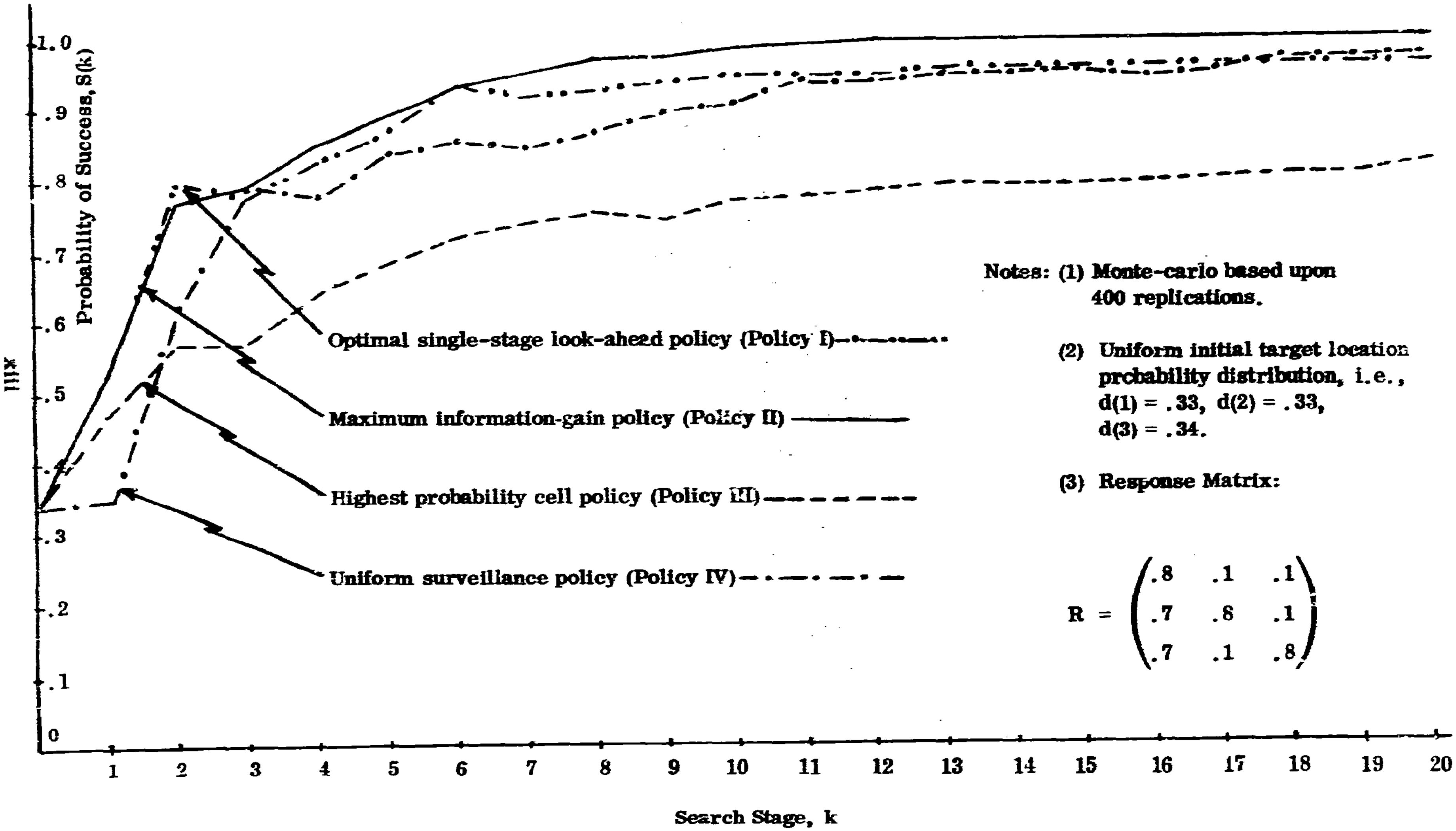


Figure S-2 indicates that in the earlier stages of search, there is little difference between the maximum information-gain policy and the optimal single-stage look-ahead policy. Asymptotically (i.e., for large k), the maximum information-gain policy appears to have a slight advantage. The uniform surveillance policy (Policy IV) also does well in this example, but the highest probability cell policy (Policy III) is not particularly attractive.

Figure S-3 shows the influence of target motion on probability of success for the maximum information-gain policy. In this case (Case I(b)), the response matrix R is the same as in Figure S-2, but the prior distribution is non-uniform with $d(1) = .75$, $d(2) = .15$, and $d(3) = .10$. The three examples shown correspond to values of the dispersion constant, $\delta = 0$, $\delta = .3$, and $\delta = 1$. The case where $\delta = 0$ corresponds to no target motion, and, consequently, this curve is the same as that given in Figure S-2 for this policy. The case where $\delta = 1$ corresponds to complete dispersion of the target location probability distribution to a uniform distribution at each stage. The first transition of the Markov process is made at the end of the first stage. Therefore, S(1) is identical for all three values of the dispersion constant. When $0 < \delta \leq 1$, the curves do not appear to approach 1 asymptotically. In these cases, it appears that equilibrium is reached for large values of k in the sense that the information gained by search is balanced by the information lost by dispersive target motion.

The principal conclusion of Chapter III is that the maximum information-gain policy appears to have very desirable characteristics in the idealized surveillance scenario considered. In all cases considered, it is the best or nearly the best of all the plans considered. Moreover, for each alternative policy, there is at least one case given where the maximum information-gain policy is much better. This conclusion appears to be at variance with some previous investigations into the value of information theory in search problems; these other investigations are reviewed briefly in the final section of Chapter III.

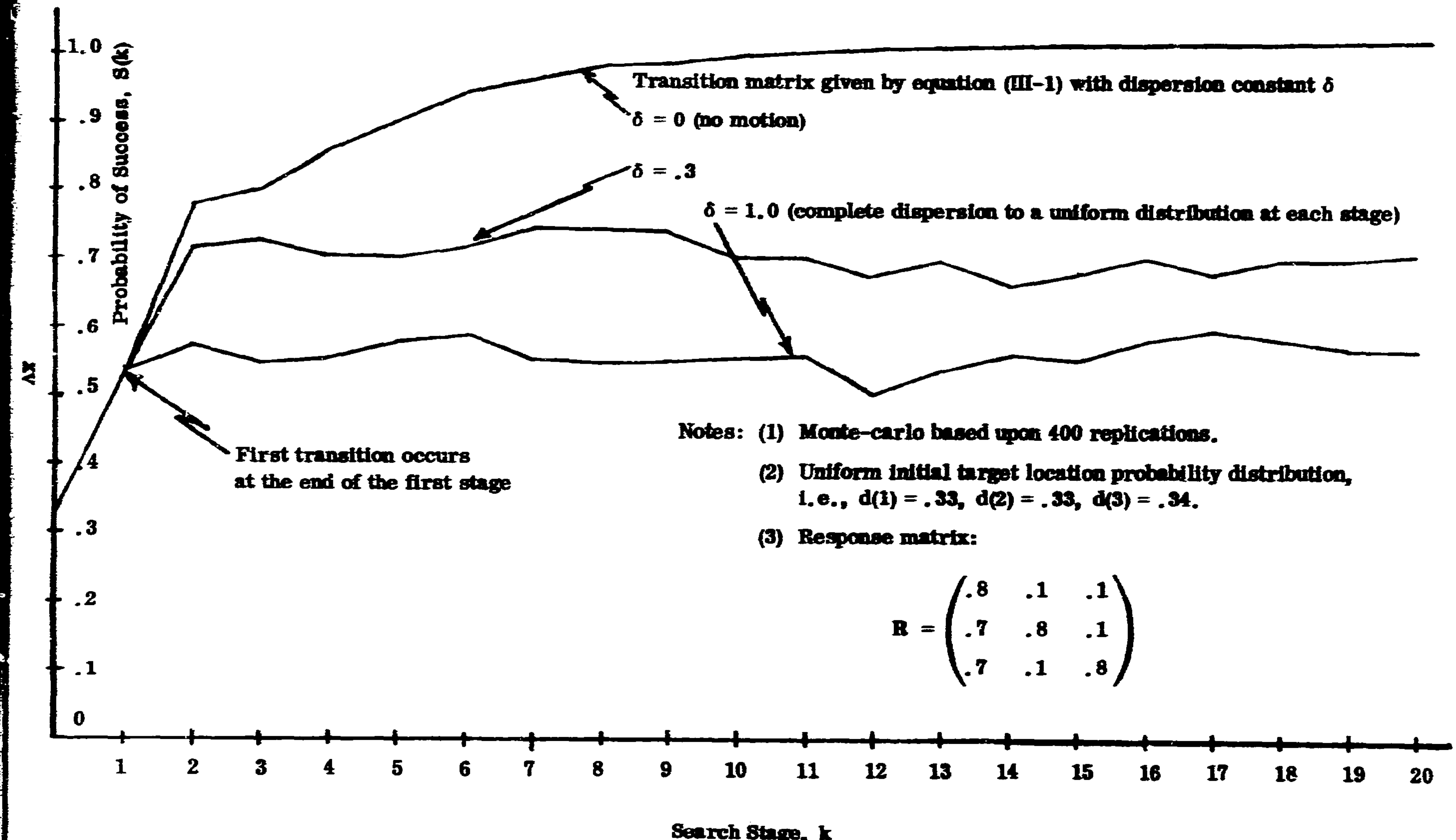
Appendices

Appendix A provides a generalized treatment of the information processing concepts described and applied in Chapter II. Knowledge of the mathematical structure of these information processing procedures makes it possible to carry out deeper investigations of their characteristics and scope. An understanding of Appendix A, however, is not required in order to undertake the development of new processing systems; the last section of Chapter II should suffice for this purpose.

Appendix B formulates the search and surveillance problem as a statistical sequential experimental design problem. The purpose of this formulation is to suggest a theoretical framework for applying information theoretic concepts to

FIGURE 8-3

THE INFLUENCE OF TARGET MOTION ON PROBABILITY OF SUCCESS FOR THE MAXIMUM INFORMATION-GAIN POLICY



surveillance problems. In particular, the maximum information-gain policy of Chapter II is shown to correspond to Lindley's approach (see reference [f]) in sequential experimental design. It is also shown that the problem of which cell to search at each stage of a surveillance operation may be viewed as a game between the search planner and nature in which the payoff to the search planner is measured in terms of the information he gains about the true state of nature.

ASW INFORMATION PROCESSING AND OPTIMAL SURVEILLANCE IN A FALSE TARGET ENVIRONMENT

CHAPTER I

INTRODUCTION

This report addresses problems pertaining to ASW information processing and optimal surveillance in a false target environment. The objective is to provide useful concepts and practical approaches for answering the question, "Where is the target?". Potential applications include central-site processing of data from fixed surveillance systems, VP mission planning and analysis at Tactical Support Centers (TSCs), ocean surveillance, and processing of diverse kinds of ASW-related information by computerized command and control systems (ASWCCS, WWMCCS, etc.).

In this report, the term sensor is used to denote the entire sensing system consisting of transponder, processor, display, and human operator. A target-like sensor response results from a decision based upon the inputs to the sensing system in favor of the hypothesis that the target is present as opposed to the alternative hypothesis that the target is not present. A target-like response may be generated by the target (a true response) or by some other non-target-related mechanism (a false response).

A predecessor report (reference [a]) deals extensively with the causes of false responses and, among other things, provides quantitative models for including false responses in ASW computer simulations.

The present report assumes that the occurrence of false responses is an unavoidable operational fact of life and focuses on the problem of what to do about them.

We are interested in utilizing the information provided by sensor responses for the purpose of making target location predictions. In parts of an operating area where there are few false response stimuli, such as those produced by shipping or biological activity, a target-like response conveys considerable information about target presence. In other areas which abound in false response stimuli, a single target-like response has less meaning and importance.

For our purposes, search is defined as the act of acquiring response/no response data from the sensors. In most treatments of search theory, the objective of search is target detection, i.e., achieving a state where the target's location (e.g., a cell in a search grid) can be stated with absolute certainty. Unfortunately, this state is seldom reached with non-visual sensors because of the possible occurrence of false responses. Thus, new approaches are required to deal realistically with these situations.

In this report, in fact, the detection state is not observable. That is, in this report it is assumed that the decision maker can never state "We have detected the target." He can only become increasingly confident that the information provided by his sensors is consistent with a particular target location or motion hypothesis.

Even within the narrow confines of ASW search and surveillance, there are a wide variety of tactical situations which might arise and which might involve many different types of ASW units, sensors, and systems. In order to treat this diversity, we have decided to emphasize concepts rather than details. Our intent is to show the potential usefulness of certain ideas rather than to present detailed algorithms for the implementation of these ideas in specific situations.

Chapter II discusses methods for centralized processing of diverse kinds of sensor data and general intelligence. These methods have been applied without false target considerations in Coast Guard search and rescue (SAR) cases (see reference [b]) and in certain ASW situations. The discussion is based upon an idealized tactical setting where a fixed distributed field of sensors provides the response data. These responses and subjective a priori information about target information are input to the processing system; the output of the processing system provides the answer to the question, "Where is the target?" in the form of target location probability maps. Illustrations are given for the cases of targets patrolling on station, targets in transit, and targets out of the area of interest entirely. In the latter case, all sensor information is false response information.

In Chapter II the subjective input takes the form of scenarios for target motion together with associated credences. The "weighted scenario" idea was introduced by Dr. John P. Craven during the Mediterranean H-bomb search in 1966 and used to develop an a priori probability target location distribution for that operation. The weighted scenario approach was used subsequently in the 1968 search for the submarine Scorpion (see reference [g]) and is presently incorporated in the operational computer-assisted search and rescue planning (CASP) system of the Coast Guard.

The methods illustrated in Chapter II also permit the input of probability distributions rather than single-valued estimates, for parameters whose values are uncertain; the "true" values of these parameters are estimated from the sensor observation data concurrent with the determination of the target location probability distributions. Appendix A supports the material in Chapter II with a more general and abstract discussion of the information processing concepts.

Chapter III is concerned with optimal utilization of the information given by the target location probability distributions, and the analysis in this chapter is intended primarily to demonstrate the potential applications of information theory to ASW surveillance in a false target environment. In this kind of environment, sensor responses do not necessarily indicate target presence, but they do provide a certain amount of information. Our results indicate that this information may be quantified, analyzed mathematically, and usefully applied in terms of the concepts of information theory.

The tactical setting considered in Chapter III is an idealized ASW surveillance situation in which one is interested in finding the sequential assignment of ASW search which will maximize the number of times that the target's position is correctly specified over an extended period of time. Four surveillance policies (i.e., sequential allocations of search effort) are compared using monte-carlo simulation. The policy which maximizes the expected information gain in the posterior target location probability distribution is found to provide the best overall results in the cases examined.

Previous studies (in particular, references [h], [i], and [j]) of the connections between search theory and information theory have reached negative conclusions. These previous studies are reviewed in the final section of Chapter III and some reasons for the apparent disagreement are offered.

Information theoretic approaches have been used extensively in statistics (see, for example, reference [k]); Appendix B relates these statistical methods to the surveillance problem from the point of view of sequential experimental design and hypothesis testing.

CHAPTER II

ASW INFORMATION PROCESSING IN A FALSE TARGET ENVIRONMENT

This chapter presents procedures for processing ASW information in a false target environment for the purpose of predicting target location. The procedures are computer-oriented and suited for use in command centers which have access to diverse kinds of ASW sensor data and intelligence. Questions pertaining to the utilization of the target location predictions are deferred to Chapter III.

The methods indicated in this chapter are Bayesian and are addressed primarily to answering the question, "Where is the target?". The results are displayed in terms of target location probability maps which are based upon subjective target-mission scenarios and upon observed sensor response data. The maps express the target location probability distributions in terms of grid-cell probabilities. Other useful results such as the probability distributions for target course and speed could be displayed if desired but are not treated in this report.

Each ASW situation has its own peculiarities, and discussion of the information processing methods in a way which would cover all contingencies would, it is believed, obscure the basic principles. Therefore, our main purpose is to demonstrate the potential usefulness of the concepts in terms of specific examples and to provide a mathematical framework for further applications.

Successful implementation of the methods will depend to a large measure upon one's ability to exploit the specifics of each application (target mission objectives and patterns of operation, own systems characteristics, crew proficiency, etc.) and to construct a mathematical model having a level of detail and realism consistent with both the data quality and the constraints imposed by computer memory size and computation speed.

As mentioned above, the information processing methods discussed in this chapter are Bayesian. Briefly described, one begins by generating a large collection of "constructs," $e_1, \ldots, e_N$. Each construct specifies a complete target track as well as any parameters of the mathematical model which are not assumed to be known exactly. For each construct e_n, there is specified a prior probability p_n that the n^{th} construct is correct. The prior probabilities reflect the validity of the constructs before any information is obtained from the various ASW sensors. Usually, $p_n = 1/N$ when the constructs are generated by monte-carlo simulation.

Preceding page blank

Sensor information is used to update the prior probabilities for the constructs in the form of a posterior distribution. This is done as follows. Let

$$q_n = \Pr\left\{\begin{array}{l}\text{sensor response patterns observed} \\ \text{throughout the time period of interest}\end{array} \middle| \begin{array}{c}\text{the } n^{th} \text{ construct} \\ \text{is correct}\end{array}\right\}$$

If p_n' denotes the posterior probability for the n^{th} construct, then according to Bayes's formula

$$p_n' = \frac{q_n p_n}{\Sigma_{m=1}^{N} q_m p_m} \quad \text{for } n = 1, \ldots, N.$$

The first section illustrates the information processing methods by applying them to hypothetical ASW situations. The second section presents the details of the mathematical procedures used to compute the illustrations. Appendix A provides a generalization of the information processing methods to more general situations.

Illustrative ASW Applications

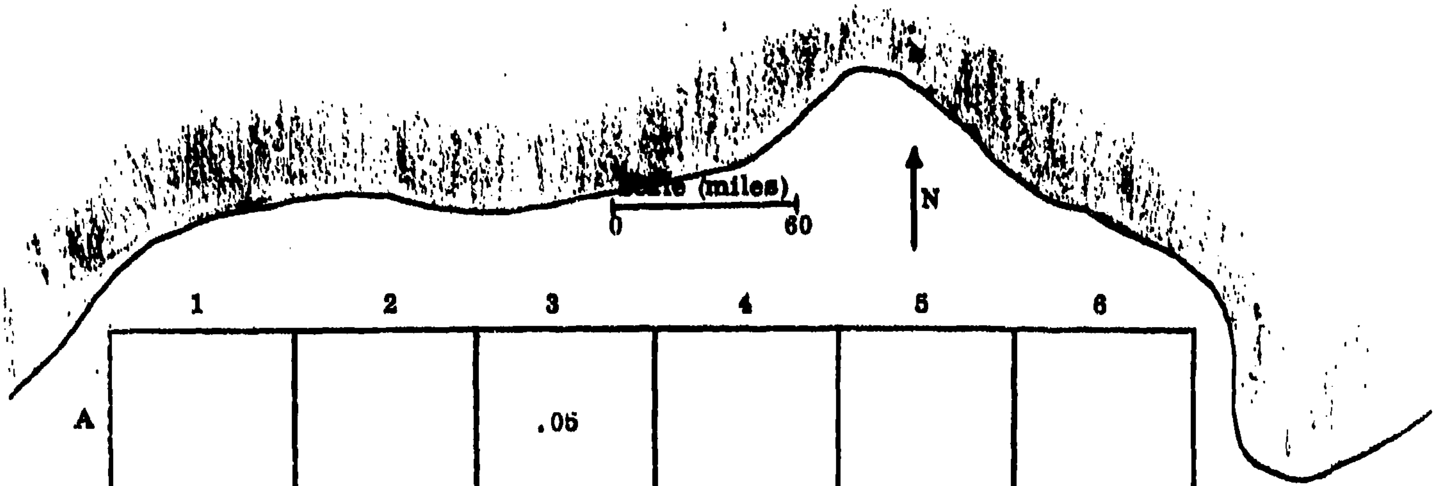
This section illustrates the results of applying certain methods for processing subjective target motion scenario information and ASW sensor response data in order to obtain estimates of the target location probability distribution and of certain other parameters of interest. The methods themselves are postponed to the second section. The target location probability distribution permits one to determine the probability that the target is contained within specific geographical regions. These probability distributions are of central importance in ASW.

In this report, the target location probability distribution is expressed in terms of grid-cell probabilities as illustrated in Figure II-1. Charts such as Figure II-1 are often referred to as target location probability maps. In the case shown, there is a 20% chance that the target is in cell C-3 and a 70% chance that the target is in the region covered by cells B-3, C-2, C-3, C-4, and D-3. Target location probabilities associated with other regions may be obtained by summing the appropriate probabilities.

Although tactical use of the target location probability distribution is not discussed in this chapter, a comment on contact investigation is in order. The usual objective of contact investigation is to detect and further localize the target. To a large extent the target location probability distribution consolidates all of the

FIGURE II-1

ILLUSTRATIVE TARGET LOCATION PROBABILITY DISTRIBUTION



	1	2	3	4	5	6
A			.05			
B		.05	.15	.05		
C		.10	.20	.10		
D		.05	.15	.05		
E			.05			
F						

relevant information needed to pursue this objective. The probability map displays the information pertaining to sensor contacts combined with the equally important information pertaining to target-mission objectives and patterns of target operation. In many cases, therefore, it is better to investigate areas associated with updated target location probabilities rather than to investigate points associated with the individual contacts. The usefulness of the latter investigation usually decreases rapidly as "time late" increases.

The first subsection below introduces the sensor response assumptions. The three subsections which follow the first provide the illustrative numerical examples. These are separately addressed to the cases where the target is (1) patrolling station, (2) in transit, and (3) out of area.

<u>Sensor assumptions.</u> Figure II-2 shows the sensor field which will be used in all of the examples given in this section. The sensors are arranged in a fixed rectangular array with 60-mile spacing between rows and columns.

The term "sensor response" will be used to indicate that a decision has been made that the sensor output contains a sufficient number of target-related cues so that the hypothesis that the target is present is preferred to the alternative hypothesis that the target is not present. A decision-theoretic discussion of this determination is given in detail in Chapter IV of reference [a], and we will not be concerned further with these details.

Sensor response decisions might be made by an individual in charge of a sensor team or, perhaps, by the programming logic of an automatic classification device. The information processing methodology presented in this chapter may be particularly useful in the latter case because the programming of an automatic classification device requires the explicit statement of classification decision rules. Such explicit rules are much easier to deal with analytically than are the less explicit rules underlying human decision making.

A "detection" is defined to be a sensor response caused by the target and a "false response" is defined to be a sensor response caused by something other than the target.

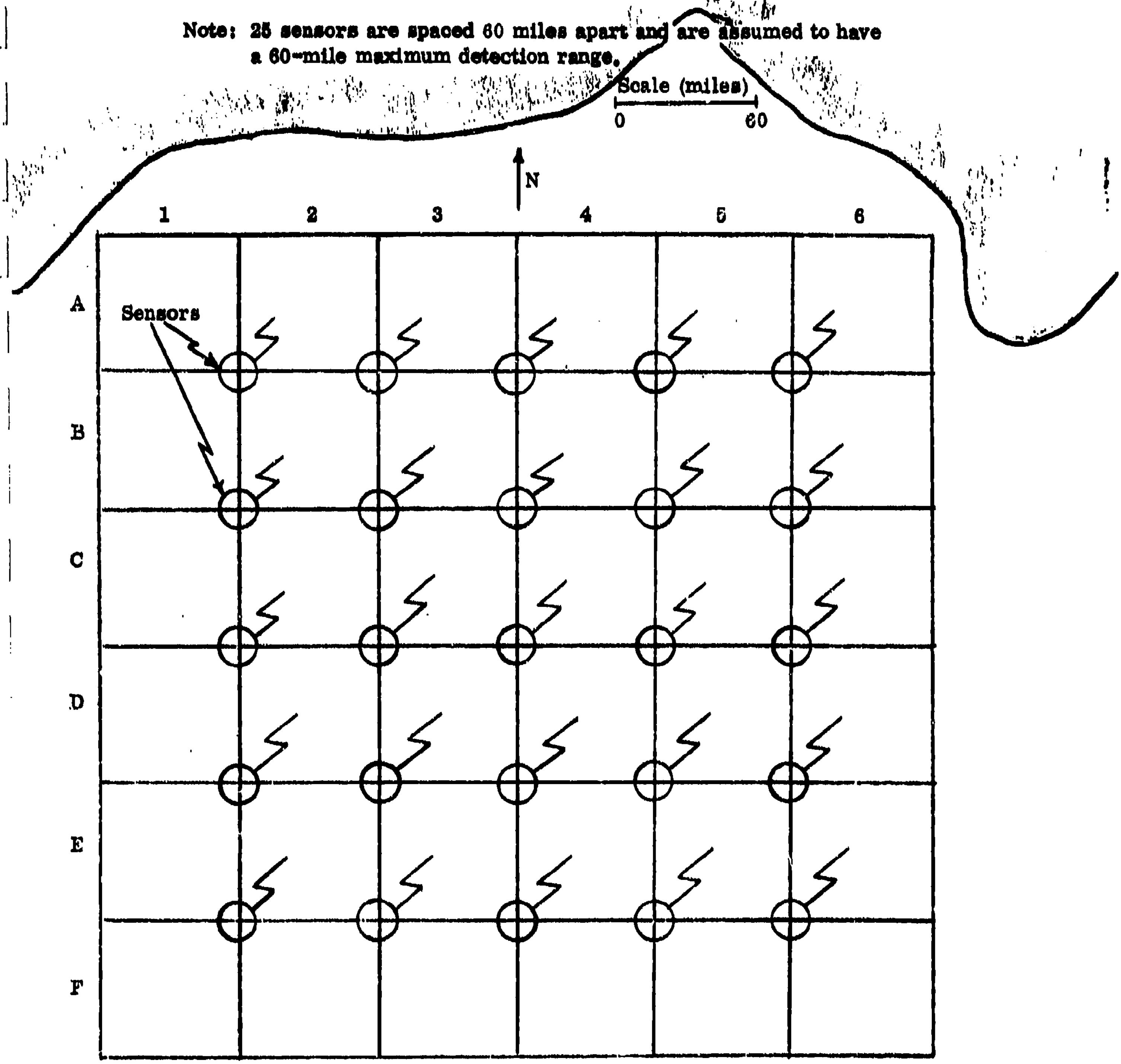
The distributed sensors are monitored at the end of 24-hour intervals. Each monitoring event is treated as a "single glimpse." Continuous field observations could also be modeled but would require more complex algorithms than those developed to compute the examples in this section.

It is assumed that each sensor has a maximum detection range of 60 miles and that the single-sensor, single-glimpse detection probability is $P_D = .8$ if the target comes within this range of a sensor. Probability of detection is assumed zero outside of 60 miles. The single-sensor, single-glimpse probability of false response is $P_A = .3$ regardless of target location. Thus, if the target is within

FIGURE II-2

ASW-SENSOR LOCATIONS

Note: 25 sensors are spaced 60 miles apart and are assumed to have a 60-mile maximum detection range.



60 miles of a sensor, then the probability of response is $1-(1-P_D)(1-P_A) = .86$, and if the target is not within 60 miles of a sensor, then the probability of response is $P_A = .3$.

It is assumed for illustration that all sensor responses are statistically independent in space and time. More complex assumptions could be made if desired in a real application. The values $P_D = .8$ and $P_A = .3$ are assumed unknown to the information processor. What is known, however, is that detection probability and false-response probability are related by an ROC (receiver operating characteristic) relationship

$$P_A = f(P_D)$$

where, for some fixed $\alpha > 0$,

$$f(p) = p^{\alpha} \quad \text{for } 0 \leq p \leq 1.$$

Note that P_A increases with P_D and that $P_A = 0$ when $P_D = 0$ and $P_A = 1$ when $P_D = 1$. The true values of P_D and P_A will be estimated from the operationally derived data as part of the processing.

Any function relating P_A and P_D could be used without significantly increasing the complexity of the processing algorithms. In fact, it would not be difficult to devise an algorithm which would permit postulation of an entire family of possible ROC relationships when there is uncertainty as to which relationship is correct. The correct relationship could then be inferred from the operationally derived data.

The above assumptions are made in order to illustrate the information processing ideas within the framework of a simple and easily understood mathematical model. They are not necessarily recommended for real-world applications. Alternative models for detection are provided, for example, in references [l], [m], [n], [o], and [p]; reference [a] provides an hierarchy of decision-theoretic models which treat detection and classification in a unified manner.

False responses result from complex interactions involving, among other things, the sensor system, the environment, and human factors (see reference [a]). At the present time, these interactions are not well understood and many of the factors which are involved (e.g., command attitudes and individual motivation) are not physically observable or measurable. Any estimate of false-response probability, therefore, could be in error by a significant amount. For this reason, it is important to develop information processing procedures which do not require exact knowledge of false-response probabilities and which are adaptive in the sense that initial estimates of these probabilities can be modified by observed sensor responses.

In order to reflect initial uncertainty about target characteristics, sensor capabilities, and environmental conditions, therefore, we shall assume a uniform probability distribution (known to the information processor) for P_D on the interval between .5 and .9. The expected value of this probability distribution is .7 and P_A is determined from P_D by means of the ROC function. (See references [q] and [r] for related analyses when the target is stationary and sensor capabilities are not known precisely.)

Example 1 -- target patrolling station. This example applies to the case of a target patrolling station. It is assumed that scenarios can be postulated for target motion based upon past observations of similar targets or knowledge of the present target's mission objectives. Associated with each scenario is a credence which expresses the scenario's relative plausibility. In the present example, a scenario specifies a probability distribution for the target's location at equally spaced points in time. The target is assumed to move along legs with constant course and speed between leg endpoints. Monte-carlo procedures are used to obtain a large number of sample target tracks for each scenario specified (more details are given in the second section). The number of tracks generated for each scenario is proportional to the associated credence. A particular target track is generated by randomly drawing the endpoints of each track leg from the specified endpoint probability distributions.

Figure II-3 presents the scenarios chosen for this example. Scenario 1 with credence .33 describes a target patrolling in a clockwise direction beginning in the south and moving west, then north, and then east. All track-leg endpoint probability distributions are assumed to be circular normal with 30-mile standard deviations. Scenario 2 with credence .33 describes a target which also is patrolling in a clockwise direction, but beginning in the north and moving east and then south and west. The endpoint probability distributions are also normal with 30-mile standard deviations. Scenario 3 with credence .34 describes a target which is patrolling in the center of the area without a regular pattern of motion. For scenario 3, all track-leg endpoint probability distributions have identical normal distributions with 60-mile standard deviations.

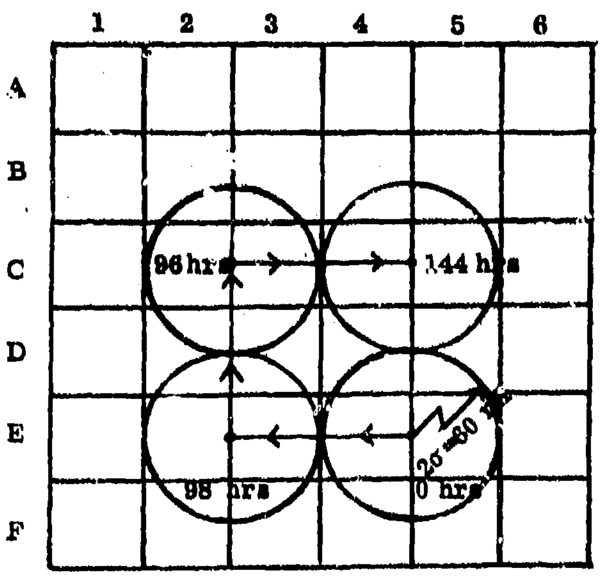
Figure II-4 shows the time history of responses from the distributed field simulated by a single replication of monte carlo. The target is assumed to follow Scenario 2 and its position as a function of time is also shown in Figure II-4. The positions were chosen to coincide with the means of the Scenario 2 distributions. A 60-mile radius circle indicating sensor detection range is drawn about the target's position so that the responses outside this circle (necessarily false) may easily be identified.

FIGURE II-3

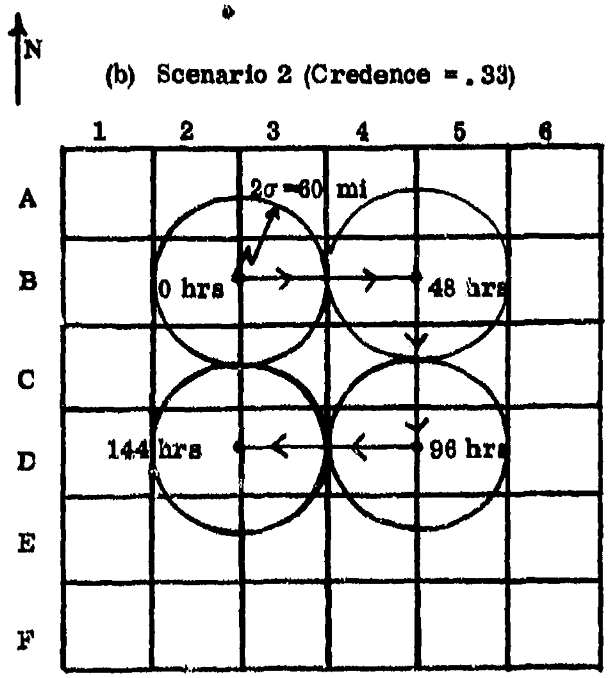
SCENARIOS FOR TARGET MOTION
EXAMPLE 1 (TARGET PATROLLING STATION)

Note: Circles indicate the 2σ-uncertainty in the circular normal distribution for target position at the endpoints of each leg.

(a) Scenario 1 (Credence = .33)



(b) Scenario 2 (Credence = .33)



(c) Scenario 3 (Credence = .34)
Note: All track endpoints have the same probability distribution.

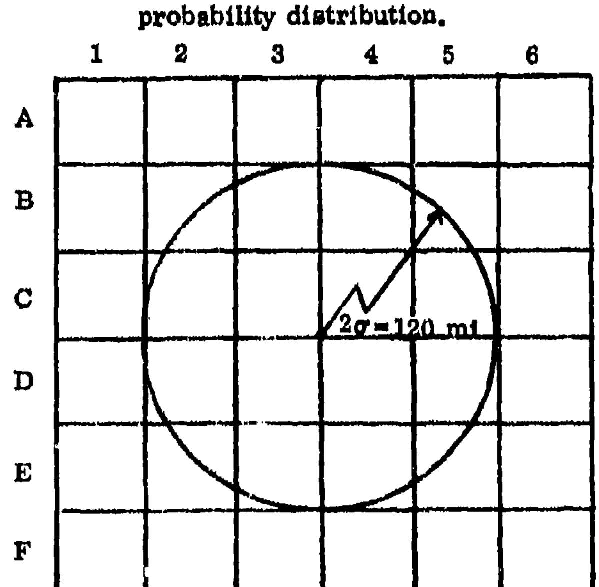
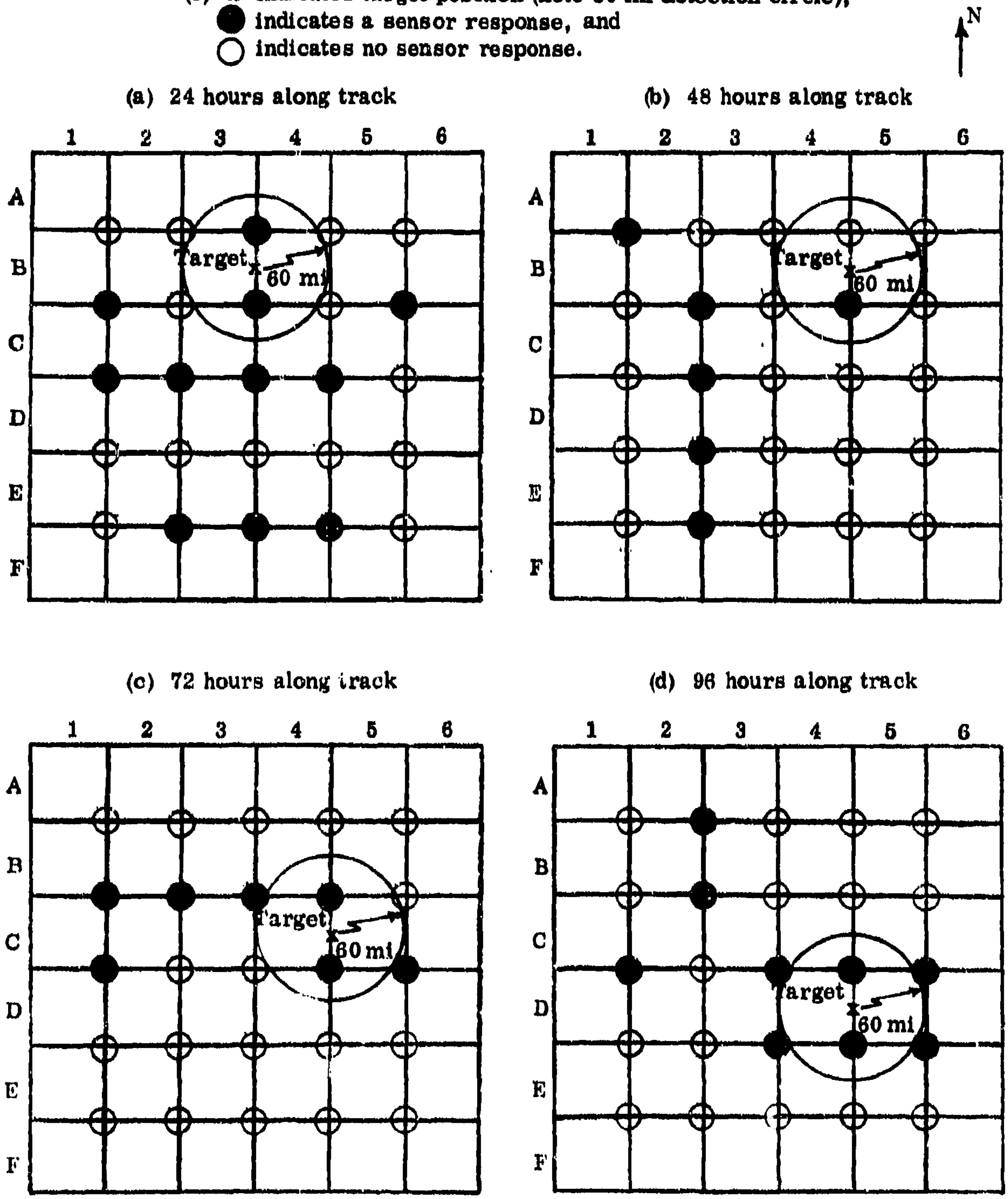


FIGURE II-4

THE HISTORY OF SENSOR RESPONSES
EXAMPLE 1 (TARGET PATROLLING STATION)

Notes: (1) Target follows Scenario 2 of Example 1.
(2) Detection probability is .8 and false-response probability is .3.
(3) x indicates target position (note 60 mi detection circle),
● indicates a sensor response, and
○ indicates no sensor response.



Figures II-5 and II-6 show the target location probability distributions before and after processing the sensor response patterns shown in Figure II-4 (500 monte-carlo replications* were used to produce these distributions). Figure II-5 shows the target location probability distributions based only upon the weighted scenarios; no sensor response information is incorporated. Figure II-6 shows the distributions which result from incorporating sensor response information. Figure II-5(a) shows the target location probability distribution for the target 24 hours along its track based solely upon the scenario formulations with no incorporation of the sensor response patterns. Note that the target actually lies on the boundary between cells B-3 and B-4 and that the sum of the probabilities in these two cells is 27%.

Figure II-6(a) shows the target location probability distribution for the same time as Figure II-5(a) but updated by incorporation of the sensor response patterns shown in Figure II-4(a). Note that the sum of probabilities in the cells B-3 and B-4, which include the target, remains about the same (28%) but that the probability distribution has become more concentrated. It should be noted from Figure II-4(a) that there were 9 false responses from the 23 sensors beyond detection range of the target. This is somewhat higher than the 6 or 7 false responses expected based upon the assumed value of $P_A = .3$ for the false-response probability.

Figure II-5(b) shows the target location probability distribution for the target at 48 hours along its track based solely upon the scenario formulations with no incorporation of any sensor response patterns. The sum of probabilities in the cells B-4 and B-5 containing the target is 22%.

Figure II-6(b) shows the updated target location probability distribution at 48 hours along the track, incorporating the sensor response patterns shown in Figure II-4(a) and in Figure II-4(b). The sensor response pattern given by Figure II-4(b) is the result of very "bad luck." Many false responses were obtained in the areas occupied by targets following Scenarios 1 and 3 while at the same time few responses were obtained in the area occupied by targets following Scenario 2. The actual target (following Scenario 2) was detected only once out of two opportunities.

As a result, the sum of probabilities in the cells B-4 and B-5 containing the target decreases to 3%. Action based on the results at this stage would not have much chance of success.

Figure II-5(c) shows the target location probability distribution for the target at 72 hours along its track based solely upon the scenario formulations. The sum of probabilities in cells C-4 and C-5 containing the target is 33%.

---

* For operational real-time applications, a much larger number of replications is suggested. In past utilization of similar systems, 2,000 to 10,000 replications have been employed routinely.

Figure II-6(c) shows the corresponding updated target location probability distribution incorporating all the sensor response patterns up to and including those shown in Figure II-4(c). The sum of probabilities in cells C-4 and C-5 is now seen to increase dramatically to 85%. Apparently, a sufficient number of patterns have been processed at this point for the updated target location distribution to begin to converge on the target's actual location.

Finally, Figures II-5(d) and II-6(d) provide the before and after comparisons corresponding to the target at 96 hours along its track. Without use of sensor response patterns, the sum of the probabilities in cells D-4 and D-5 containing the target is 26% as given in Figure II-5(d). After use of all sensor response patterns shown in Figure II-4, the sum of the probabilities in cells D-4 and D-5 is 97%.

Table II-1 shows the influence of the sensor responses on the scenario credences and the mean detection probability. Recall that initially the scenario credences were assumed equal; the sensor detection probability was assumed uniformly distributed between .5 and .9 with mean value of .7. Processing of the sensor response patterns at 24 hours and 48 hours decreases the credence associated with Scenario 2 from .330 to .211 and .044, respectively. The results improve as the response patterns for 72 hours and 96 hours are incorporated. The updated weight for Scenario 2, the actual scenario, rises to .952 following incorporation of the sensor responses obtained at 96 hours.

The mean detection probability $\hat{P}_D$ before processing is .7. In the present example, the actual but unknown detection probability is P_D = .8. After processing all the sensor responses, the updated mean detection probability is .77.

Thus, in spite of an unknown and relatively high false-response probability, the processing algorithms produce (in this example) a very accurate indication of the true target scenario and single-sensor detection probability. Moreover, after processing the sensor response patterns, the target location probability distribution becomes quite concentrated about the true location of the target.

Example 2 -- target in transit. This example considers the problem of localizing a target as it transits through an area covered by the distributed sensor field. The target (if it shows up) is expected to begin its transit through the area between time 0 and time 72 hours, but the exact time of transit is unknown. It is desired to use the sensor response patterns to detect the target's presence in the area and to localize it as it moves through.

Figure II-7 presents the scenarios formulated for this example. Once again, the location of the target at the endpoint of each leg is specified by a normal probability distribution. The distributions are elongated in the east-west direction, however, in order to represent the uncertainty in the target's location across a "front." The standard deviation in the east-west direction is 60 miles and the standard deviation in the north-south direction is 30 miles.

FIGURE II-5

TARGET LOCATION PROBABILITY DISTRIBUTIONS
EXAMPLE 1 (TARGET PATROLLING STATION) - NO SENSOR INFORMATION USED

Notes: (1) Target follows Scenario 2 of Example 1.
(2) Detection probability is .8 and false-response probability is .3.
(3) x indicates target position (note 60 mi detection circle).

↑ N

(a) 24 hours

	1	2	3	4	5	6
A		Target		.02		
B		.01	.12	x .15	.01	
C		.01	.09	.08	.01	
D		.01	.09	.07	.01	
E			.17	.12		
F			.02	.01		

(b) 48 hours

	1	2	3	4	5	6
A			Target	.02	.01	
B		.01	.03	.12	x .10	.01
C	.01	.03	.04	.06	.03	
D	.01	.06	.07	.04	.01	
E		.12	.11	.03		
F		.03	.04			

(c) 72 hours

	1	2	3	4	5	6
A						
B			.01	.02	.01	Target
C		.02	.07	.18	x .15	
D		.17	.20	.07	.05	
E		.01	.03	.02		
F						

(d) 96 hours

	1	2	3	4	5	6
A						
B		.02	.05	.02		
C	.01	.13	.15	.05	.04	Target
D		.04	.09	.18	x .08	.01
E		.01	.01	.05	.03	
F						

FIGURE II-6

TARGET LOCATION PROBABILITY DISTRIBUTIONS
EXAMPLE 1 (TARGET PATROLLING STATION) - ALL SENSOR INFORMATION USED

Notes: (1) Target follows Scenario 2 of Example 1.
(2) Detection probability is .8 and false-response probability is .3.
(3) x indicates target position (note 60 mi detection circle).

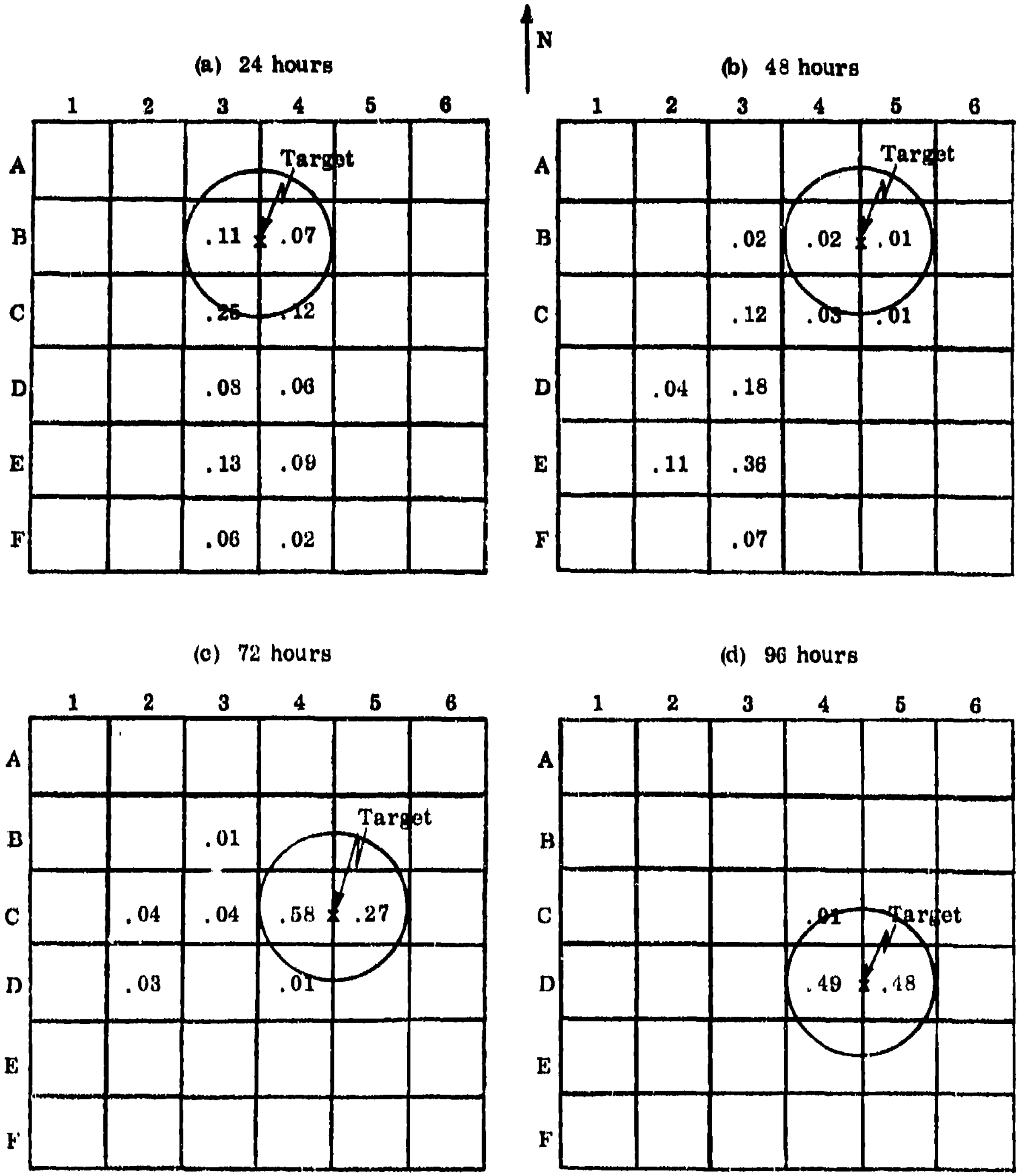


TABLE II-1

THE INFLUENCE OF SENSOR RESPONSES ON ESTIMATED PARAMETER VALUES EXAMPLE 1 (TARGET PATROLLING STATION)

Notes: (1) Target follows Scenario 2 of Example 1.

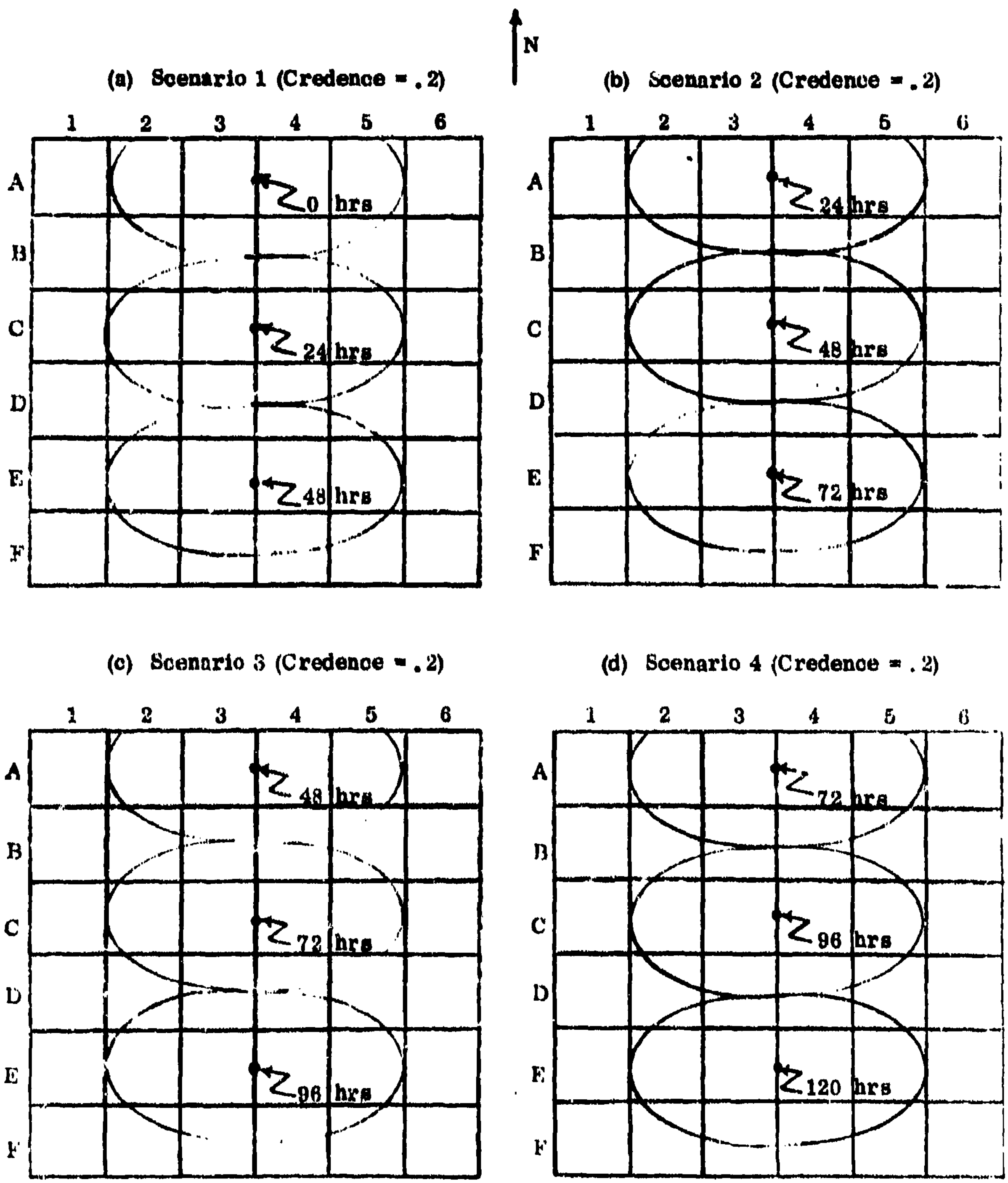
(2) True single-sensor, single-glimpse detection probability is .8 and false-response probability is .3.

		Scenario Credences			Mean of Single-Sensor, Single-Glimpse Detection Probability Distribution (true value is .8)
		Scenario: 1	2	3	
Initial Assumptions		.330	.330	.340	.700
Time of Field Response	24 hrs	.299	.211	.490	.823
	48 hrs	.525	.044	.431	.811
	72 hrs	.036	.712	.252	calculation not available
	96 hrs	.001	.952	.047	.770

FIGURE II-7

SCENARIOS FOR TARGET MOTION
EXAMPLES 2 AND 3 (TARGET IN TRANSIT)

Notes: (1) Ellipses indicate the 2σ-uncertainty in target position at the endpoints of each leg.
(2) Scenario 5 (Credence = .2) corresponds to no target transit during the time period of interest.



Scenarios 1 through 4 differ only in the assumed time of entry into the area, i.e., Scenarios 1 through 4 are based, respectively, on the target reaching the midpoint of the first row of cells at 0 hours, 24 hours, 48 hours, and 72 hours. Scenario 5 (not shown) is included to cover the contingency that there is no target transit during the time of interest. The initial credences for all five scenarios are equal.

In certain cases, the processing algorithms will produce good results even when the actual target motion does not conform to any of the scenarios specified. In order to demonstrate this fact, the target is assumed actually to begin penetration at 12 hours. This is midway between the assumptions of Scenario 1 and Scenario 2.

Figure II-8 shows the target's actual position and the sensor response patterns at 24-hour intervals. The sensor response assumptions are the same as those used in Example 1.

Figures II-9 and II-10 are based on 500 monte-carlo replications.

Figure II-9 shows the target location probability distributions when no sensor response patterns are processed by the system. Probabilities associated with locations outside of the grid are not shown. These probability distributions are based solely upon the initial scenario formulations. Note that the actual target's position lies within a cell with probability .03 throughout the transit. In order to demonstrate the flexibility of the algorithms, the computation of this example was based upon the assumption that the endpoint probability distributions are correlated so that the simulated target tracks through the area will be straight lines (no zig zags). This accounts for the fact that the target location probability distributions in Figure II-9 have the appearance of a single distribution sliding through the area.

Figure II-10 shows the target location probability distributions resulting from processing all sensor response patterns. Note that the target is located in cells having relatively large probabilities and that the probability distributions are much more concentrated than was the case in Figure II-9. It is also of interest to contrast the probability that a transit has begun (given in the notes corresponding to each time period) with the initial probabilities based on the scenarios only (given by the first general note). Once the target actually penetrates the area, these probabilities are substantially higher than the corresponding probabilities based upon the initial scenario assumptions alone. For example, at 24 hours, the probability is .40 that a transit has begun based upon the scenarios only. The corresponding probability making use of the sensor responses is .92.

It is also interesting to contrast Figure II-10(e) with Figure II-9(e). The target has completed transit of the area at this time; this is quite apparent in Figure II-10(e) which shows only 2% probability of the target being in the area.

Table II-2 shows the influence of the sensor responses on the scenario credences and the mean detection probability. Recalling that the target enters the grid area midway between the times specified by Scenarios 1 and 2, we see that at times corresponding to 24 hours and thereafter, Scenarios 1 and 2 together account for more than 90% of the total scenario probability. The sum of initial credences for these scenarios is .40.

As in Example 1, the mean of the detection probability distribution appears to be converging towards the true value .8.

Example 3 -- target out of grid area. This example is based upon the same scenarios as Example 2 but corresponds to the case where no target penetrates the area, i.e., where Scenario 5 is the correct scenario. As in Examples 1 and 2, 500 monte-carlo replications are used.

Figure II-11 shows the patterns of sensor responses which, under the present assumptions, are all false. No target location probability distributions were computed for this example.

Table II-3 shows the influence of the sensor responses on the scenario credences and the mean of the detection probability distribution. The shaded area in the table corresponds to scenarios specifying that the target has not yet entered the area. Note that as more sensor response patterns are processed, the probability tends to shift towards the "shaded" region and that at the end of 96 hours the largest scenario credence is associated with Scenario 5--the correct scenario.

As in Examples 1 and 2, the mean of the detection probability distribution appears to be converging towards the correct value of .8. Use of the ROC curve permits detection probability to be estimated from false-response data when the target is not in the area.

Information Processing Procedures

This section describes the information processing procedures used to obtain the results given in the examples in the preceding section. A more general treatment is given in Appendix A. Our purpose here is to explain the concepts in terms of the simple model used in the preceding section so that the reader may construct suitable models for other applications.

The processing system consists of two information input files, SCENE and DETECT, two state information files, UFILE and WGHT, and four computer programs, START, MAP, TRANS, and OBSERV, which operate on the state information files. These files and computer programs are discussed in the

FIGURE II-8

TIME HISTORY OF SENSOR RESPONSES
EXAMPLE 2 (TARGET IN TRANSIT)

Notes: (1) Target is a 12-hour late Scenario 1 or a 12-hour early Scenario 2 of Example 2.
(2) Detection probability is .8 and false-response probability is .3.
(3) x indicates target position (note 60 mi. detection circle),
● indicates a sensor response, and
○ indicates no sensor response.

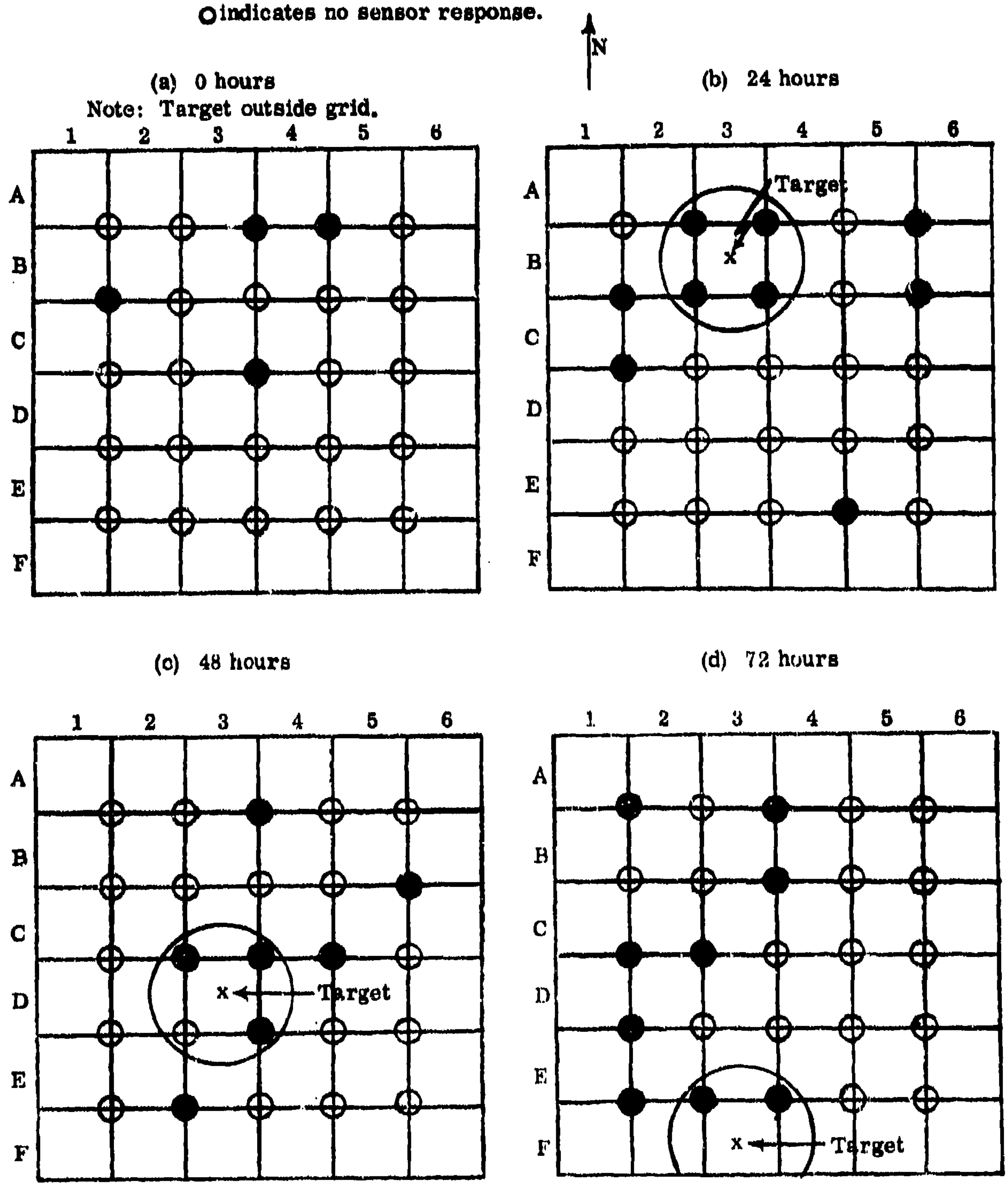


FIGURE II- 8 (continued)

(e) 96 hours
Note: Target outside grid.

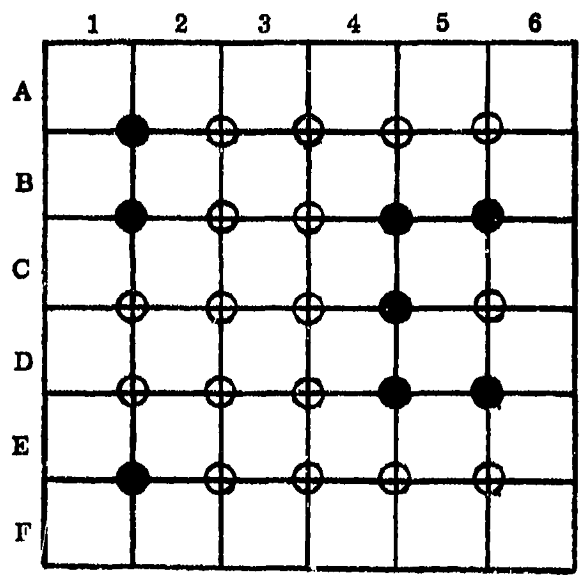


FIGURE II-9

TARGET LOCATION PROBABILITY DISTRIBUTIONS
EXAMPLE 2 (TARGET IN TRANSIT) - NO SENSOR INFORMATION USED

Notes: (1) Assumed prior probabilities that target transit has begun:

Time (hrs):	0	24	48	72	96
Probability:	.20	.40	.60	.80	.80

(2) Only probabilities within the grid are shown.

(3) x indicates target position.

N

(a) 0 hours

Note: Target is north of grid.

	1	2	3	4	5	6
A		.02	.03	.07	.02	
B				.02		
C						
D						
E						
F						

(b) 24 hours

	1	2	3	4	5	6
A		.02	[illegible]	.04	.02	
B	Target →	.01	x .03	.03	.01	
C		.02	.05	.07	.02	
D				.02		
E						
F						

(c) 48 hours

	1	2	3	4	5	6
A		.02	.05	.03	.02	
B		.01	.03	.02		
C		.02	[illegible]	.04	.02	
D	Target →	.01	x .03	.03	.01	
E		.02	[illegible]	.07	.02	
F				.02		

(d) 72 hours

	1	2	3	4	5	6
A		.01	.03	.04	.02	
B		.01	.03	.02		
C		.02	.05	.03	.02	
D		.01	.03	.02		
E		.02	[illegible]	.04	.02	
F	Target →	.01	x .03	.03	.01	

FIGURE II-9 (continued)

(e) 96 hours
Note: Target is south of grid.

	1	2	3	4	5	6
A						
B		.01		.01		
C		.01	.03	.04	.02	
D		.01	.03	.02		
E		.02	.05	.03	.02	
F		.01	.03	.02		

FIGURE II-10

TARGET LOCATION PROBABILITY DISTRIBUTIONS
EXAMPLE 2 (TARGET IN TRANSIT) - ALL SENSOR INFORMATION USED

Notes: (1) Assumed prior probabilities that target transit has begun:

Time (hrs):	0	24	48	72	96
Probability:	.20	.40	.60	.80	.80

(2) Only probabilities within the grid are shown.

(3) x indicates target position.

N ↑

(a) 0 hours

Notes: (1) Target is north of grid.
(2) Estimated probability that transit has begun is .73.

	1	2	3	4	5	6
A			.01	.55	.02	
B				.13		
C						
D						
E						
F						

(b) 24 hours

Note: Estimated probability that transit has begun is .92.

	1	2	3	4	5	6
A			.04			
B		.11	x .64	.02		
C		.07	.03			
D						
E						
F						

(c) 48 hours

Note: Estimated probability that transit has begun is .94.

	1	2	3	4	5	6
A				.01		
B						
C			.01			
D			x .36	.36		
E			.12			
F		.02	.03			

(d) 72 hours

Note: Estimated probability that transit has begun is .97.

	1	2	3	4	5	6
A			.04			
B						
C						
D	.07	.04				
E		.09	.06			
F		.13	x .24	.06		

FIGURE II- 10 (continued)

(e) 96 hours

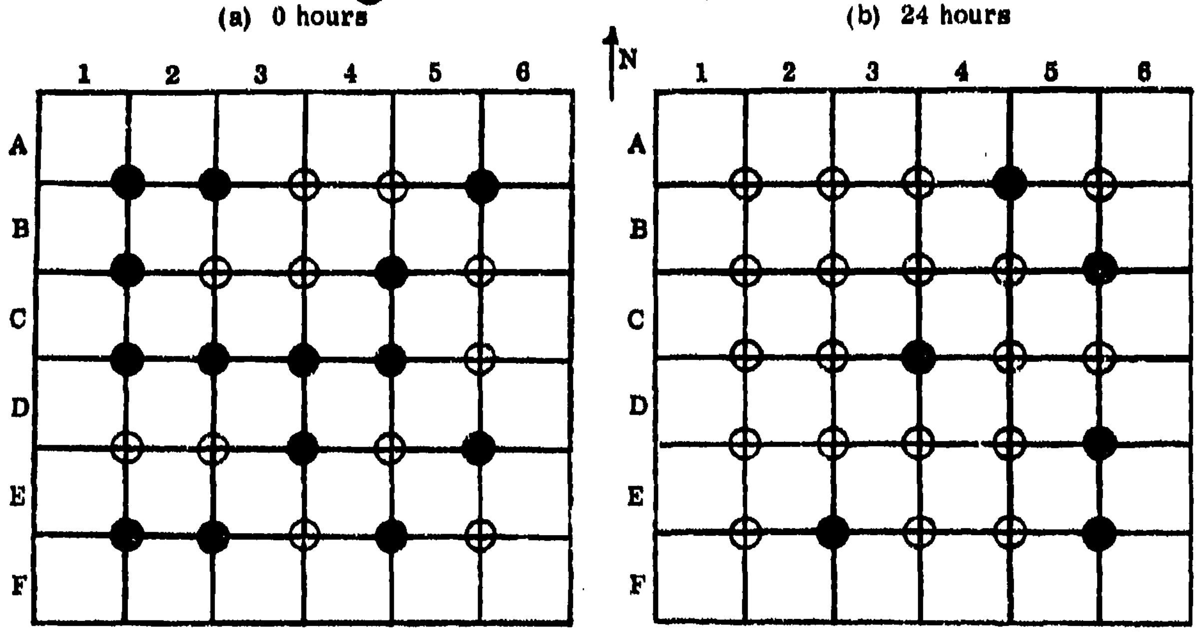
Notes: (1) Target is south of grid.
(2) Estimated probability that transit has begun is .95.

	1	2	3	4	5	6
A						
B						
C						
D						
E						
F	.01		.01			

FIGURE II-11

TIME HISTORY OF SENSOR RESPONSES
EXAMPLE 3 (TARGET OUT OF GRID AREA)

Notes: (1) Target is out of grid area covered by sensors.
(2) All responses are false. The false response probability is .3.
(3) ● indicates a sensor response, and
○ indicates no sensor response.



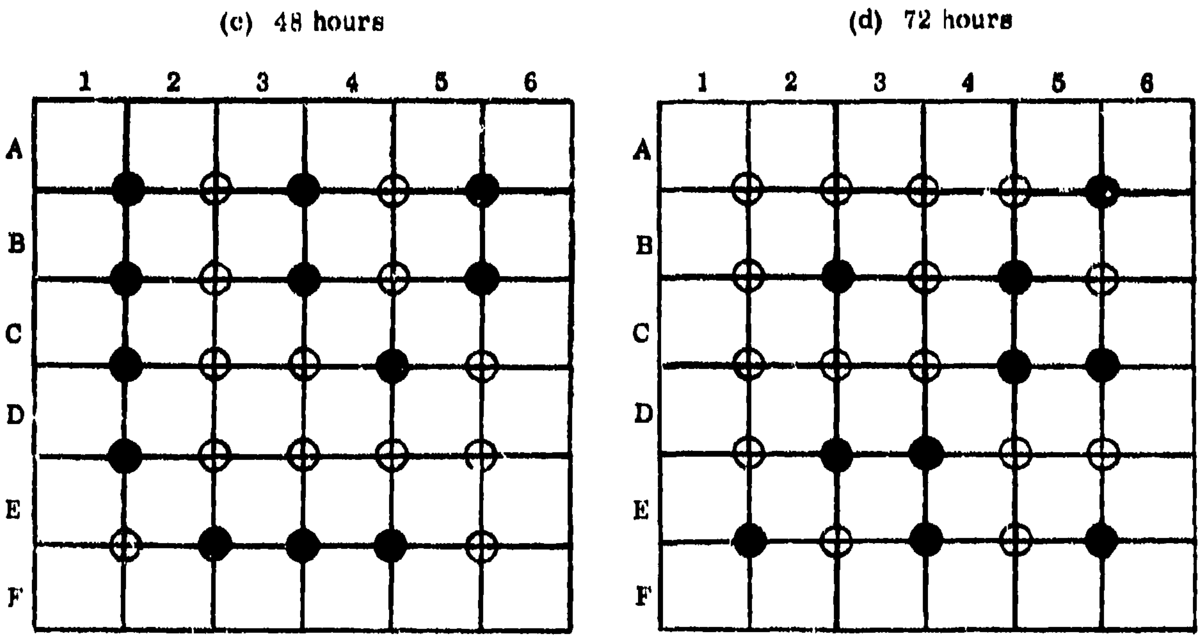


FIGURE II-11 (continued)

(e) 96 hours

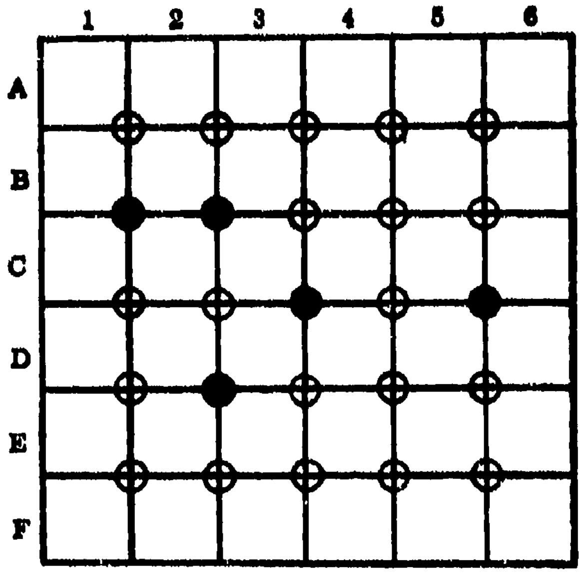


TABLE II-2

THE INFLUENCE OF SENSOR RESPONSES ON ESTIMATED PARAMETER VALUES EXAMPLE 2 (TARGET IN TRANSIT)

Notes: (1) Target is a late Scenario 1 or an early Scenario 2 of Example 2.

(2) True single-sensor, single-glimpse detection probability is .8 and false-response probability is .3.

		Scenario Credences					Mean of Single-Sensor, Single-Glimpse Detection Probability Distribution (true value is .8)
	Scenario:	1	2	3	4	5	
Initial Assumptions		.20	.20	.20	.20	.20	.70
Time of Field Response	0 hrs	.73	.07	.05	.07	.08	.67
	24 hrs	.85	.07	.01	.04	.03	.72
	48 hrs	.75	.18	.01	.03	.03	.73
	72 hrs	.45	.45	.00	.06	.04	.77
	96 hrs	.33	.60	.00	.02	.05	.78

TABLE II-3

THE INFLUENCE OF SENSOR RESPONSES ON ESTIMATED PARAMETER VALUES EXAMPLE 3 (TARGET OUT OF GRID AREA)

Notes: (1) Target is out of grid area (Scenario 5 of Example 3). All sensor responses are false responses.

(2) True single-sensor, single-glimpse detection probability is .8 and false-response probability is .3.

(3) Shading indicates scenarios placing target out of area at the specified times.

	Scenario Credences — Scenario: 1	2	3	4	5	Mean of Single-Sensor, Single-Glimpse Detection Probability Distribution (true value is .8)
Initial Assumptions	.20	.20	.20	.20	.20	.70
Time of Field Response: 0 hrs	.20	.17	.33	.19	.11	.87
24 hrs	.02	.12	.31	.30	.25	.83
48 hrs	.00	.04	.37	.33	.26	.85
72 hrs	.00	.03	.41	.17	.39	.84
96 hrs	.02	.06	.07	.08	.77	.83

following subsections. A processing system flow chart is provided in Figure II-12. The file UFILE contains the "constructs" mentioned in the heuristic description given at the beginning of the chapter. The file WGHT contains numbers ("weights") which are proportional to the posterior probabilities of the constructs.

Files SCENE and DETECT. All scenario information is stored in the information input file SCENE. This information consists of the value of the credence c_i for the i^{th} scenario for all $1 \leq i \leq I$, the time δ specified for the target to complete each of the K single legs, and the parameters for the track-leg endpoint probability distributions $\Delta_i(k)$ for $0 \leq k \leq K$ and $1 \leq i \leq I$. Here, $k\delta$ indicates the endpoint time and i indicates the scenario.

The information input file DETECT provides the known parameters characterizing the detection mechanism (i.e., the bounds on $\tilde{P}_D$ and the parameters of the ROC function.

Files UFILE and WGHT. The file UFILE contains N_r records which provide the samples from the monte-carlo simulation of target position and other parameters. The contents of UFILE vary with time and, therefore, it is convenient to let UFILE(t) denote the contents of UFILE at simulation time t. Each of the N_r records in UFILE(t) contains statistical sample values for the following random variables:

$\tilde{z}_1(t)$ = target's latitude (degrees) at time t

$\tilde{z}_2(t)$ = target's longitude (degrees) at time t

$\tilde{s}_1(t)$ = target's velocity component (degrees/hour) in the north-south direction

$\tilde{s}_2(t)$ = target's velocity component (degrees/hour) in the east-west direction

$\tilde{i}$ = target's scenario index

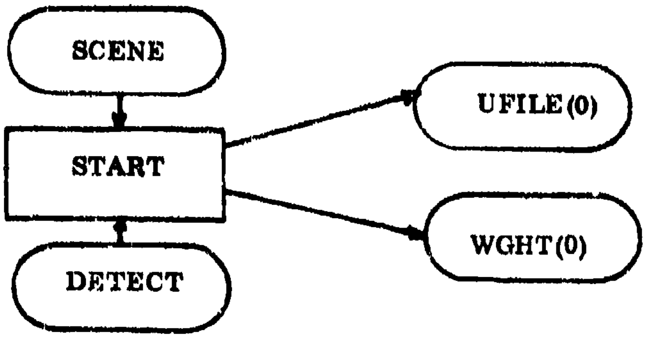
$\tilde{P}_D$ = target's probability of being detected by a single sensor on a single glimpse given that target is within detection range of the sensor.

WGHT(t) denotes the contents of file WGHT at simulation time t and contains N_r records, each providing the "weight" for the corresponding record of UFILE(t). The weights are calculated using Bayes's formula and indicate the extent to which the records of UFILE(t) are consistent with the observed sensor field responses.

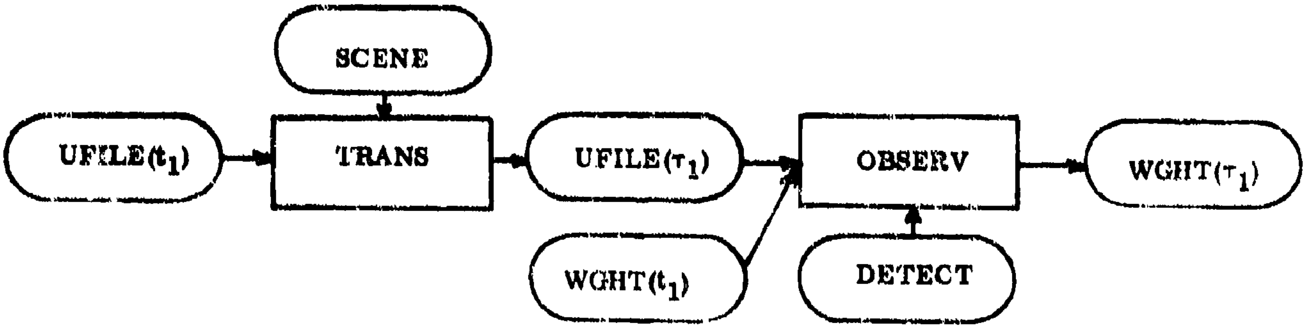
FIGURE II-12

PROCESSING SYSTEM FLOW CHART

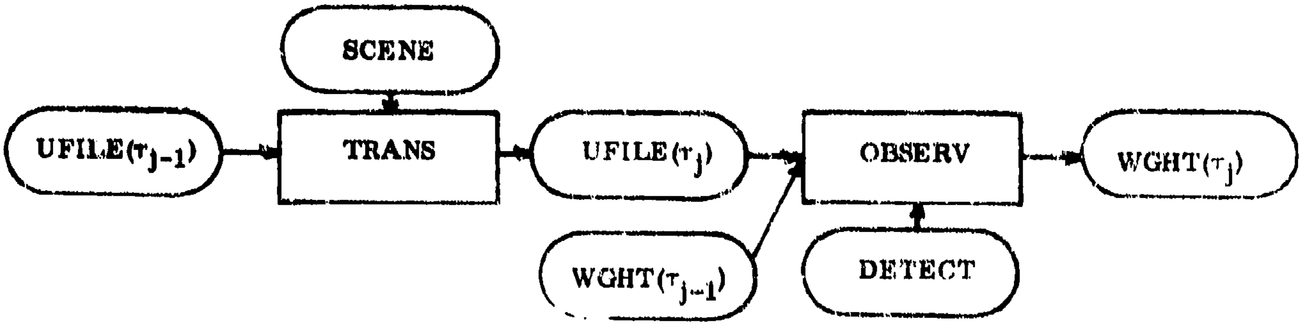
(a) Initial Processing Step



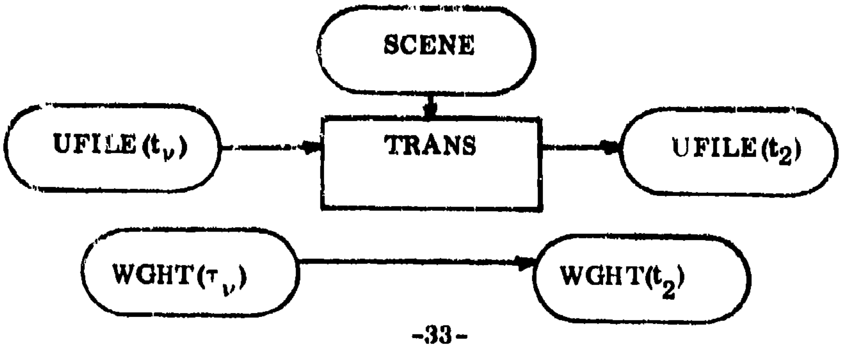
(b) Update from time t_1 to time τ_1



(c) Update from time τ_{j-1} to time τ_j for $j = 2, \ldots, \eta$



(d) Update to time $t_2 > \tau_\eta$



Large weights correspond to a high degree of consistency. The method of computing the weights is discussed below.

Program START. The computer program START creates UFILE(0) and WGHT(0). To do this, it uses the data in the scenario input file SCENE and the data in the detection characteristics input file DETECT.

The records of UFILE(0) are created one after the other. To create a given record (say, the n^{th} record), a sample scenario index $\hat{\iota}^n$ is drawn in accordance with the prescribed credences.

For the n^{th} record, the values $\hat{z}_1^n(0)$ and $\hat{z}_2^n(0)$ of the random variables $\tilde{z}_1(0)$ and $\tilde{z}_2(0)$ are found by sampling from the probability distribution $\Delta_{\hat{\iota}^n}(0)$.

The speed components $\hat{s}_1^n(0)$ and $\hat{s}_2^n(0)$ corresponding to the random variables $\tilde{s}_1(0)$ and $\tilde{s}_2(0)$ are found by sampling from $\Delta_{\hat{\iota}^n}(1)$ to obtain $\hat{z}_1^n(\delta)$ and $\hat{z}_2^n(\delta)$ and then computing

$$\hat{s}_1^n(0) = \frac{1}{\delta} \left(\hat{z}_1^n(\delta) - \hat{z}_1^n(0)\right)$$

and

$$\hat{s}_2^n(0) = \frac{1}{\delta} \left(\hat{z}_2^n(\delta) - \hat{z}_2^n(0)\right).$$

The n^{th} record of UFILE(0) is completed by determining the sample value $\hat{P}_D^n$ for the random variable $\tilde{P}_D$. This is done by taking the minimum value A and the maximum value B for $\tilde{P}_D$ from the input file DETECT and computing

$$\hat{P}_D^n = A + \hat{\zeta}(B-A)$$

where, here and in what follows, $\hat{\zeta}$ denotes an independent draw from a uniform distribution on the interval [0, 1].

The file WGHT(0) is generated by program START so that all weights are equal to unity, i.e., if $w^n(0)$ denotes the content of the n^{th} record of WGHT(0), then

$$w^n(0) = 1 \quad \text{for } 1 \leq n \leq N_r. \tag{II-1}$$

Equation (II-1) reflects the fact that all records of UFILE(0) are considered to be equally likely a priori. The weights will change, however, whenever information is obtained by observing the sensors.

<u>Program TRANS</u>. The computer program TRANS updates file UFILE to reflect target motion in accordance with the scenario information provided in the scenario input file SCENE. No change in file WGHT is made by TRANS since no new sensor information is input to the system during this operation.

In order for TRANS to update UFILE, it is assumed that the tracks between leg endpoints are straight when expressed in coordinates of latitude and longitude. That is, if t_1 and t_2 are times corresponding to target positions on the same leg, then for $l = 1$ and 2 and for $t_1 \leq t \leq t_2$,

$$\tilde{z}_l(t) = \frac{t_2 - t}{t_2 - t_1} \tilde{z}_l(t_1) + \frac{t - t_1}{t_2 - t_1} \tilde{z}_l(t_2).$$

Suppose that TRANS is to update UFILE from simulation time t_1 to simulation time t_2. For the n^{th} record of UFILE(t), let μ and ν be chosen so that $\mu\delta \leq t_1 \leq (\mu+1)\delta$ and $\nu\delta \leq t_2 \leq (\nu+1)\delta$.

If $\nu = \mu$, then the target has not moved to another leg; consequently, for $l = 1$ and 2

$$\hat{z}_l^n(t_2) = \hat{z}_l^n(t_1) + (t_2 - t_1)\hat{s}_l^n(t_1)$$

and

$$\hat{s}_l^n(t_2) = \hat{s}_l^n(t_1).$$

If $\nu = \mu+1$, then the target has moved to the next leg and one must sample from the probability distribution $\Delta_n(\nu+1)$ in order to obtain the target's position $(\hat{z}_1^n(\nu\delta+\delta), \hat{z}_2^n(\nu\delta+\delta))$ at time $(\nu+1)\delta$. Since, for $\nu = \mu+1$ and $l = 1$ and 2,

$$\hat{z}_l^n(\nu\delta) = \hat{z}_l^n(t_1) + (\nu\delta - t_1)\hat{s}_l^n(t_1),$$

the target's position at time $\nu\delta$ as well as $(\nu+1)\delta$ is known, and, therefore, the velocity components on the leg between times $\nu\delta$ and $(\nu+1)\delta$ can be computed. Thus, for $l = 1$ and 2,

$$\hat{s}_l^n(t_2) = \frac{\hat{z}_l^n(\nu\delta + \delta) - \hat{z}_l^n(\nu\delta)}{\delta} \tag{II-2}$$

and target position at time t_2 is given by

$$\hat{z}_l^n(t_2) = \hat{z}_l^n(\nu\delta) + (t_2 - \nu\delta)\, \hat{s}_l^n(t_2). \tag{II-3}$$

Finally, if $\nu \geq \mu+2$, then the target's position at time t_2 is statistically independent of the target's position at time t_1. The velocity and position of the target at time t_2 are found by sampling the probability distributions $\Delta\hat{\Upsilon}_n(\nu\delta)$ and $\Delta\hat{\Upsilon}_n(\nu\delta+\delta)$ to determine $\hat{z}^n(\nu\delta)$ and $\hat{z}^n(\nu\delta+\delta)$ for $l = 1$ and 2 and then by using equations (II-2) and (II-3).

Program OBSERV. The computer program OBSERV updates file WGHT to reflect information gained from the sensor field. Suppose that WGHT(t_1) is to be updated to time t_2. Let $\tau_1, \ldots, \tau_\eta$ denote the times at which the sensor field is observed between times t_1 and t_2 and assume that $t_1 < \tau_1 < \tau_2 < \cdots < \tau_\eta \leq t_2$.

Let $t' = t_1$ and $\tau' = \tau_1$. The first step is to update UFILE(t') to time τ'. This is done using the program TRANS described above. Then the updated file UFILE(τ'), file WGHT(t'), and file DETECT are input to program OBSERV. The weight $\hat{w}^n(\tau')$ corresponding to the n^{th} record of WGHT(τ') is determined from $\hat{w}^n(t')$ corresponding to the n^{th} record of WGHT(t') by multiplication by the conditional probability of observing the actual field responses. That is, $\hat{w}^n(\tau')$ is computed by the formula (essentially Bayes's formula without normalization)

$$\hat{w}^n(\tau') = \hat{w}^n(t')\,[1 - \hat{P}_A^n]^{\iota_1}\,[(1 - \hat{P}_A^n)(1 - \hat{P}_D^n)]^{\iota_2}\,[\hat{P}_A^n]^{\iota_3}\,[1 - (1 - \hat{P}_A^n)(1 - \hat{P}_D^n)]^{\iota_4} \tag{II-4}$$

where the detection probability $\hat{P}_D^n$ is taken from the n^{th} record of UFILE(τ') and the false-response probability $\hat{P}_A^n = f(\hat{P}_D^n)$ is determined from $\hat{P}_D^n$ by use of the "ROC" function f. In our examples, f is defined for simplicity by $f(p) = p^\alpha$; the parameter α describing f is obtained from the input file DETECT.

The exponents $\iota_1, \ldots, \iota_4$ depend upon the position of the target $(\hat{z}_1^n(\tau'), \hat{z}_2^n(\tau'))$ given in the n^{th} record of UFILE(τ'). The exponents are defined as follows:

ι_1 is the number of non-responding sensors which are beyond detection range of the target,

ι_2 is the number of non-responding sensors which are within detection range of the target,

ι_3 is the number of responding sensors which are beyond detection range of the target, and

ι_4 is the number of responding sensors which are within detection range of the target.

Once files UFILE($\tau_{\nu-1}$) and WGHT($\tau_{\nu-1}$) are completed for any $2 \leq \nu \leq \eta$, files UFILE(τ_ν) and WGHT(τ_ν) are obtained by repeating the procedure described above with $t' = \tau_{\nu-1}$ and $\tau' = \tau_\nu$.

When files UFILE(τ_η) and WGHT(τ_η) are obtained, the computation is complete if $t_2 = \tau_\eta$. If $t_2 > \tau_\eta$, then the final update consists of using TRANS to operate on file UFILE(τ_η) in order to generate UFILE(t_2). Since no new observations occur between times τ_η and t_2, WGHT(t) is a replica of WGHT(τ_η).

Program MAP and other output. Let the random vector $\tilde{U}(t)$ be defined by

$$\tilde{U}(t) = (\tilde{z}_1(t),\ \tilde{z}_2(t),\ \tilde{s}_1(t),\ \tilde{s}_2(t),\ \tilde{k},\ \tilde{P}_D).$$

Each record $\hat{U}^n(t) = (\hat{z}_1^n(t),\ \hat{z}_2^n(t),\ \hat{s}_1^n(t),\ \hat{s}_2^n(t),\ \hat{k}^n,\ \hat{P}_D^n)$ of UFILE(t) is then an independent sample of $\tilde{U}(t)$.

Any probability statement associated with the random variable $\tilde{U}(t)$, conditioned upon observation of the sensors, may be estimated using files UFILE(t) and WGHT(t) and the formula (B is a set representing an event)

$$\Pr\{\tilde{U}(t) \in B \mid \text{sensor observations}\} \approx \frac{\sum_{n \in I(B)} \hat{w}^n(t)}{\sum_{n=1}^{N_r} \hat{w}^n(t)} \qquad \text{(II-5)}$$

where $I(B) = \{n \mid \hat{U}^n(t) \in B\}$. For example, the computer program MAP operates on files UFILE(t) and WGHT(t) to produce a probability map for the target's location.

To do this, MAP accepts inputs which define a grid over the geographical region of interest. This grid might consist, for example, of cells each covering one degree of latitude and one degree of longitude.

Let the grid cells be denoted G_j for $1 \leq j \leq J$ and let the corresponding target location probabilities determined by MAP from files UFILE(t) and WGHT(t) be denoted $\hat{L}_j(t)$ for $1 \leq j \leq J$. Then the values of $\hat{L}_j$ are computed by MAP for B defined by

$$B = \{(b_1, \ldots, b_6) \mid (b_1, b_2) \in G_j\}.$$

The updated credence for the i^{th} scenario is given by equation (II-5) for B defined by

$$B = \{(b_1, \ldots, b_6) \mid b_5 = i\}.$$

Finally, the mean of the updated probability distribution for detection probability is computed by

$$\text{Exp}[\tilde{P}_D \mid \text{sensor observations}] \approx \frac{\sum_{n=1}^{N_r} \hat{P}_D^n \hat{w}^n(t)}{\sum_{n=1}^{N_r} w^n(t)}. \tag{II-6}$$

CHAPTER III

THE APPLICATION OF INFORMATION THEORY TO OPTIMAL SURVEILLANCE IN A FALSE TARGET ENVIRONMENT--AN EXPLORATORY ANALYSIS

This chapter examines in an exploratory way the application of information theory to optimal allocation of surveillance resources in a false target environment. The objective here is to investigate methods of allocating ASW resources for the purpose of shaping the target location probability distribution to serve certain tactically useful purposes. The preceding chapter provides methods for updating the target location probability distribution to incorporate the sensor information resulting from these allocations.

We have been motivated by a strong heuristic attraction to policies which build up the information content of the target location probability distribution. We are aware, however, that much of this attraction is due to semantics (i.e., the fact that the language of information theory is so suggestive in the present context) and we have tried to exhibit more substantive reasons why further development of "maximum information-gain policies" may be desirable from an operational point of view.

The term "optimal" is used in the title of this chapter to reflect a desire rather than to state an accomplishment. We desire to find the best surveillance policy within the context of our tactical scenario, and, to this end, we formulate several policies and examine their properties. The policy based on maximizing the information content of the target location probability distribution appears to be close to optimal (among those plans considered) in all cases examined. Further work is required, however, before more precise statements can be made.

The investigation is approached numerically and theoretically. The numerical work is based upon monte-carlo simulation of the properties of selected surveillance policies. These results are presented in this chapter. The theoretical work has been directed towards establishing the connection between the application of information theory to surveillance and the application of information theory to statistical hypothesis testing. The latter applications have an extensive literature. The results of this theoretical work and review are rather technical and are presented in Appendix B.

From both perspectives, numerical and theoretical, we feel that the definition of information used in information theory has the promise of providing a useful measure of effectiveness for judging the utility of alternative allocations of diverse ASW resources in certain kinds of surveillance missions.

The first section presents the tactical setting and basic assumptions underlying the numerical analysis. The surveillance policies considered are described in the second section, and the numerical results are given in the third section. The fourth section presents conclusions and provides a brief review of related analyses which have appeared in the operations research literature. The related statistical literature is discussed in Appendix B.

Tactical Setting and Basic Assumptions

We shall assume that there are J search cells, one of which contains the target. The target may move from cell to cell in the course of the search, and target motion is modeled as a Markov process as described below. The surveillance procedure consists of assigning ASW search effort to a selected grid cell and then estimating the target's location (designating the cell containing the target) based upon the search results. This is similar to the "whereabouts" searches discussed in reference [e].

The surveillance operation is carried out sequentially in stages where each stage consists of assigning search effort to a single cell, evaluating the search results, and then estimating the target's location. Changes in target position only take place between stages.

As an example of a potential application, consider a VP operation where each day one or more flights are sent to an area specified for that day. At the end of the day, the search results are evaluated, the area for the next day's flight is determined, and the best estimate of the target's location (specified by a grid cell) is passed to the operational commander.

Sensor-response assumptions. If a sensor response is obtained in a cell searched, this does not necessarily mean that the target is located in that cell. Because of the possibility of false responses, one never knows with certainty that the target has been detected.

It is assumed that the performance of the ASW search system is idealized in terms of a J × J response array,

$$R = \left(R(i, j)\right),$$

where R(i, j) is the probability that an increment of search effort applied to the j^{th} cell will result in a response, given that the target is located in the i^{th} cell. Here, as usual, i is the row index and j is the column index. As in Chapter II, a response is a decision, based upon the available information, in favor of the hypothesis that the target is present as opposed to the alternative hypothesis that the target is not present.

At the end of each stage, a cell is selected to contain the target. For all policies examined in this report, the cell selected is the one having the highest target location probability based upon evaluation of the search results.

The problem is to determine a surveillance policy (a procedure for assigning search effort and estimating target location) which will maximize effectiveness over an extended period of time.

Measure of effectiveness. In order to measure surveillance effectiveness, let S(k) denote the probability of correctly selecting the cell containing the target at the end of the k^{th} stage. The function S is called the "success function." A surveillance policy which maximizes S(k) is called a "k-optimal" surveillance policy and a surveillance policy which maximizes S(k) for all $k \geq 1$ is referred to as a "uniformly optimal" surveillance policy.

A success occurs if the correct cell is selected, although this fact can never be confirmed, since any sensor response is possibly due to a non-target cause. Confidence in the specified target locations can only be obtained by an accumulation of evidence, no single item of which is decisive.

Target motion assumptions. Target motion is assumed to be a Markov process described by an initial probability row vector d and a transition matrix D. The resulting target motion stochastic process may or may not be a stationary process, depending upon whether or not d is the stationary vector for the process. In more general non-Markovian situations, the methods of Chapter II can be used to model the motion of the target.

For this illustration, it will be assumed that D is a circulant matrix (see, for example, page 51 of reference [s]) having the form

$$D = \begin{pmatrix} 1-\frac{(J-1)}{J}\delta, & \frac{\delta}{J}, & \ldots, & \frac{\delta}{J} \\ \frac{\delta}{J}, & 1-\frac{(J-1)}{J}\delta, & \frac{\delta}{J}, & \ldots, \frac{\delta}{J} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \frac{\delta}{J}, & \ldots, & \frac{\delta}{J}, & 1-\frac{(J-1)}{J}\delta \end{pmatrix}, \qquad \text{(III-1)}$$

where $0 \leq \delta \leq 1$. The process with transition matrix D given by equation (III-1) depends upon a single parameter δ and is stationary if and only if d is the uniform distribution $d(j) = 1/J$ for $j = 1, \ldots, J$.

If d^k denotes the target's distribution over the J cells after the k^{th} transition, then it is not difficult to show that for $j = 1, \ldots, J$, and $k = 1, 2, \ldots$,

$$d^k(j) = \frac{1}{J} + (1-\delta)^k \left[d(j) - \frac{1}{J} \right].$$

Thus, each component in the target distribution vector converges monotonically to the uniform vector. We shall call δ the dispersion constant. Note that if $\delta = 0$, then the target is motionless, and if $\delta = 1$, then the target distribution disperses to the uniform distribution in one step.

The object of the tracking policy is to overcome the dispersive effects of random target motion by the expenditure of search effort.

Formulation in terms of stochastic control. It is useful to look at this surveillance problem as a problem of controlling a Markov process (for background in stochastic control see, for example, reference [t]).

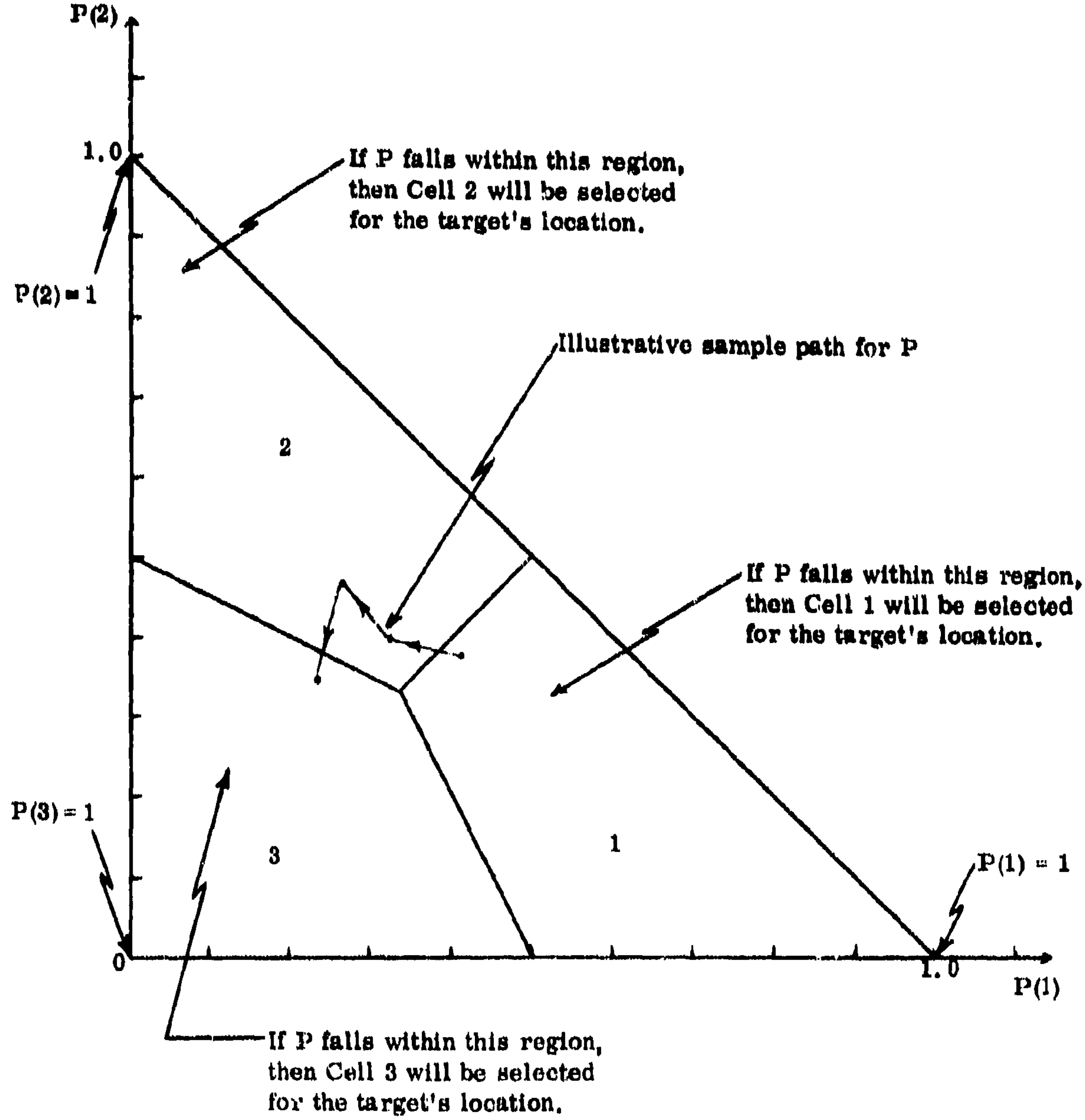
Consider a dynamic system whose state is the probability vector P for the target's location. We are particularly interested in this vector at the beginning of each stage. Except for the first stage where P is assumed known, P depends upon random sensor observations, and, therefore, P is itself a random variable. In fact, the time behavior of P is Markovian when d, D, and R are assumed known without error. For three cells ($J = 3$), it is possible to visualize P as a point (P(1), P(2)) in the plane since $\Sigma_{j=1}^{3} P(j) = 1$. Figure III-1 shows the state space for P based upon this interpretation.

The object of the tracking policy is to provide information which will permit one to correctly select the cell containing the target at the end of each stage. Since the predetermined selection rule is to pick the cell with the highest posterior probability as determined by the search results obtained during the stage, we can consider the state space of P to be the union of three disjoint (except for boundaries) regions labeled 1, 2, and 3 in Figure III-1. If the point falls in region j at the end of a stage, then the j^{th} cell is selected as the cell containing the target.

The "control" is a decision function or policy which depends upon P at the beginning of a stage and which indicates the cell to be searched during that stage. If the target is actually in cell j during a stage, then the target is visualized as occupying the vertex determined by $P(j) = 1$. The purpose of the control is to guide the point P as often as possible into the set which contains the target.

FIGURE III-1

STATE SPACE DIAGRAM FOR TARGET LOCATION PROBABILITY VECTOR



Let P_B denote the vector P at the beginning of a stage, and let P_A denote P at the end of a stage. The target is assumed to move only between stages, and the P_B for a given stage is computed by $P_B = P_A D$ where P_A is the vector P at the end of the previous stage and D is the transition matrix.

For the response matrix R and a decision to search cell m, the vector P_A will depend upon whether or not a response is obtained. If a response is not obtained, then

$$P_A(j) = \frac{P_B(j)\ (1 - R(j, m))}{\Sigma_{i=1}^{J}\ P_B(i, m)\ (1 - R(i, m))}, \quad \text{for } 1 \leq j \leq J.$$

If a response is obtained, then

$$P_A(j) = \frac{P_B(j)\ R(j, m)}{\Sigma_{i=1}^{J}\ P_B(i, m)\ (1 - R(i, m))}, \quad \text{for } 1 \leq j \leq J.$$

It is theoretically possible to design a control which will maximize probability of success S(k) at the end of the k^{th} stage. This would be the generalization of the single-stage look-ahead policy described in the next section. In fact, by working backwards in time, this k-optimal control can be found by dynamic programming, although the solution is rather complicated.

Our main interest, however, is in the situation where all stages are important and where it is not natural to establish a fixed terminal time. In order to gain insight into this situation, we will examine the behavior of four decision policies (i.e., controls); these are described in the next section. Two of the policies, the single-stage optimal look-ahead policy (Policy I) and a control based upon maximizing the information content of the posterior distribution (Policy II), are chosen for their intuitive appeal. The other two policies, a policy based on searching the highest probability cell (Policy III) and a policy based upon searching the cells in a regular rotation, are chosen because they are simple and easy to compute and they give us "bench marks" for comparison.

It should be noted that the highest probability cell policy has been mentioned as optimal in a closely related scenario examined in reference [u]. In reference [w], however, the search stops as soon as the first response is obtained and the search is successful if and only if the response occurs in the cell containing the target. Our measure of effectiveness is quite different.

Description of the Surveillance Policies

This section describes the four different surveillance policies considered. These are the optimal single-stage look-ahead policy (Policy I), the maximum information-gain policy (Policy II), the highest probability cell policy (Policy III), and the uniform surveillance policy (Policy IV). They are described individually in the subsections below.

The success function defined in the preceding section is taken to be the measure of effectiveness and is computed by monte-carlo simulation. For the k^{th} search stage of the n^{th} monte-carlo replication, let

$$\hat{S}(n,k) = \begin{cases} 1 & \text{if the target cell is correctly specified} \\ 0 & \text{otherwise.} \end{cases}$$

For N replications, the k^{th} stage success probability S(k) is estimated by the formula

$$S(k) \approx \frac{1}{N} \sum_{n=1}^{N} \hat{S}(n,k).$$

For each comparison, an initial target location probability distribution d and a transition matrix D are specified to describe the target's movements and a response matrix R is specified to describe the search environment and the sensor system.

The monte-carlo calculation begins by drawing a random number to pick the cell for the target's initial location. This selection is made in accordance with the initial target location probability distribution d.

The search policy specifies a search cell for each stage based upon the current before-search target location probability distribution P_B; then the search results are simulated in accordance with the target's actual location and the probabilities given by the response array R. Next, the after-search target location probability distribution P_A is determined from the simulated search results, and, finally, the after-search highest probability cell (based upon P_A) is selected as the target cell. The target position is then updated in accordance with the target motion transition matrix D at the end of the stage and a new estimate of the before-search probability distribution P_B is obtained by computing $P_B = P_A D$. The process is then repeated.

<u>The optimal single-stage look-ahead policy (Policy I).</u> The optimal single-stage look-ahead policy is to search in the cell which, based upon the estimated vector P_B, maximizes the probability of correctly selecting the target's cell at the end of the stage. This is a generalization of the optimal whereabouts plan formulated in reference [e] for searches without false responses.

More precisely, at the beginning of any stage, let $P_B(i)$ denote the before-search probability that the target is located in cell i and let R be the response matrix. Let $p_A(r, i, j)$ be the conditional after-search probability that the target is located in the i^{th} cell given that the j^{th} cell was searched and result r was obtained. Here, r = 1 indicates a target-like response and r = 0 indicates no target-like response. Let $Q(r, i, j)$ denote the probability of obtaining search result r given that the target is in cell i and that cell j is searched. Then

$$Q(r, i, j) = \begin{cases} R(i, j) & \text{for } r = 1 \\ 1 - R(i, j) & \text{for } r = 0. \end{cases}$$

The probability function p_A is determined from P_B and Q by the equation

$$p_A(r, i, j) = \frac{P_B(i)\, Q(r, i, j)}{\sum_{m=1}^{J} P_B(m)\, Q(r, m, j)}.$$

Let $X(r, j)$ denote the cell selected to contain the target given that cell j was searched and result r was obtained. In view of the selection rule, which states that the cell with the highest target location probability should be chosen, we have

$$p_A(r, X(r, j), j) \geq p_A(r, i, j) \qquad \text{for } 1 \leq i \leq J.$$

Let B(k) denote the before-search probability that if the k^{th} cell is searched, then the target cell will be correctly selected based upon the search results. If

$$\epsilon(x) = \begin{cases} 1 & \text{if } x = 1 \\ 0 & \text{otherwise,} \end{cases}$$

$$B(j) = \sum_{i=1}^{J} \sum_{r=0}^{1} P_B(i)\, R(r, i, j)\, \epsilon(i - X(r, j))$$

$$= \sum_{r=0}^{1} P_B(X(r, j))\, Q(r, X(r, j), j)$$

$$= \max\{P_B(i)\, R(i, j) : 1 \leq i \leq J\} + \max\{P_B(i)\, [1 - R(i, j)] : 1 \leq i \leq J\}.$$

The optimal single-stage look-ahead policy is to search in cell j^* for which

$$B(j^*) \geq B(j) \quad \text{for } 1 \leq j \leq J.$$

If more than one cell qualifies, then from the qualifying cells the cell with the highest probability according to P_B is chosen. If more than one cell still qualifies, then the search cell is selected randomly from the qualifying cells according to a uniform distribution.

<u>The maximum information-gain policy (Policy II)</u>. The maximum information-gain policy is to search in the cell which maximizes the expected information content (or equivalently minimizes the expected entropy) of the posterior after-search target location probability distribution.

More precisely, let $P_B(j)$ and $p_A(r, i, j)$ be defined as above, and let the entropy (see reference [d]) H(P) of any probability vector P over J cells be defined by

$$H(P) = - \sum_{j=1}^{J} P(j) \ln P(j).$$

Intuitively speaking, as the entropy of a distribution increases, the distribution flattens. It is well known (see, for example, reference [d]) that maximum entropy* is attained by the uniform distribution. The information content of a distribution P is defined to be -H(P) + C where C is some fixed constant. We are interested only in changes in information and, hence, the value of C is not important.

The expected entropy U(j) of the posterior target location distribution given search in cell j is given by

---

* This result holds for probability distributions on a finite number of points but not for probability distributions on a countably infinite number of points. In this latter case, a uniform probability distribution is not defined.

$$U(j) = \sum_{i=1}^{J} \sum_{r=0}^{1} P_B(i)\, Q(r, i, j)\, H(p_A(r, \cdot, j))$$

$$= - \sum_{i=1}^{J} \sum_{r=0}^{1} P_B(i)\, Q(r, i, j) \ln[p_A(r, i, j)].$$

The maximum information-gain policy is to search in any cell j^* for which

$$U(j^*) \le U(j) \quad \text{for } 1 \le j \le J.$$

This corresponds to a Lindley procedure (see reference [f]) for sequential experimental design as discussed in Appendix B. If more than one cell satisfies the above inequality, then the search cell is selected randomly from the qualifying cells according to a uniform distribution.

<u>The highest probability cell policy (Policy III)</u>. The highest probability cell policy is to search in the cell with the highest before-search probability. That is, if at the beginning of a stage $P_B(j)$ is the before-search probability that the target is located in cell j, then the highest probability cell policy is to search in any cell j^* for which

$$P_B(j^*) \ge P_B(j) \quad \text{for } 1 \le j \le J.$$

If more than one cell qualifies, then the search cell is selected randomly from the qualifying cells according to a uniform distribution.

It is interesting to note that the before-search target location probability distribution is identical to the expectation (with respect to P_B) of the after-search target location probability distribution regardless of the cell searched. In order to show this, let $\xi(i, j)$ denote the expected after-search probability that the target is located in cell i, given that cell j is searched. Then

$$\xi(i, j) = \sum_{n=1}^{J} \sum_{r=0}^{1} P_B(n)\, Q(r, n, j)\, p_A(r, i, j)$$

$$= \sum_{n=1}^{J} \sum_{r=0}^{1} P_B(n)\, Q(r, n, j) \frac{P_B(i)\, Q(r, i, j)}{\sum_{m=1}^{J} P_B(m)\, Q(r, m, j)}$$

$$= \sum_{r=0}^{1} P_B(i)\, Q(r, i, j)$$

$$= P_B(i).$$

The uniform surveillance policy (Policy IV). The uniform surveillance policy is to search systematically through all search cells in a fixed rotation. In mathematical notation, the j^{th} cell is searched during the k^{th} stage where $j \equiv 1 + (k-1) \pmod J$ for $k = 1, 2, \ldots,$ and J equal to the number of search cells. That is, one searches the J cells in order and then repeats the search as often as required.

If the target does not move, it is not difficult to prove that the success function for this policy will converge to 1 whenever the rows of the response matrix are distinct (the usual case).

Numerical Comparison of Surveillance Policies

This section provides a numerical comparison of the four surveillance policies described in the preceding section. Five surveillance cases are considered corresponding to different assumptions about d, D, and R. In order to reduce complexity and make it easier to interpret the results, the search grid is limited to three cells in the first four cases and nine cells in the fifth case. Moving targets are considered only in the first case.

Three response matrices are examined. The first is

$$R = \begin{pmatrix} .8 & .1 & .1 \\ .7 & .8 & .1 \\ .7 & .1 & .8 \end{pmatrix} \qquad \text{(III-2)}$$

and is used in Cases I, II, and III. Recall that $R(i, j)$ is the probability of obtaining a response from a search of cell j given that the target is in cell i. This particular form of R is chosen in order to simulate a situation where search in one cell (the first) produces very little information gain. In this cell, the true-response and false-response probabilities are nearly equal (.8 and .7, respectively).

Three initial target location probability distributions are used with the response matrix of equation (III-2); these are a uniform distribution (Case I) given by $d(1) = .33$, $d(2) = .33$, and $d(3) = .34$, a "highly" non-uniform distribution (Case II) given by $d(1) = .75$, $d(2) = .15$, and $d(3) = .10$, and a "moderately" non-uniform distribution (Case III) given by $d(1) = .5$, $d(2) = .3$, and $d(3) = .2$.

The second response matrix,

$$R = \begin{pmatrix} .8 & .6 & .3 \\ .6 & .8 & .6 \\ .3 & .6 & .8 \end{pmatrix}, \qquad \text{(III-3)}$$

corresponds physically to a situation where the three cells are arranged in a row and where the response probabilities increase the "closer" one gets to the target cell. The uniform target location probability distribution is used with the response matrix given by equation (III-3) in Case IV.

The third response matrix,

$$\begin{pmatrix} .8 & \cdots & .8 & .1 \\ .8 & \cdots & .8 & .2 \\ & \cdots & & \\ .8 & \cdots & .8 & .9 \end{pmatrix}, \qquad \text{(III-4)}$$

is considered in Case V and has some features in common with the response matrices given by equations (III-2) and (III-3). It is similar to the response matrix given by equation (III-2) in that little information is gained from searching certain cells. Equation (III-4) represents an extreme case in which no information is gained from searching cells 1 through 8. It is similar to the response matrix given by equation (III-3) in that one may think of cells 1 through 9 arranged in a row with the probability of a response from a search of cell 9 increasing with decreasing distance from the target.

The initial target location probability distribution used in Case V is $d(1) = .2$ and $d(j) = .1$ for $2 \leq j \leq J$.

The numerical results are given in the following subsections. In Cases I through IV, 400 monte-carlo replications are used for each curve, and in Case V, 50 replications are used.

Case I(a) -- stationary target. As mentioned above, there are three grid cells and the initial target location probability distribution is uniform. The response matrix is given by equation (III-2). For all cells, if the target is in the cell searched, then the probability of response is .8. The probability of false response is .7 in the first cell and the probability of false response is .1 in the second and third cells. These false-response probabilities do not depend upon the location of the target.

Figure III-2 provides the estimated probability of response curve for Policies I through IV. Notice that the success probability for the maximum information-gain policy (Policy II) approaches 1 asymptotically. In the earlier stages of search, there appears to be no statistically significant difference between the maximum information-gain policy and the optimal single-stage look-ahead policy (Policy I). Asymptotically, however, the maximum information-gain policy appears to have a slight advantage. It is interesting to note that the uniform surveillance policy (Policy IV) also does well in this example. The highest probability cell policy (Policy III) is not particularly attractive in contrast to the other policies.

A problem with Policy I occurs when the target location probability distribution becomes very concentrated. When this happens, the after-search estimate of target location will not depend upon the cell searched nor upon the search results. Since only one stage is considered, all search cells appear equally attractive and the before-search highest probability cell is chosen. This can lead to trouble as we shall see in Case V. Table III-1 displays the details of one of the monte-carlo replications for Policy I in order to illustrate this point. Notice that the state of indeterminacy is reached at the third search stage.

Case I(b) -- moving target. The same initial target location probability distribution and sensor response assumptions are made as in Case I(a). In the present case, however, the target is permitted to move between stages according to a Markov process specified by the dispersive transition matrix given by equation (III-1). The comparison is limited to Policies II and IV.

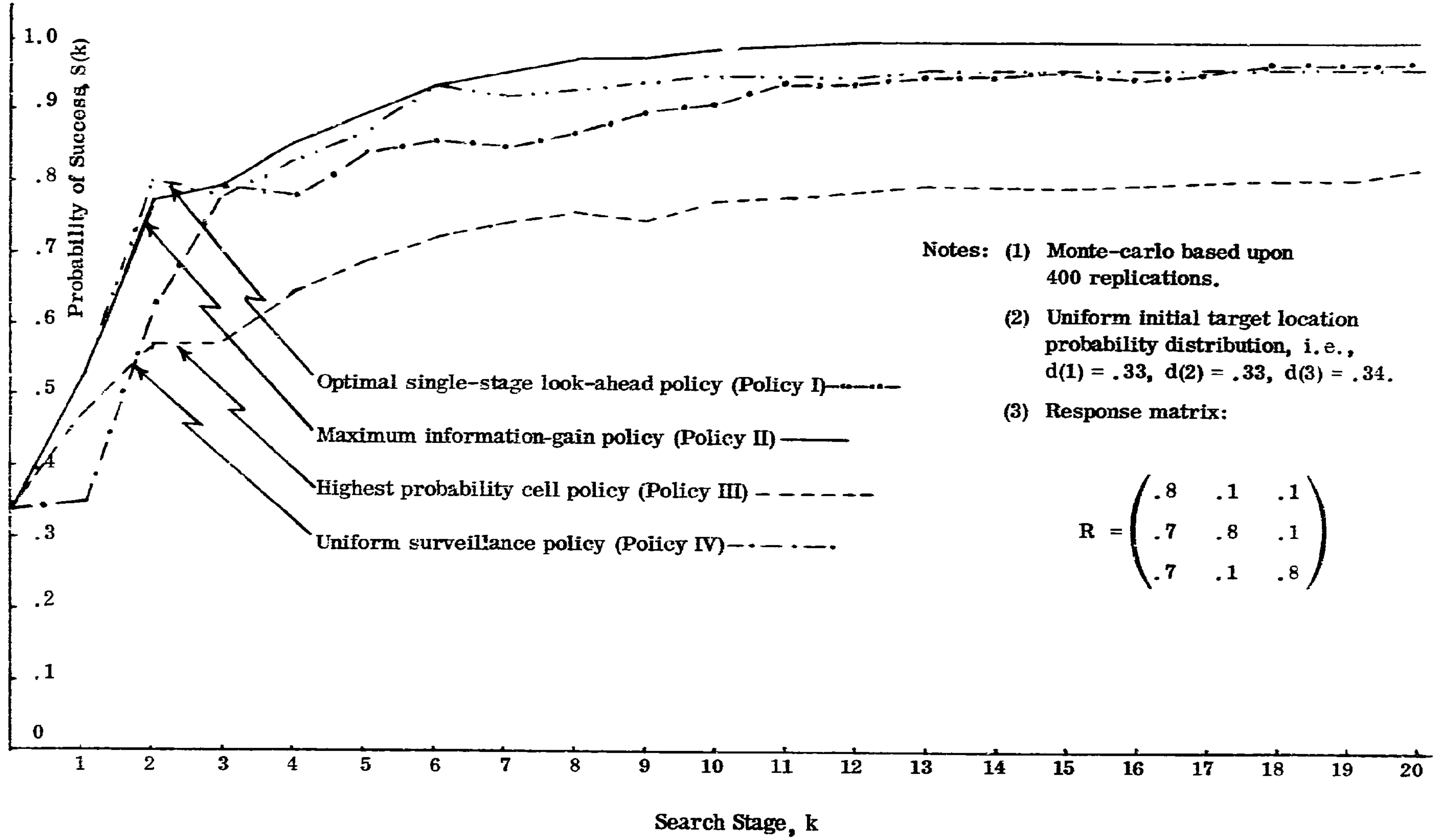
Figure III-3 shows the influence of target motion on probability of success when the maximum information-gain policy (Policy II) is used. Three examples are considered corresponding to values of the dispersion constant, $\delta = 0$, $\delta = .3$, and $\delta = 1$. The example where $\delta = 0$ corresponds to no target motion, and, therefore, this curve is the same as that given in Figure III-2 for Policy II. The example where $\delta = 1$ corresponds to complete dispersion of the target location probability distribution to a uniform distribution at each stage. Here, even if the target's position is known with certainty at the end of a stage, the ensuing motion will produce a uniform distribution for target location at the beginning of the next stage.

The first transition of the Markov process is made at the end of the first stage. Therefore, S(1) is identical for all three values of the dispersion constant.

Figure III-4 shows the influence of target motion on probability of success when the uniform surveillance policy is used. Results are shown for the same values of the dispersion constant as used in Figure III-3. The striking irregularity of the curves given in Figure III-4 is due to the fact that the uniform surveillance policy considered here is a regular rotation of search through the three cells in the grid. As noted before, Cell 1 is a particularly poor cell to search because of

FIGURE III-2

COMPARISON OF SURVEILLANCE POLICIES (CASE I(a))



Notes: (1) Monte-carlo based upon 400 replications.

(2) Uniform initial target location probability distribution, i.e., d(1) = .33, d(2) = .33, d(3) = .34.

(3) Response matrix:

$$R = \begin{pmatrix} .8 & .1 & .1 \\ .7 & .8 & .1 \\ .7 & .1 & .8 \end{pmatrix}$$

TABLE III-1

ILLUSTRATIVE SURVEILLANCE OPERATION USING THE OPTIMAL SINGLE-STAGE LOOK-AHEAD POLICY

Notes: (1) In this illustration, the target is in cell 3.

(2) Response matrix:

$$R = \begin{pmatrix} .8 & .1 & .1 \\ .7 & .8 & .1 \\ .7 & .1 & .8 \end{pmatrix}$$

Stage	Before-Search Target Location Probability P_B in Cell j			Estimated Probability of Success B(j) Given Search in Cell j			Cell Selected for Search	Search Results	After-Search Target Location Probability in Cell			Estimated Target Cell
	j = 1	2	3	k = 1	2	3			1	2	3	
1	.330	.330	.340	.366	.570	.569	2	No Response	.444	.099	.457	3
2	.444	.099	.457	.492	.491	.765	3	Response	.106	.023	.871	3
3	.106	.023	.871	.871	.871	.871	3	Response	.015	.003	.981	3

FIGURE III-3

THE INFLUENCE OF TARGET MOTION ON PROBABILITY OF SUCCESS FOR THE MAXIMUM INFORMATION GAIN POLICY

Case I(b)

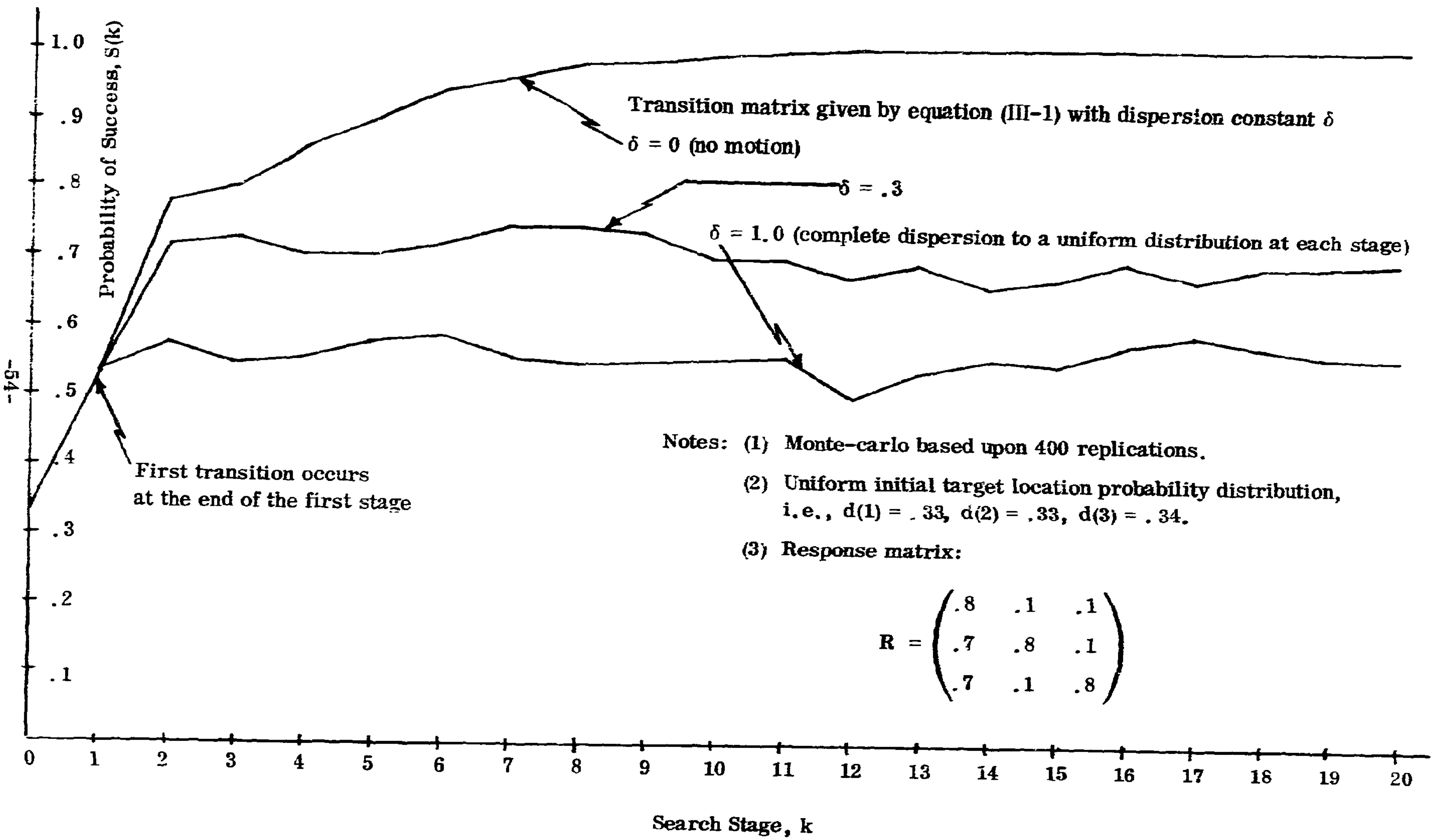
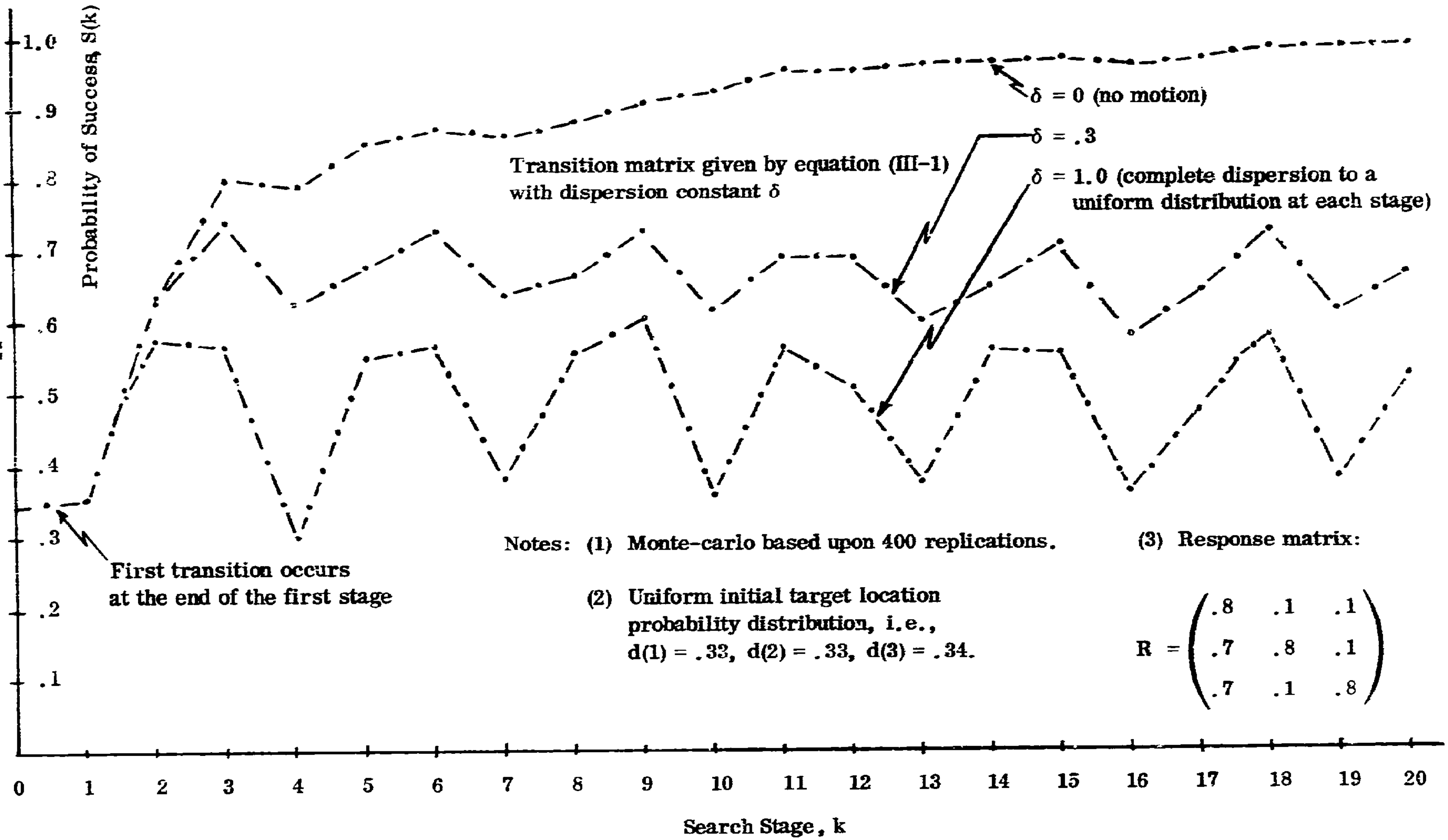


FIGURE III-4

THE INFLUENCE OF TARGET MOTION ON PROBABILITY OF SUCCESS FOR THE UNIFORM SURVEILLANCE POLICY

Case I(b)



the high probability of false response in this cell. The dips in the curves correspond to search of Cell 1 at stages 3k+1 for k = 0, 1, 2, ... These dips become increasingly pronounced as the dispersion constant increases.

As in Figure III-3 the curves coincide at the first stage since there has been no target motion up to this point.

It should be noted that if the uniform surveillance policy were implemented by selecting the search cells at random according to a uniform distribution, then the curves would be smoother and would not exhibit the periodic dips. This would not, however, improve the average performance of the policy.

Case II. In this case the initial target location probability distribution is non-uniform with the highest probability assigned to the first cell, i.e., d(1) = .75, d(2) = .15, and d(3) = .10. The response matrix is the same as in Case I.

This case is presented to illustrate a situation where searching the highest probability cell is clearly not a good policy. Here, the first cell has a very high initial target location probability but very little is learned from a search of this cell because of the high false-response probability.

Figure III-5 provides the estimated probability of success curves for Policies I through IV. As anticipated, the highest probability cell policy does not appear to be very good. In fact, it is only slightly better than the trivial policy which would select the target cell at random in accordance with the initial target location probability distribution d and reselect the same cell at each stage. In this case the trivial policy would select the first cell with probability .75, the second with probability .15, and the third with probability .1.

Once again, Policy II does very well, and as one might expect, Policy I initially does better than Policy II with the latter catching up in the latter stages. Once again, we note that S for Policy IV will always converge to one when the target is stationary and the rows of the response matrix are distinct.

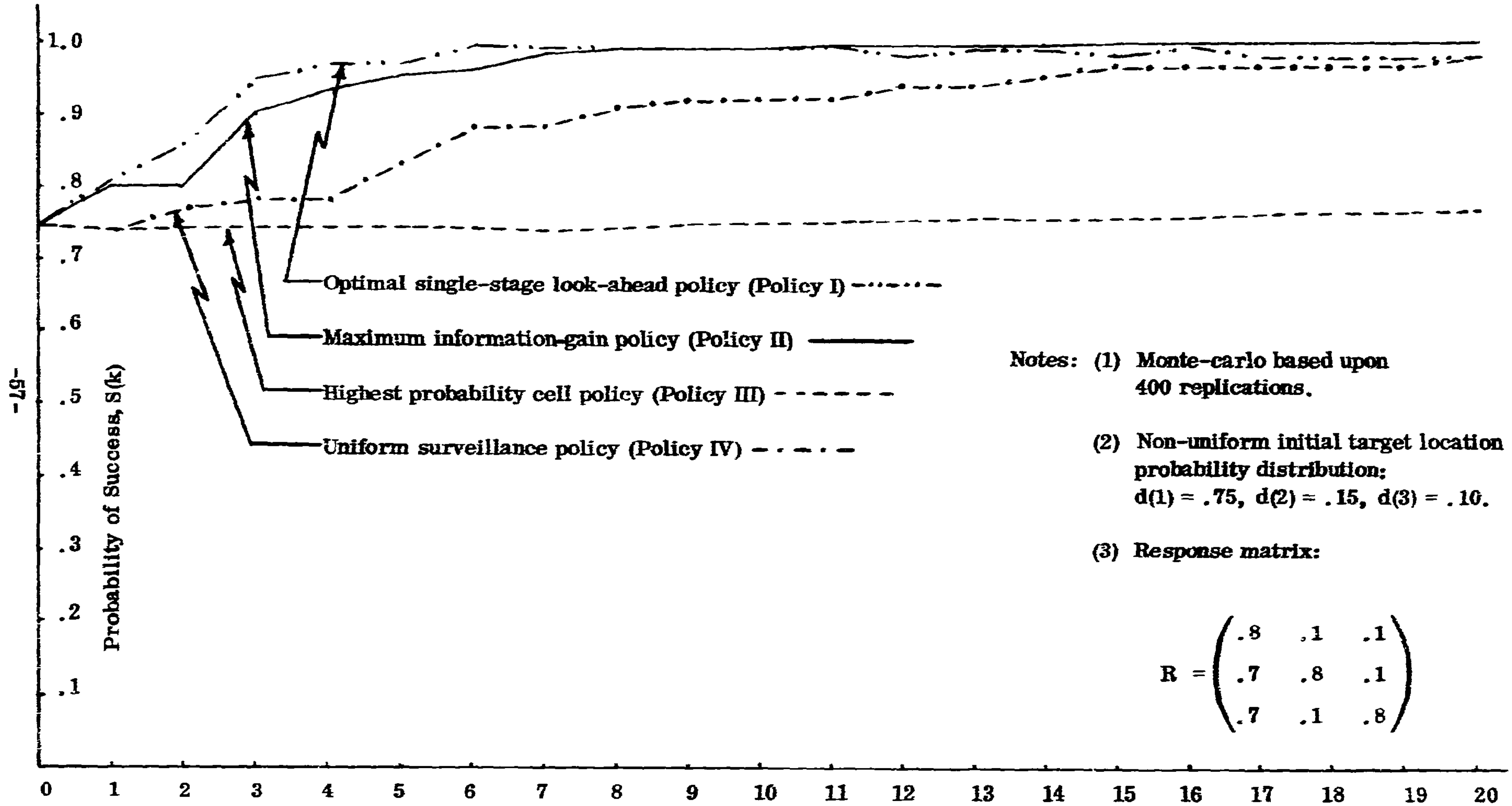
Case III. As in Case II, it is assumed that the initial target location probability distribution is non-uniform and given by d(1) = .5, d(2) = .3, and d(3) = .2. Figure III-6 provides the estimated probability of success curves for Policies I through IV.

Once again, we see that Policy II, the maximum information-gain policy, appears to be better than the others.

Case IV. In this case, the initial target location probability distribution is uniform, i.e., d(1) = .33, d(2) = .33, and d(3) = .34, and the response matrix is given by equation (III-3). Here, Cells 2 and 3 have relatively high false-response probabilities in contrast to the previous cases.

FIGURE III-5

COMPARISON OF SURVEILLANCE POLICIES (CASE II)



Notes: (1) Monte-carlo based upon 400 replications.

(2) Non-uniform initial target location probability distribution: d(1) = .75, d(2) = .15, d(3) = .10.

(3) Response matrix:

$$R = \begin{pmatrix} .8 & .1 & .1 \\ .7 & .8 & .1 \\ .7 & .1 & .8 \end{pmatrix}$$

FIGURE III-6

COMPARISON OF SURVEILLANCE POLICIES (CASE III)

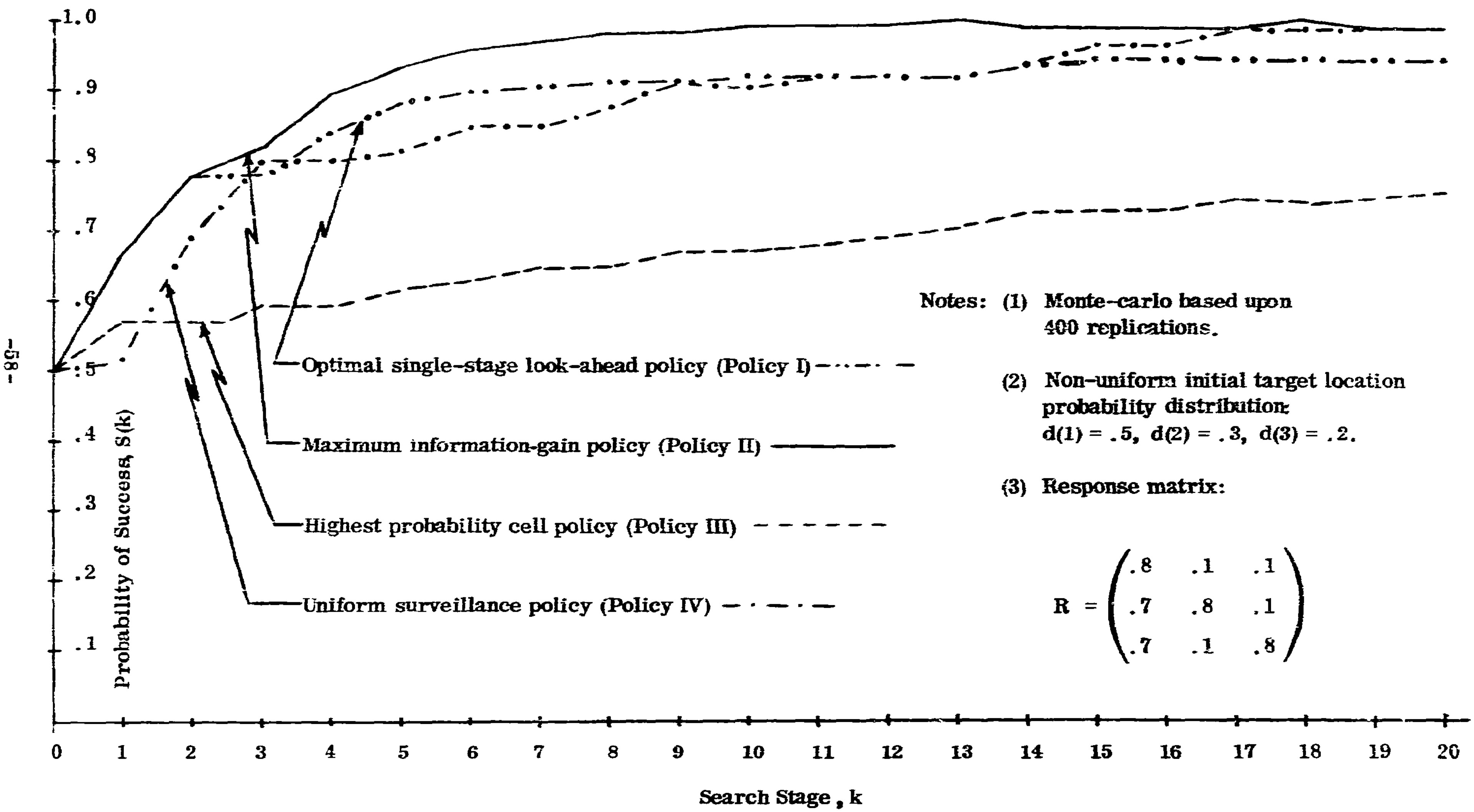


Figure III-7 provides the probability of success curves for Policies I through IV. In contrast to the other cases considered thus far, there appears to be little difference in the probability of success curves.

Case V. The purpose of this case is to examine a situation where the maximum information-gain policy (Policy II) is definitely superior to the other plans considered. It is assumed that there are 9 search cells and that the initial target location probability distribution d is given by d(1) = .2 and d(j) = .1 for j = 2, . . . , 9. The response matrix is given by equation (III-4). The results are shown in Figure III-8 for Policies I, II, and IV. Policy III is extremely poor in this case and is not shown. It continually picks the first cell for search and its probability of success function remains constant with a value of .2.

According to this response matrix, no information is gained by searching in Cells 1 through 8. Since the uniform surveillance policy (Policy IV) rotates search through all cells, a considerable amount of time will be lost when this plan is used. Although the probability of success function for Policy IV is guaranteed to converge to 1, the convergence will be slow.

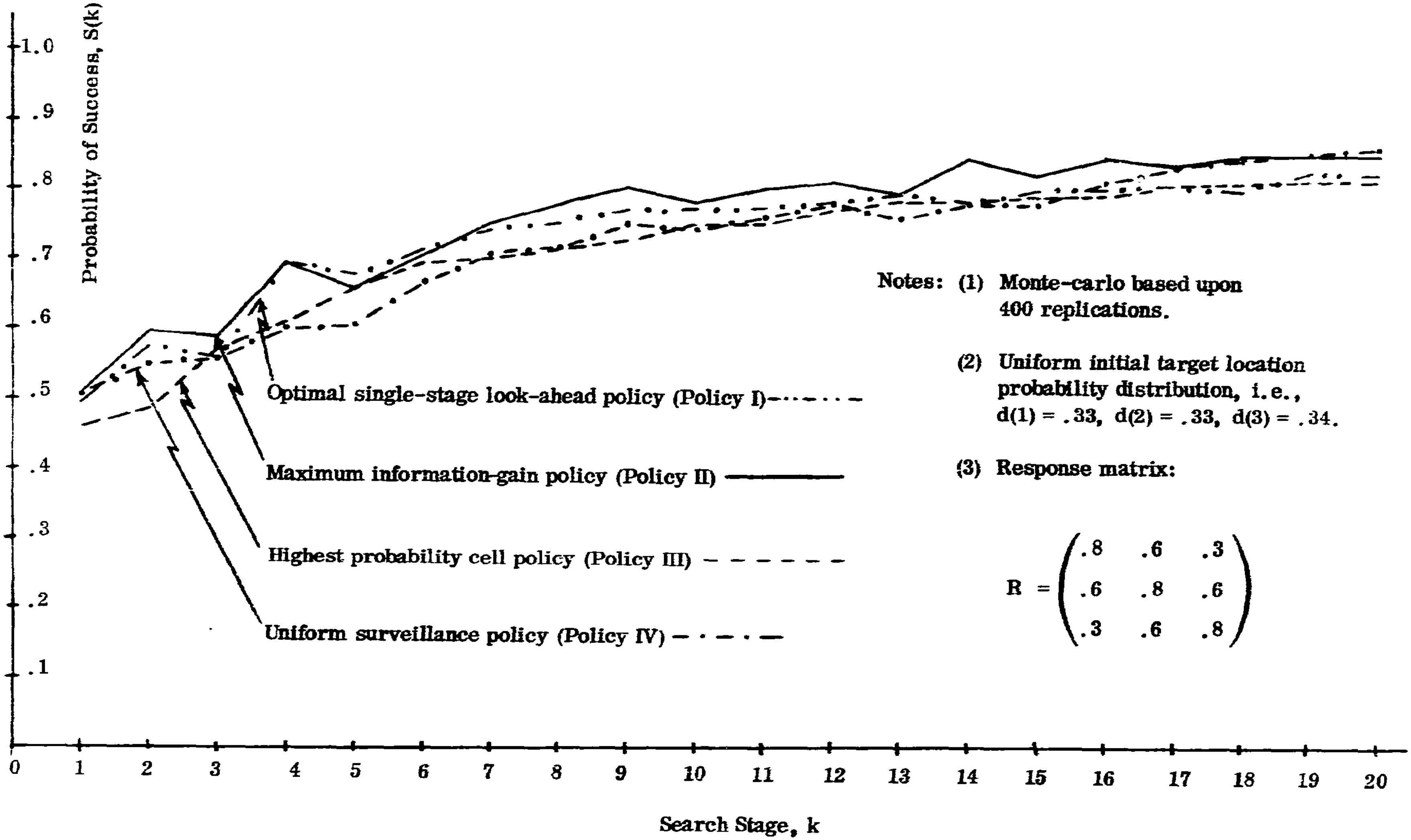
The optimal single-stage look-ahead policy will also have difficulty in this case, and, in fact, the probability of success function for this plan does not appear to converge to 1. The reason for this is that a state of indeterminacy is eventually reached by this plan; this behavior was previously noted in Case I and illustrated in Table III-1. When the before-search target location probability distribution P_B is driven to the state where the after-search selection of the target cell is the same regardless of which cell is searched or what response is obtained, then the cell with the largest before-search probability is searched. However, if the highest probability cell is among the first 8, then no information is gained by the search and the after-search probability distribution is the same as the before-search probability distribution. This means that the same cell will be searched continually in succeeding stages and progress will stop.

Conclusions and Related Operations Research Studies

The principal conclusion based upon the numerical examples in the preceding section is that the maximum information-gain policy (Policy II) appears to have very desirable characteristics in the idealized surveillance scenario considered. Among these characteristics (as measured by the success function) are good initial behavior in the early stages and good asymptotic behavior in the later stages. The initial behavior is measured principally by comparison with the optimal single-stage look-ahead policy (Policy I) which is designed to be good in the early stages. The asymptotic behavior is measured principally by comparison with the uniform surveillance policy (Policy IV) which, for a stationary target, is guaranteed to converge to 1 as the number of stages increases indefinitely (provided the rows of R are distinct).

FIGURE III-7

COMPARISON OF SURVEILLANCE POLICIES (CASE IV)



Notes: (1) Monte-carlo based upon 400 replications.

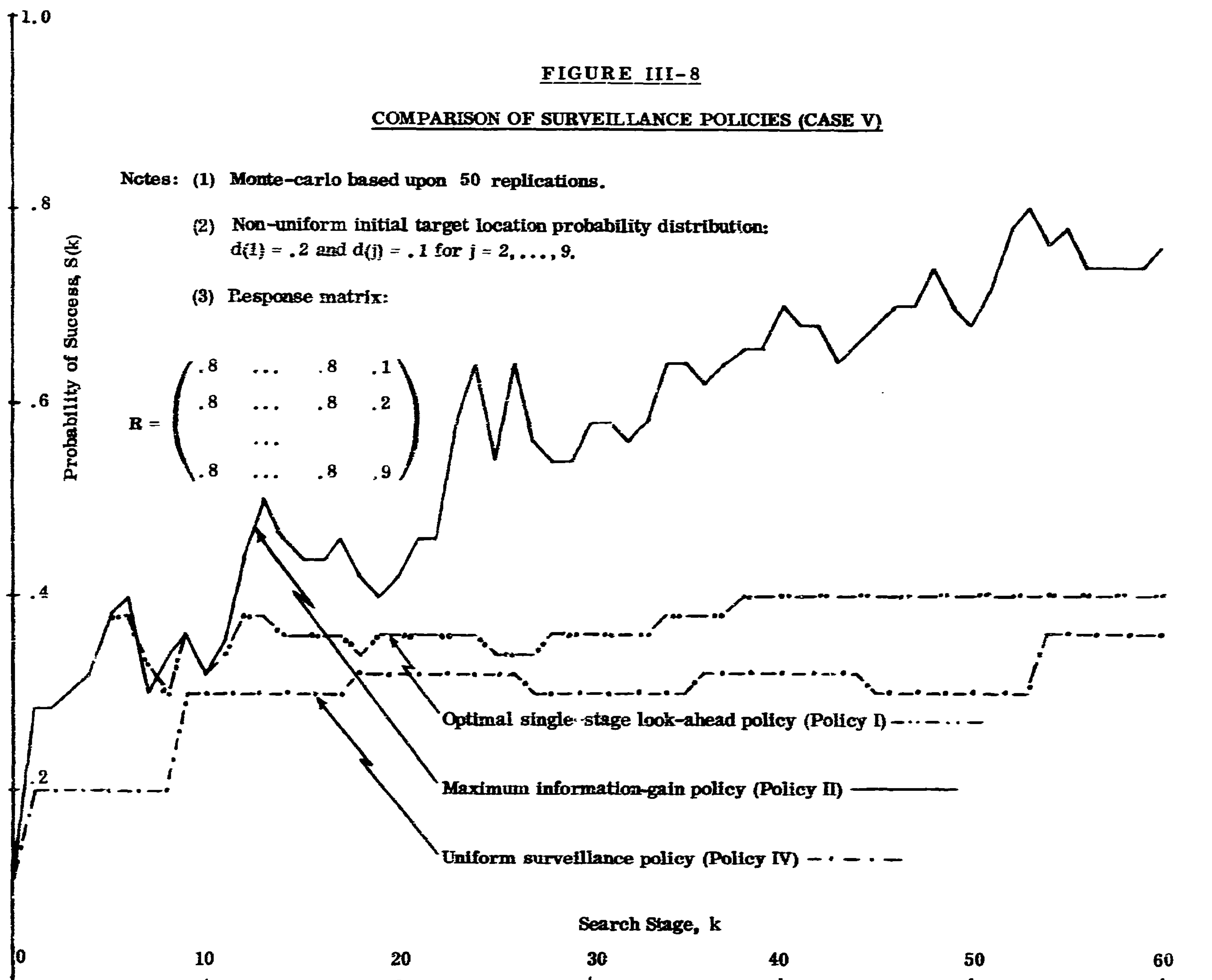
(2) Uniform initial target location probability distribution, i.e., d(1) = .33, d(2) = .33, d(3) = .34.

(3) Response matrix:

$$R = \begin{pmatrix} .8 & .6 & .3 \\ .6 & .8 & .6 \\ .3 & .6 & .8 \end{pmatrix}$$

FIGURE III-8

COMPARISON OF SURVEILLANCE POLICIES (CASE V)



We conjecture that Policy II will also perform well in cases where an incorrect prior distribution d is used, i.e., that Policy II is robust with respect to errors in d. More analysis is needed to verify this conjecture, however.

Policy I appears to have very good behavior until the "saturation" period is reached (see Table III-1) where the search cell can no longer be uniquely chosen by picking the cell which maximizes the value of the single-stage success probability function B. The ad hoc rule of choosing the highest probability cell at this point produces poor asymptotic performance in some situations (see Case V) and a better rule should be employed. Switching to randomized uniform surveillance at the saturation point would be an improvement.

Even better, perhaps, would be the extension of Policy II to optimal multi-stage look-ahead policies. In this regard, the theory of optimal stochastic control might offer some useful insights.

The highest probability cell policy (Policy III) has little to commend it, in general, although in certain special cases (e.g., Case IV) it may produce satisfactory results. Its poor behavior, in general, results from the fact that it does not make good use of the information in the response matrix.

It should be noted that none of the policies considered make non-trivial use of the information in the Markov transition matrix D. It seems worthwhile to formulate and evaluate surveillance plans which anticipate target motion by explicit consideration of D or, more generally, consideration of whatever stochastic mechanism is used for updating target location.

In the results presented in this chapter, it has been assumed that the response matrix R is known exactly. Since this is unlikely to be true in practice, it would be useful to relax this assumption and develop policies which estimate R and target location simultaneously. This kind of adaptive estimation (see reference [c]) is illustrated in the examples in Chapter II (without optimization considerations, however); there the single-glimpse probability of detection P_D is treated as a random variable and estimated from the sensor observations.

In view of the good performance of Policy II based upon maximizing the expected information gain in the after-search target location probability distribution, it is somewhat surprising that there has been so little utilization of information theory in search and surveillance problems in operations research. In fact, the relevant work which has been carried out and reported in the literature has not reflected favorably upon the use of information theory as a tool for the analysis of these problems.

One of the earliest readily accessible papers on the subject (reference [h] which appeared in 1961) discusses the connection between information theory and search theory and concludes with the statement "Thus, search theory should be considered in connection with the general theory of statistical decisions rather than with information theory." This statement is repeated and reaffirmed with further examples in reference [i] which appeared in 1971.

Both references [h] and [i] examine the search plans which maximize expected information gain. As discussed below, we believe this is the correct approach for surveillance but not for search.

In reference [j] which appeared in 1968, it is stated that "Ever since the mid-nineteen-forties when the theories of information and of search became subjects of general interest, attempts have been made to apply the theory of information to problems of search. These have proved disappointing; neither the formulas nor the concepts of the former theory have found a place in clarifying the problems of the latter."

Why do our results convey the opposite impression? The answer, we believe, is that one must make a clear distinction between search, where the objective is detection of the target, and surveillance, where the objective is knowledge of the target's location. The concepts of information theory can be applied to both types of problems but in different ways.

For the search problem (but not the surveillance problem treated in this chapter), we believe that the proper way to draw the connection between information theory and search theory is to think of an optimal search plan as one which maximizes (rather than minimizes) the entropy of the posterior target location probability distribution. Viewed this way (which is different from the approach of references [h] and [i]), search effort is used to extract information from the distribution rather than to add information to the distribution.

For the surveillance problem, however, it seems appropriate to maximize the information gain (minimize entropy). This is especially true in multi-stage scenarios, such as those we have examined in this chapter, where success can be achieved without detection of the target in the usual sense. The scenarios discussed in references [h] and [i] are limited to a single stage and thus the time behavior of the search policies is not apparent. Another point of difference between our analysis and those of references [h], [i], and [j] is that the latter do not consider the possibility of false responses.

Reference [v] makes use of information theory concepts to consider the optimal distribution of reconnaissance effort against targets in the presence of decoys. This analysis is addressed to aerial reconnaissance against land targets and is closely related to our present study. There is, however, an important point of difference which is discussed below.

Part I of reference [v] has the most in common with our present study. Part II considers questions related to enemy hindrance of the operations--a problem which we do not consider.

The basic assumption in reference [v] is that there are J regions (cells) and that in the i^{th} region there are "Npossible objects" with a priori probabilities $p_1^j, \ldots, p_N^j$ is subject to the constraint

$$\sum_{n=1}^{N} p_n^j \quad 1 \qquad \text{for } j \quad 1, \ldots, N. \tag{III-5}$$

For example, in one important case, $N = 3$ and the objects are a missile installation, a decoy, and nothing of interest.

The uncertainty in the j^{th} region is defined in reference [v] to be

$$U^j(p_1^j, \ldots, p_N^j) = -c^j \left(\sum_{n=1}^{N} p_n^j \log p_n^j \right)$$

where c^j is some positive constant. The uncertainty in the entire map is defined to be

$$U(p_1^1, \ldots, p_N^1; \ldots; p_1^J, \ldots, p_N^J) = \sum_{j=1}^{J} U^j(p_1^j, \ldots, p_N^j). \tag{III-6}$$

Reference [v] introduces and discusses assumptions pertaining to the optimal allocation of reconnaissance effort in order to minimize the uncertainty given by equation (III-6) subject to the constraint given by equation (III-5).

Although closely related to our problem, a critical difference is that we also make use of the knowledge that there is a single target present in the area of interest. This is an extremely important piece of information for it allows sensor responses and other information obtained on scene to be correlated with target motion considerations.

In the scheme of reference [v], the case where it is known that there is a single target present corresponds to N 2, where p_1^j is the a priori probability that the target is present in the j^{th} cell and $p_2^j = 1 - p_2^j$. The complication arises from the additional constraint that

$$\sum_{j=1}^{J} p_1^j = 1.$$

This constraint is necessary and important in our tactical setting but, unfortunately, it transforms the separable allocation problem considered in reference [v] into a non-separable problem. In general, non-separable problems are much more difficult to solve than separable problems. In this chapter, we have avoided this difficult allocation problem by restricting the search policy to examination of a single cell at each stage. Further work is needed to devise efficient computational algorithms for obtaining optimal multi-cell allocations which maximize the expected information gain.

REFERENCES

[a] False Target Considerations in ASW, Daniel H. Wagner, Associates Report to Naval Analysis Programs, Office of Naval Research, by C. A. Persinger and B. H. Rhodes, dated January 10, 1972, Unclassified.

[b] Computer-Assisted Search and Rescue System, Daniel H. Wagner, Associates Memorandum to Advanced Technology Division, Commandant, U. S. Coast Guard, by H. R. Richardson, dated May 28, 1970, Unclassified.

[c] Uncertain Dynamic Systems, F. C. Schweppe, Prentice-Hall, Inc., 1973.

[d] Mathematical Foundations of Information Theory, A. I. Khinchin, Dover Publications, Inc., 1957.

[e] "Optimal Whereabouts Search," J. B. Kadane, Operations Research 19, (1971), pp. 894-904.

[f] "On a Measure of the Information Provided by an Experiment," by D. V. Lindley, Annals of Mathematical Statistics, Vol. 26, 1956, pp. 686-704.

[g] "Operations Analysis During the Underwater Search for Scorpion," H. R. Richardson and L. D. Stone, Naval Research Logistics Quarterly, June 1971, Vol. 18, No. 2, pp. 141-157.

[h] "Information Theory and Search Theory as Special Cases of Decision Theory," D. F. Mela, Operations Research 9 (1961), pp. 907-909.

[i] "Search Detection and Subsequent Action: Some Problems on the Interfaces," S. M. Pollock, Operations Research, May-June 1971, pp. 559-585.

[j] Search and Information Theory, Part of Final Report on Stochastic Processes in Certain Naval Operations, B. O. Koopman, Columbia University.

[k] Information Theory and Statistics, by S. Kullback, Dover Publications, New York, 1968.

Preceding page blank

[l] Theory of Cumulative Detection Probability, Daniel H. Wagner, Associates Report to Naval Underwater Sound Laboratory, by E. P. Loane, H. R. Richardson, and E. S. Boylan, dated November 10, 1964, Unclassified.

[m] Cumulative Detection Probability for Continuous-Parameter Stochastic Processes, Daniel H. Wagner, Associates Report to Naval Air Development Center, by E. C. Curtis, H. R. Richardson, and E. P. Loane, dated July 30, 1966, Unclassified.

[n] Cumulative Detection Probability for Passive Sonobuoy Fields, Daniel H. Wagner, Associates Report to U. S. Naval Air Development Center, by L. K. Arnold and H. R. Richardson, dated December 14, 1967, Unclassified.

[o] Detection Models for Air ASW Simulation Programs, Daniel H. Wagner, Associates Memorandum to Naval Air Development Center, by L. K. Arnold, dated December 8, 1969, Unclassified.

[p] A Comparison of Detection Models Used in ASW Operations Analysis (U), Daniel H. Wagner, Associates Report to Office of Naval Research, by B. J. McCabe and B. Belkin, dated October 31, 1973, Confidential.

[q] "Optimal Search With Uncertain Sweep Width," by H. R. Richardson and B. Belkin, Operations Research, Vol. 20, No. 4, July-August, 1972, pp. 764-784.

[r] Manual for the Operations Analysis of Deep Ocean Search, Supervisor of Salvage Publication NAVSHIPS 0994-010-7010, by H. R. Richardson and L. D. Stone (in collaboration with F. A. Andrews, consultant to Ocean Systems, Incorporated), dated March 25, 1971, For Official Use Only.

[s] Random Processes, M. Rosenblatt, Oxford University Press, 1962.

[t] Introduction to Stochastic Control, H. Kushner, Holt, Rinehart and Winston, Inc., 1971.

[u] An Optimum Discrete Space Sequential Search Procedure Which Considers False Alarm and False Dismissal Instrument Errors, M. W. Smith, Technical Report No. 35, Department of Statistics THEMIS Contract, Southern Methodist University.

[v] "A Theory of Reconnaissance: I," J. M. Danskin, Operations Research, May-June 1962, pp. 285-299.

[w] Nonparametric Methods in Statistics, by D. A. S. Fraser, John Wiley & Sons, New York, 1957.

[x] "Sequential Design of Experiments," by H. Chernoff, Annals of Mathematical Statistics, Vol. 30, 1959, pp. 755-770.

[y] "Uncertainty, Information, and Sequential Experiment," by M. H. DeGroot, Annals of Mathematical Statistics, Vol. 33, 1962, pp. 404-419.

[z] "Discrimination Among Mechanistic Models," by G. E. P. Box and W. J. Hill, Technometrics, Vol. 9, 1967, pp. 57-71.

APPENDIX A

GENERALIZED TREATMENT OF THE INFORMATION PROCESSING SYSTEM

This appendix presents a generalized treatment of the mathematical technique used to calculate the examples in Chapter II. The general mathematical model is presented and discussed in the first section. The second section specializes the analysis to Markov models. Among other things, this specialization leads to the development of recursive computational procedures. The next section discusses the mathematical model of Chapter II in terms of the more general formalism presented in this appendix. This is followed by two sections addressed, respectively, to reduction of state space dimensionality and numerical computation.

General Mathematical Model

In this section the term "model" will be used to refer to the N_M-dimensional vector-valued process $\tilde{M} = (\tilde{m}_1, \ldots, \tilde{m}_{N_M})$ whose components comprise all of the stochastic processes which are relevant to the information processing situation under consideration.

The probability structure for the model is given by the triple (Ω, α, Pr) where Ω is the probability space, α is a σ-field* of subsets of Ω, and Pr is a probability distribution (a measure) defined on α. Thus,

$$\tilde{M}(t) : \Omega \to S$$

for $0 \leq t < \infty$ where the "state space" S is an N_M-dimensional Euclidean space. The σ-field of Borel sets of S is denoted by β.

The model $\tilde{M}$ consists of "observable" and "non-observable" stochastic processes, which we explain in turn.

---

* All random variables are α-measurable functions (perhaps vector-valued) defined on Ω. When explicit dependence on $\omega \in \Omega$ is shown, ω will appear as the last argument on the right, e.g., $\tilde{M}(t, \omega)$ is abbreviated as $\tilde{M}(t)$.

An observable process is associated with a physical phenomenon whose characteristics can be expressed in quantitative terms and can be assumed known to the processing system. The response processes of acoustic and non-acoustic sensors are specific examples of observable processes which are of particular interest in ASW. In simplest terms, sensor responses may be treated as (0, 1)-valued processes where 1 denotes a response and 0 denotes no response. The specification of a 0 or 1 for a sensor at any particular time might be the result, for example, of a human judgment or the output of an automatic classification device.

In more complicated formulations, the observable processes might correspond to more basic quantities such as voltages generated by the sensor hydrophones.

Regardless of complexity, however, the probability structure (Ω, α, Pr) must be established explicitly so that, among other things, one may compute the probabilities associated with the events associated with the mutual interactions of stochastic processes within the model.

The nonobservable stochastic processes consist both of "physical processes" which describe, for example, target radiated noise, position, course, and speed, and "non-physical processes" which are required to insure that the model is logically self-consistent and possesses certain desirable mathematical properties. The non-physical processes are not susceptible to physical measurement and verification in the way that the physical processes are, but nevertheless play an essential role in the operation of the information processor.

An example of a commonly used non-physical process is the time-correlated stochastic process often introduced in models to represent random fluctuations in the signal-to-noise ratio of acoustic sensors (see reference [p]). This time-correlated process is a logical necessity in cases where sensors are observed continuously since otherwise unreasonable results are obtained, e.g., if one assumes the random variables of the fluctuation process are mutually statistically independent in time (white noise).

The processes of the model $\tilde{M}$ are ordered so that $\tilde{M}$ $(\tilde{U}, \tilde{V})$, where the non-observable processes of $\tilde{M}$ are collected into one N_U-dimensional vector-valued process $\tilde{U}$ $(\tilde{u}_1, \ldots, \tilde{u}_{N_U})$, and the observable processes are collected into one N_V-dimensional vector-valued process $\tilde{V}$ $(\tilde{v}_1, \ldots, \tilde{v}_{N_V})$.

At times it will be useful to write S as the Cartesian product of the N_U-dimensional space S^1 and the N_V-dimensional Euclidean space S^2, i.e., to write S $S^1 \times S^2$. A superscripted symbol for a point or a set will indicate membership in S^1 or S^2. Points and sets without superscripts will generally be associated with S. For example, we might write $A^1 \subset S^1$, $A^2 \subset S^2$, and A $A^1 \times A^2 \subset S^1 \times S^2$.

It is assumed that the observable processes are monitored at discrete time instants $\tau_1 < \tau_2 < \ldots$ and that $\tau_1 > 0$. Continuous observations over intervals of time are not considered here, but with somewhat more effort they could be included within the general processing framework under discussion.

Let θ_t be the sub σ-field of α where θ_t is generated by the collection of observable random variables $\{ \tilde{V}(\tau_k) : \tau_k \leq t \}$. Events in θ_t correspond to observations which have occurred at or before time t. In the most abstract terms, we are interested in calculating the conditional probabilities (Pc denotes a conditional probability operator)

$$\text{Pc}\{ A \mid B \} = \frac{\Pr\{ A \cap B \}}{\Pr\{ B \}} \qquad \text{(A-1)}$$

whenever $A \in \alpha$, $B \in \theta_t$, and $\Pr\{ B \} > 0$. More generally, if $p_c\{ A \mid \theta_t \}$ denotes the conditional probability of $A \in \alpha$ given the σ-field θ_t, then $p_c\{ A \mid \theta_t \}$ is a θ_t-measurable function and

$$\Pr\{ A \cap C \} = \int_C \text{pc}\{ A \mid \theta_t \} \, d\Pr,$$

for all $C \in \theta_t$. This formulation $\text{pc}\{ A \mid \theta_t \}$ of conditional probability is required, for example, in cases where $\Pr\{ B \} = 0$ in equation (A-1).

It is not usually necessary to compute $\text{Pc}\{ A \mid B \}$ (or $\text{pc}\{ A \mid \theta_t \}$) for all $A \in \alpha$. Substantial reduction of computing cost and computer memory can be achieved if events A are restricted to smaller σ-fields. Eventually, in fact, we will restrict attention to computation of $\text{Pc}\{ A \mid B \}$ for $A \in \phi_t$, where ϕ_t is the sub σ-field generated by the non-observable random variable $\tilde{U}(t)$. Notice that ϕ_t pertains only to events which are associated with $\tilde{U}$ at a single time t.

Markov Models

In order to develop efficient recursive computational procedures, it is useful to structure $\tilde{M}$ as a Markov process. This is not as restrictive as it may seem, since in many cases what appears to be a non-Markov process can be transformed into a Markov process by enlargement of the state space and by other devices.

Let us assume that $\tilde{M}$ is Markovian and that $G(0, \cdot)$ denotes the initial probability measure of $\tilde{M}$ induced on the state space S and that Γ denotes the Markov transition function. Let t_1 and t_2 denote two instants of time ($t_1 \leq t_2$). The transition function has the following properties by definition:

(1) $\Gamma(t_1, X\ ;\ t_2, \cdot)$ is a probability distribution on β-measurable subsets of S for $X \epsilon$ S.

(2) $\Gamma(t_1, \cdot\ ;\ t_2, A)$ is a β-measurable function on S for each measurable subset A of S.

(3) Γ satisfies the Chapman-Kolmogorov integral equation, i.e., if $t_1 < t' < t_2$, then

$$\Gamma(t_1, X\ ;\ t_2, A) = \int_S \Gamma(t_1, X\ ;\ t', dY)\ \Gamma(t', Y\ ;\ t_2, A).$$

Let $G(t, \cdot)$ denote the probability distribution induced on S by $\tilde{M}(t)$ conditioned upon the observations which have taken place at times up to and including t, i.e., given events in the σ-field θ_t.

For sets in θ_t of the form $\{ \omega : \tilde{V}(\tau_1, \omega) \epsilon A_1^2, \ldots, \tilde{V}(\tau_\eta, \omega) \epsilon A_\eta^2$ and $0 \leq \tau_1 < \tau_2 < \ldots < \tau_\eta \leq t \}$, the Markov structure permits expression of $G(t, B)$ explicitly in terms of the functions $G(0, \cdot)$ and Γ. Letting $A_j = S^1 \times A_j^2$ for $j = 1, \ldots, \eta$, $G(t, B)$ is given by

$$\frac{\int_S \int_{A_1} \cdots \int_{A_\eta} G(0, dX_0)\ \Gamma(0, X_0; \tau_1, dX_1) \cdots \Gamma(\tau_{\eta-1}, X_{\eta-1}; \tau_\eta, dX_\eta)\ \Gamma(\tau_\eta, X_\eta; t, B)}{\int_S \int_{A_1} \cdots \int_{A_\eta} \int_S G(0, dX_0)\ \Gamma(0, X_0; \tau_1, dX_1) \cdots \Gamma(\tau_{\eta-1}, X_{\eta-1}; \tau_\eta, dX_\eta)}, \quad \text{(A-2)}$$

where the denominator is assumed to be non-zero (this is always true in our applications).

In most situations of interest, equation (A-2) does not lend itself to easy computation. There are two principal problems. The first is that the state space S has very high dimension, and the second reason is that the functions $G(0, \cdot)$ and Γ are not usually conveniently expressed in terms of mathematical formulas. For example, the transition function Γ might be expressed in terms of scenarios which specify the stochastic assumptions for target behavior in the mission under consideration. As in Chapter II, these statements are most directly translated into monte-carlo computer programs, rather than into "analytical formulas" suitable for substitution in equation (A-2).

The Model of Chapter II

In order to motivate the introduction of additional mathematical structure for the purpose of overcoming the two problems stated above, the following three subsections will describe, in the formalism of this appendix, the model used to calculate the examples given in Chapter II.

The principal objective in Chapter II is to compute and update the probability distributions for target location making use of target-motion scenarios and sensor response data. To do this, the components of the observable process $\tilde{V} = (\tilde{v}_1, \ldots, \tilde{v}_{N_V})$ are defined to be (0, 1)-stochastic processes which describe the time history of the sensor responses, i.e., for $1 \leq n \leq N_V$,

$$\tilde{v}_n(t) = \begin{cases} 1 & \text{if the } n^{th} \text{ sensor is responding at time t, and} \\ 0 & \text{otherwise.} \end{cases}$$

Target motion assumptions. Let $\tilde{z}_1$ and $\tilde{z}_2$ denote the stochastic processes for target latitude and longitude, respectively. In the model used in the examples, the two-dimensional target location stochastic process $(\tilde{z}_1, \tilde{z}_2)$ by itself is not Markovian. However, by addition of the target velocity stochastic process $(\tilde{s}_1, \tilde{s}_2)$ and the scenario random variable $\tilde{k}$, the augmented five-dimensional $\tilde{Z} = (\tilde{z}_1, \tilde{z}_2, \tilde{s}_1, \tilde{s}_2, \tilde{k})$ becomes Markovian. In other words, given $\tilde{Z}(t)$, the future $\{ \tilde{Z}(t') \}_{t' > t}$ is statistically independent of the past $\{ \tilde{Z}(t') \}_{t' < t}$.

The transition function for $\tilde{Z}$ is specified in terms of the scenario descriptions and is realized by monte-carlo simulation (see the discussion of program TRANS in Chapter II).

Sensor-response assumptions. This subsection presents the sensor-response assumptions for the model of Chapter II.

The single-sensor, single-glimpse probability of detection and false response are assumed in Chapter II to be themselves random variables. This is done to call attention to the fact that in most ASW situations there is not sufficient information about target characteristics, sensor performance, and environmental conditions (including non-target shipping) to provide high confidence inputs to a detection or false-response calculation. Using the methodology presented in Chapter II, one begins with an initial probability distribution for the uncertain parameters and then modifies this distribution adaptively by utilizing the information obtained from the sensor responses. In the language of systems theory (see reference (c)), this is an example of adaptive state estimation and system identification.

In the illustrative model, the sensor response (0, 1)-random variables $\{ \tilde{v}_n(\tau_k) : \text{for } 1 \leq n \leq N_V \text{ and } 1 \leq k \leq \eta \}$ are assumed to be mutually statistically independent, conditioned upon knowledge of the target location and the detection and false-response probability random variables $\tilde{P}_D$ and $\tilde{P}_A$. That is*,

* ξ (no tilde) denotes a specific value of a random variable $\tilde{\xi}$.

$$\text{Pc}\left\{\begin{matrix}\tilde{V}_n(\tau_\eta) = V_n(\tau_\eta) \\ \text{for } 1 \leq n \leq N_V\end{matrix} \middle| \begin{matrix}\tilde{V}_i(\tau_k) \text{ for } 1 \leq i \leq N_V \text{ and } 1 \leq k < \eta, \\ \text{and } \tilde{z}_1(\tau_\eta),\ \tilde{z}_2(\tau_\eta),\ \tilde{P}_A,\ \tilde{P}_D\end{matrix}\right\}$$

$$= \prod_{n=1}^{N_V} \text{Pc}\{\tilde{V}_n(\tau_\eta) = V_n(\tau_\eta) \mid \tilde{z}_1(\tau_\eta),\ \tilde{z}_2(\tau_\eta),\ \tilde{P}_A,\ \tilde{P}_D\}. \qquad \text{(A-3)}$$

Thus, knowledge of the target position at time τ_η and of the probabilities of detection and false response make the current observation random variables mutually statistically independent and independent of their past values.

The observable process $\tilde{V}$ by itself is usually not Markovian because, among other things, the sensor responses depend upon the target's position which is not observable. However, when the unobservable process $\tilde{U}$ is also specified, the model $\tilde{M} = (\tilde{U}, \tilde{V})$ is Markovian even though $\tilde{V}$ is not.

The unobservable process. The unobservable process for the model of Chapter II is defined by

$$\tilde{U}(t) = (\tilde{z}_1(t), \tilde{z}_2(t), \tilde{s}_1(t), \tilde{s}_2(t), \tilde{k}, \tilde{P}_D) = (\tilde{Z}, \tilde{P}_D).$$

The process $\tilde{U}$ (but not the observable process $\tilde{V}$) is Markovian, since, as we have noted, if $\tilde{Z}(t)$ is known, then the statistical properties of $\tilde{Z}(t')$ are determined for all $t' > t$ and do not depend upon values of $\tilde{Z}$ before time t. The random variables $\tilde{k}$ and $\tilde{P}_D$ are not time dependent and, hence, constitute trivial Markov processes.

Reduction of State Space Dimensionality

The assumptions in this section are made in order to reduce state space dimensionality. Large state space dimensionality is one of the problems previously mentioned concerning the evaluation of equation (A-2).

Let A^i be a β^i-measurable subset of S^i for $i = 1$ and 2 (β^i is the Borel field of S^i). Assume that $\tilde{U}$ is Markovian and that $G(0, \cdot)$ and Γ may be expressed in the special form

$$G(0, A^1 \times A^2) = \int_{A^1} L(0, dX_0^1)\, H(X_0^1, A^2)$$

and

$$\Gamma(t_1, X_{t_1}, t_2, A^1 \times A^2) \quad \int_{A^1} \Lambda(t_1, X^1_{t_1}, t_2, dX^1_{t_2})\, H(X^1_{t_2}, A^2)$$

where $L(0,\cdot)$ and Λ are, respectively, the initial probability distribution and the transition function for the unobservable Markov process $\tilde{U}$. The value $H(X^1, A^2)$ is the conditional probability that $\tilde{V}(t) \in A^2$ given that $\tilde{U}(t) = X^1$. The function $H(X^1, \cdot)$ is assumed to be a probability distribution on β^2 for every $X^1 \in S^1$ and $H(\cdot, A^2)$ is assumed to be a β^1-measurable function on S^1 for each β^2-measurable set $A^2 \subset S^2$.

Under the above assumptions, it can be shown that for subsets of S of the form $B^1 \times S^2$, where B^1 is any β^1-measurable subset of S^1, the function $G(t,\cdot)$ defined by equation (A-2) can be rewritten

$$G(t, B^1 \times S^2) = \kappa_t \int_{S^1} \int_{S^1} \cdots \int_{S^1} \left[\prod_{j=1}^{\eta} H(X^1_j, A^2_j) \right] L(0, dX^1_0) \tag{A-4}$$

$$\Lambda(0, X^1_0; \tau_1, dX^1_1) \cdots \Lambda(\tau_{\eta-1}, X^1_{\eta-1}; \tau_\eta, dX^1_\eta)\, \Lambda(\tau_\eta, X^1_\eta; t, B^1)$$

for $t > 0$, where κ_t is a normalizing constant defined so that $G(t, S^1 \times S^2) = 1$.

The significance of equation (A-4) is that probabilities associated with the unobservable process $\tilde{U}$ and conditioned upon the observable process $\tilde{V}$ may be computed by integrations over the state space of S^1 of $\tilde{U}$ rather than by integrations over the state space S of $\tilde{M}$ as required by equation (A-2). Among other things, this decreases the amount of computer memory required for processing the data and usually can be expected to increase computing speed.

Another advantage is that equation (A-4) may be computed recursively. Let $B^1 \in \beta^1$ and $A^2_j \in \beta^2$ for $1 \leq j \leq \eta$. Further, let $t > 0$ and $\tau_0 = 0$. For notational convenience, define $L(t, B^1) \; G(t, B^1 \times S^2)$. Then one can show that for $j \leq \eta - 1$

$$L(\tau_{j+1}, B^1) = \kappa_{\tau_{j+1}} \int_{S^1} \int_{B^1} L(\tau_j, dX^1_j)\, \Lambda(\tau_j, X^1_j, \tau_{j+1}, dX^1_{j+1})\, H(X^1_{j+1}, A^2_{j+1}), \tag{A-5}$$

and for $t > \tau_\eta$ (the last time of sensor observation)

$$L(t, B^1) \;\; \kappa_2 \int_{S^1} L(\tau_\eta, dX^1_\eta)\, \Lambda(\tau_\eta, X^1, t, B^1). \tag{A-6}$$

The factor $\kappa_{\tau_{j+1}}$ appearing in equation (A-5) is a normalizing constant defined so that $L(\tau_{j+1}, S^1) = 1$.

Equations (A-5) and (A-6) indicate that, in a sense, all relevant past information about target motion and sensor response is contained in the most recent probability distribution $L(t,\cdot)$ defined on S^1.

Numerical Computation

In applications of these concepts to large-scale, multi-sensor, multi-platform operations, the conditional probabilities $L(t, B^j)$ have been computed from equations (A-5) and (A-6) by using monte-carlo simulation of $\tilde{U}$ and analytic determination of $H(X_j^1, A_j^2)$. This section briefly outlines these computational procedures in terms of an idealized computer-processing system.

The principal advantages of the computational procedures discussed in this section and employed in Chapter II are as follows:

(1) Realistic target motion scenarios and descriptions of sensor behavior may be used when formulating the processing algorithms since monte-carlo simulation reduces the need for introducing artificial mathematical assumptions in order to obtain closed-form solutions.

(2) A minimum of computer core memory is required, since most data are stored peripherally and processed sequentially.

(3) In many cases, certain expressions can be precomputed, making use of existing models such as the large-scale ASW simulation models APAIR and APSURV. Off-line precomputation, when feasible, results in rapid processing, which is particularly useful in real-time tactical applications.

The reader should refer to the section of Chapter II entitled "Information Processing Procedures" for a more detailed discussion in terms of the illustrative model.

All information concerning past target movements and sensor responses is contained in two external files UFILE and WGHT. The processing consists of reading these files into the computer in parallel and updating records a pair at a time, one from each file.

Let UFILE(t) and WGHT(t) denote the contents of UFILE and WGHT, respectively, at time t.

The file UFILE(t) contains N_r monte-carlo samples of $\tilde{U}(t)$, and the file WGHT(t) contains N_r "weights," each pertaining to the corresponding record of UFILE(t).

Let $\hat{U}^n(t)$ denote the n^{th} simulated sample function of $\tilde{U}(t)$ for $0 \leq t$ and $1 \leq n \leq N_r$, where N_r denotes the number of replications. Since $\tilde{U}$ is Markovian, knowledge of $\hat{U}^n(t)$ statistically determines the values of $\hat{U}^n(t')$ for $t' > t$ without reference to values of $\hat{U}^n(t')$ for $t' < t$. For $1 \leq n \leq N_r$, the n^{th} weight $\hat{w}^n(t)$ contained in WGHT(t) is the probability $\prod_{j=1}^{\eta} H(\hat{U}^n(\tau_j), A_j^2)$ based upon the observed sensor responses.

START denotes the computer program which creates UFILE(0) and WGHT(0). The file UFILE(0) is created by generating N_r monte-carlo samples from the initial probability measure $L(0, \cdot)$ of $\tilde{U}$.

Since by definition no observations are associated with time $t = 0$, each record of the initial file WGHT(0) contains the probability 1. These weights indicate that at time $t = 0$, all samples of UFILE(0) are considered equally likely.

Now suppose that UFILE(t_1) and WGHT(t_1) associated with time t_1 are to be updated to time t_2. As above, the times at which observations are obtained are denoted $\tau_1, \ldots, \tau_\eta$, and $\tau_\eta \leq t_2$ represents the time of the most recent observation. Assume that $t_1 < \tau_1 < \ldots < \tau_\eta \leq t_2$. The first step is to update UFILE(t_1) to time τ_1.

Let TRANS denote a computer program which updates UFILE by implementing the transition function Λ. Let $t' = t_1$ and $\tau' = \tau_1$. The first record $\hat{U}^1(t')$ of UFILE(t') is read into the computer. The probability distribution $\Lambda(t', \hat{U}^1(t'), \tau', \cdot)$ is then sampled by monte-carlo and the result $\hat{U}^1(\tau')$ becomes the first record of UFILE(τ'). This procedure is repeated for each record of UFILE(t') until all records have been updated to τ'.

The next step is to update the file WGHT(t'). OBSERV denotes the idealized computer program for this purpose. The inputs to OBSERV include the newly created file UFILE(τ') and the file WGHT(t'); as with UFILE, updating is carried out one record at a time. A pair of values $\hat{U}^n(\tau')$ and $w^n(t')$ is then used to compute $w^n(\tau')$ using the formula

$$w^n(\tau') = w^n(t')\, H(\hat{U}^n(\tau'), A_j^2)$$

which follows from equation (A-5).

Once files UFILE(τ_{j-1}) and WGHT(τ_{j-1}) are completed for any $2 \leq j \leq \eta$, files UFILE(τ_j) and WGHT(τ_j) are obtained by repeating the procedure with $t' = \tau_{j-1}$ and $\tau' = \tau_j$. This continues until files UFILE(τ_η) and WGHT(τ_η) are generated.

If $t_2 > \tau_\eta$, then the final update consists of using TRANS to operate on file UFILE(τ_η) in order to generate UFILE(t_2) (see equation (A-6)). Since no new observations occur between times τ_η and t, WGHT(t_2) is a replica of WGHT(τ_η).

Any probability associated with the random variable $\tilde{U}(t_2)$ conditioned upon the observed process $\tilde{V}$ may be estimated using files UFILE(t_2) and WGHT(t_2) and the formula

$$\Pr\{\tilde{U}(t_2) \in B^1 \mid \tilde{V}(\tau_j) \in A_j^2 \text{ for } 1 \le j \le \eta\} \approx \frac{\sum_{n \in I(B^1)} \hat{w}^n(t_2)}{\sum_{n=1}^{N_r} \hat{w}^n(t_2)} \qquad \text{(A-7)}$$

where $I(B^1) = \{ n : \hat{U}^n(t) \in B^1 \}$.

APPENDIX B

FORMULATION OF THE SEARCH AND SURVEILLANCE PROBLEM AS A STATISTICAL SEQUENTIAL EXPERIMENTAL DESIGN PROBLEM

by

Thomas L. Corwin

The purpose of this appendix is to suggest a theoretical framework in which to relate information-theoretic concepts to surveillance in a false target environment. It is shown that the problem of which cell to search at each stage of a surveillance operation may be viewed as a game between the search planner and Nature in which the payoff to the search planner is measured in terms of the information he gains about the true state of Nature for a particular choice of cell to search. Two sequential design procedures are examined in this context.

In the first section the surveillance problem is stated as a problem in statistical hypothesis testing. In the second section some fundamental concepts of the theory of sequential experimental design and of information theory are introduced. The third section is devoted to the discussion of a general measure of the information content of an experiment, called the discriminator function. In the fourth section it is shown that the values assumed by the discriminator function may be viewed as the potential payoffs to the experimenter in the play of a certain type of two-person game. Discussion of the sequential design procedures of Chernoff and Lindley as particular examples of such games is presented in this section.

Introduction

Let a region in N-dimensional Euclidean space be divided into J non-null measurable sets θ_j (the search cells) for $1 \leq j \leq J$.

Let θ_T denote the cell containing the target. In this appendix we assume that the target is stationary. Let the parameter space be given by $\{1, \ldots, J\}$. Assume also the existence of conditional probabilities $R(j,k)$ for $1 \leq j \leq J$ and $1 \leq k \leq J$, where $R(j,k)$ is the probability of a response upon searching in θ_k given the target is located in θ_j.

It is then desired to test the following simple hypothesis against the attending composite alternatives:

$$\left.\begin{array}{ll} H_o^i : i = T & \\ H_a^i : j = T & \text{for some } j \neq i \end{array}\right\} \text{ for } 1 \leq j \leq J \text{ and } 1 \leq i \leq J.$$

In the ensuing discussion the points of the parameter space $\{1, \ldots, J\}$ will often be referred to as "states of Nature."

<u>Preliminaries</u>

Let us consider a measurable space $(\mathcal{X}, \mathcal{Z})$ [reference [w], p. 2], i.e., $\mathcal{X}$ is a basic set of elements $x \in \mathcal{X}$ and $\mathcal{Z}$ a σ-algebra of subsets of $\mathcal{X}$. We regard $\mathcal{X}$ as the sample space of an experiment and $\mathcal{Z}$ as the set of all possible events made up of elements of the sample space. Now let us consider the construction of J probability spaces. For each possible state of Nature $j \in \{1, \ldots, J\}$, let Ξ_j be a probability measure defined on $\mathcal{Z}$. We will assume that the probability measures are mutually absolutely continuous and distinct. Thus, essentially we are considering J probability spaces $(\mathcal{X}, \mathcal{Z}, \Xi_j)$, $j \in \{1, \ldots, J\}$.

For example, in the application mentioned in the introduction, the sample space for an experiment in which all of the cells $\{O_j : 1 \leq j \leq J\}$ are simultaneously searched over and in which $J = 3$, is given by $\mathcal{X} = \{x_1, x_2, \ldots, x_8\}$, where

$$\begin{aligned} x_1 &= (NR, NR, NR) \\ x_2 &= (NR, NR, R) \\ x_3 &= (NR, R, NR) \\ x_4 &= (NR, R, R) \\ x_5 &= (R, NR, NR) \\ x_6 &= (R, NR, R) \\ x_7 &= (R, R, NR) \\ x_8 &= (R, R, R). \end{aligned}$$

Here an "R" in the k^{th} entry of x_j indicates a reponse in cell k and an "NR" in the k^{th} entry of x_j indicates no response in cell k, for $1 \leq j \leq 8$.

The measures $\{\Xi_1, \ldots, \Xi_J\}$ may be constructed in this case by defining the value of Ξ_j on each element of the sample space. The value of Ξ_j on a particular element of the sample space is simply the probability that the particular sequence of R's and NR's will be observed given that the true state of Nature is j, i.e., $T = j$, (or given that the target is in cell θ_j). Thus, if it is assumed that cells are searched independently, Ξ_j in the example presented above is defined as follows:

$$\Xi_j(x_1) = [1 - R(j,1)]\,[1 - R(j,2)]\,[1 - R(j,3)]$$

$$\Xi_j(x_2) = [1 - R(j,1)]\,[1 - R(j,2)]\,R(j,3)$$

$$\Xi_j(x_3) = [1 - R(j,1)]\,R(j,2)\,[1 - R(j,3)]$$

$$\Xi_j(x_4) = [1 - R(j,1)]\,R(j,2)\,R(j,3)$$

$$\Xi_j(x_5) = R(j,1)\,[1 - R(j,2)]\,[1 - R(j,3)]$$

$$\Xi_j(x_6) = R(j,1)\,[1 - R(j,2)]\,R(j,3)$$

$$\Xi_j(x_7) = R(j,1)\,R(j,2)\,[1 - R(j,3)]$$

$$\Xi_j(x_8) = R(j,1)\,R(j,2)\,R(j,3)\ ; \qquad \text{for } 1 \le j \le J.$$

Let us now consider a set of M random variables defined on the sample space χ, denoted $\tilde{Y}_1, \ldots, \tilde{Y}_M$. In the above example if the experimenter is allowed to search only one cell, we could define a set of M = J random variables on the sample as follows:

$$\tilde{Y}_m(x_i) = \begin{cases} 1 & \text{if the } m^{th} \text{ entry in } x_i \text{ is an "R,"} \\ 0 & \text{if the } m^{th} \text{ entry in } x_i \text{ is an "NR,"} \quad \text{for } 1 \le i \le 2^J,\ 1 \le m \le J. \end{cases}$$

Let us now consider the following problem: Let us assume that for a particular experiment there are J possible states of nature. For each j, $1 \le j \le J$, there exists a probability space $(\chi, \mathfrak{C}, \Xi_j)$ and on the sample space χ are defined M random variables, $\tilde{Y}_1, \ldots, \tilde{Y}_M$. Now let us also assume that available to the experimenter are N trials in which he may observe any one of these M random variables in order to make inference about which one of the J states of Nature is the true one. In the terminology of the problem stated in the introduction, the search planner has available to him N trials in each of which he may search any of J cells in order to make a determination about the actual location of the target. In this case M = J. The problem then is to determine which random variable should be sampled at each trial in order to optimize his ability to discern the actual state of Nature at the end of N trials. In the terminology of Chapter III, the search planner is interested in maximizing S(N).

To this end, let us discuss the likelihood ratio statistic. For purposes of simplicity, let us assume that for $1 \leq m \leq M$ the random variable $\tilde{Y}_m$ is real valued. Let us also assume that the probability measure of $\tilde{Y}_m$, $1 \leq m \leq M$, is absolutely continuous with respect to some fixed measure μ defined on the Borel field of the real numbers. For each state of Nature $j \in \{1, \ldots, J\}$, let us denote the density of the random variable $\tilde{Y}_m$ by $w_{m,j}$ for $1 \leq m \leq M$. Then the likelihood ratio statistic for testing the hypothesis that j is the true state of Nature against the alternative that k is the true state of Nature is given by

$$L(j, k, m_1, \ldots, m_N, \hat{Y}_{m_1}, \ldots, \hat{Y}_{m_N}) = \sum_{n=1}^{N} \log\left[\frac{w_{m_n, j}(\hat{Y}_{m_n})}{w_{m_n, k}(\hat{Y}_{m_n})}\right]$$

for $1 \leq j \leq J$ and $1 \leq k \leq J$,

where

(i) $\hat{Y}_{m_n}$ is the sampled value of the random variable $\tilde{Y}_{m_n}$, the random variable sampled at the n^{th} step, and

(ii) $w_{m_n,j}(\cdot)$ is the density of the random variable $\tilde{Y}_{m_n}$ under the hypothesis that j is the true state of Nature.

Thus, at the n^{th} step, the increment in the likelihood ratio statistic is given by

$$\delta(j, k, m_n, \hat{Y}_{m_n}) = \log\left[\frac{w_{m_n, j}(\hat{Y}_{m_n})}{w_{m_n, k}(\hat{Y}_{m_n})}\right] \quad \text{for } 1 \leq j \leq J \text{ and } 1 \leq k \leq J.$$

Intuitively, if j is the true state of Nature, $\delta(j, k, m_n, \hat{Y}_{m_n})$ represents the additional ability, obtained through sampling Y_{m_n} at the n^{th} step, to discriminate between the hypotheses j and k. Roughly speaking, at the n^{th} step one would prefer to sample the random variable $\tilde{Y}_{m^*}$ which maximizes this increment on the average.

In the terminology of Kullback and Leibler, reference [k], given the m random variables $\{\tilde{Y}_m : 1 \leq m \leq M\}$ from which to choose at the n^{th} step, the experimenter would prefer to choose that experiment which maximizes the information number $I(j, k, m)$ defined as follows:

$$I(j,k,m) = \int_{-\infty}^{\infty} \delta(j,k,m,x)\, w_{m,j}(x)\, d\mu(x) \quad \text{for } 1 \le j \le J,\ 1 \le k \le J,\ 1 \le m \le M. \tag{B-1}$$

Implicit in this definition is the assumption that j is the true state of Nature.

The Discriminator Function

Since at the n^{th} step of any sequential scheme the experimenter, in general, does not know the true state of Nature, he is faced with the problem of choosing one of the random variables $\{\tilde{Y}_m : 1 \le m \le M\}$ to maximize his ability to discriminate between some estimate of the true parameter value T and the remaining parameters. Thus, one is led to consider maximization of analogs of the expected increment in the likelihood ratio statistic or the Kullback-Liebler information number presented in expression (B-1). The analogs to be considered here have the following form:

(i) Let $\{\lambda_i : 1 \le i \le J\}$ be a set of real numbers such that $0 \le \lambda_i \le 1$ for $1 \le i \le J$ and $\Sigma_{i=1}^{J} \lambda_i = 1$.

(ii) For each $m \in \{1, \ldots, M\}$, let $\phi(m,\cdot)$ be a density with respect to the measure μ.

Then define the "discriminator" function D as follows:

$$D(m, \phi(m,\cdot), \lambda_1, \ldots, \lambda_J) = \int_{-\infty}^{\infty} \sum_{j=1}^{J} \lambda_j \log\left[\frac{\phi(m,x)}{w_{m,j}(x)}\right] \phi(m,x)\, d\mu(x)$$

$$\text{for } 1 \le m \le M. \tag{B-2}$$

Many of the important discriminator functions discussed in the literature appear as special cases of the discriminator D, for specific choices of the function ϕ and the real numbers $\lambda_1, \ldots, \lambda_J$. For instance:

(i) Let j and k be two elements of the sample space $\{1, \ldots, J\}$; then define the function η by

$$\eta(j,m,x) = w_{m,j}(x) \quad \text{for } -\infty < x < \infty,\ 1 \le m \le M,\ 1 \le j \le J \tag{B-3}$$

and define

$$\omega(k, i) = \begin{cases} 1 & \text{for } i = k \\ 0 & \text{for } i \neq k \end{cases} \quad \text{for } 1 \leq i \leq J \text{ and } 1 \leq k \leq J. \tag{B-4}$$

Then the number $D(m, \eta(j, m, \cdot), \omega(k, 1), \ldots, \omega(k, J))$ is the expected increment in the likelihood ratio statistic for testing the hypothesis that j is the true state of Nature against the alternative that k is the true state of Nature when sampling the random variable $\tilde{Y}_m$. If i^*_{n-1} is the mode of the posterior probabilities defined on the parameter space $\{1, \ldots, J\}$ at the $(n-1)^{st}$ step and i^{**}_{n-1} is the mode of the posterior probabilities defined on the parameter space at the $(n-1)^{st}$ step restricted to the set $\{1, \ldots, J\} - \{i^*_{n-1}\}$, then the number $D(m, \eta(i^*_{n-1}, m, \cdot), \omega(i^{**}_{n-1}, 1), \ldots, \omega(i^{**}_{n-1}, J))$ is a form of the discrimination number used in Chernoff's procedure A (reference [x]) to be discussed later.

(ii) Let us modify example (i) above slightly to produce a different discrimination function D. Instead of the function ω defined in (B-4), let us use a real-valued function $\bar{\omega}$ defined on the sample space $\{1, \ldots, J\}$, satisfying

(a) $0 \leq \bar{\omega}(i) \leq 1$, for $i \in \{1, \ldots, J\} - \{i^*_{n-1}\}$

(b) $\bar{\omega}(i^*_{n-1}) = 0$

(c) $\Sigma^J_{i=1} \bar{\omega}(i) = 1$

(d) $D(m, \eta(i^*_{n-1}, m, \cdot), \bar{\omega}(1), \ldots, \bar{\omega}(J))$

$$\inf_{(\lambda_1, \ldots, \lambda_J) \in \Lambda(i^*_{n-1})} D(m, \eta(i^*_{n-1}, m, \cdot), \lambda_1, \ldots, \lambda_J),$$

where

$$\Lambda(k) = \{(\lambda_1, \ldots, \lambda_J) : 0 \leq \lambda_i \leq 1 \text{ for } i \in \{1, \ldots, J\}, \sum_{i=1}^{J} \lambda_i = 1, \text{ and } \lambda_k = 0\}. \tag{B-5}$$

and define

$$\omega(k, i) = \begin{Bmatrix} 1 & \text{for } i = k \\ & \\ 0 & \text{for } i \neq k \end{Bmatrix} \quad \text{for } 1 \leq i \leq J \text{ and } 1 \leq k \leq J. \tag{B-4}$$

Then the number $D(m, \eta(j, m, \cdot), \omega(k, 1), \ldots, \omega(k, J))$ is the expected increment in the likelihood ratio statistic for testing the hypothesis that j is the true state of Nature against the alternative that k is the true state of Nature when sampling the random variable $\tilde{Y}_m$. If i^*_{n-1} is the mode of the posterior probabilities defined on the parameter space $\{1, \ldots, J\}$ at the $(n-1)^{st}$ step and i^{**}_{n-1} is the mode of the posterior probabilities defined on the parameter space at the $(n-1)^{st}$ step restricted to the set $\{1, \ldots, J\} - \{i^*_{n-1}\}$, then the number $D(m, \eta(i^*_{n-1}, m, \cdot), \omega(i^{**}_{n-1}, 1), \ldots, \omega(i^{**}_{n-1}, J))$ is a form of the discrimination number used in Chernoff's procedure A (reference [x]) to be discussed later.

(ii) Let us modify example (i) above slightly to produce a different discrimination function D. Instead of the function ω defined in (B-4), let us use a real-valued function $\bar{\omega}$ defined on the sample space $\{1, \ldots, J\}$, satisfying

(a) $0 \leq \bar{\omega}(i) \leq 1$, for $i \in \{1, \ldots, J\} - \{i^*_{n-1}\}$

(b) $\bar{\omega}(i^*_{n-1}) = 0$

(c) $\Sigma^J_{i=1} \bar{\omega}(i) = 1$

(d) $D(m, \eta(i^*_{n-1}, m, \cdot), \bar{\omega}(1), \ldots, \bar{\omega}(J))$

$$\inf_{(\lambda_1, \ldots, \lambda_J) \in \Lambda(i^*_{n-1})} D(m, \eta(i^*_{n-1}, m, \cdot), \lambda_1, \ldots, \lambda_J),$$

where

$$\Lambda(k) = \{(\lambda_1, \ldots, \lambda_J) : 0 \leq \lambda_i \leq 1 \text{ for } i \in \{1, \ldots, J\}, \sum_{i=1}^{J} \lambda_i = 1, \text{ and } \lambda_k = 0\}. \tag{B-5}$$

As will be discussed later, the discrimination number $D(m, \eta(i^*_{n-1}, m, \cdot), \omega(i^{**}_{n-1}, 1), \ldots, \omega(i^{**}_{n-1}, J))$ represents the payoff to the experimenter in a particular game outlined in Chernoff's procedure A for his choice of random variable $\tilde{Y}_m$ at the n^{th} step when Nature may choose from among only a certain set of her pure strategies (see fourth section); while the discriminator $D(m, \eta(i^*_{n-1}, m, \cdot), \omega(1), \ldots, \omega(J))$ represents the payoff to the experimenter in the same game for his choice of random variable $\tilde{Y}_m$ when Nature is allowed to choose from among a wider class of her mixed strategies.

(iii) Let $\{\alpha^{n-1}(i) : 1 \leq i \leq J\}$ be the set of posterior probabilities defined on the parameter space $\{1, \ldots, J\}$ at the $(n-1)^{st}$ step. Then define the function E as follows:

$$E(m,x) = \sum_{i=1}^{J} \alpha^{n-1}(i)\, w_{m,i}(x) \quad \text{for } -\infty < x < \infty \text{ and } 1 \leq m \leq M.$$

Then for $1 \leq m \leq M$, the number $D(m, E(m, \cdot), \alpha^{n-1}(1), \ldots, \alpha^{n-1}(J))$ is the expected decrease in the entropy of the posterior probabilities on the parameter space at the n^{th} step given that the experimenter samples random variable $\tilde{Y}_m$. Notice here that the choice of m to optimize $D(m, E(m, \cdot), \alpha^{n-1}(1), \ldots, \alpha^{n-1}(J))$ is essentially an attempt to increase the average expected power to discriminate between the current estimate of the density of $\tilde{Y}_m$ given by $\Sigma_{i=1}^{J} \alpha^{n-1}(i)\, w_{m,i}(\cdot)$ and the densities of $\tilde{Y}_m$ under the assumption that each of the parameters $\{1, \ldots, J\}$ represents the true state of Nature.

A number of simple results regarding the function D may be proved using the results of Kullback.

THEOREM B-1. For $1 \leq m \leq M$,

$$D(m, \phi, \lambda, \ldots, \lambda_J) \geq 0. \tag{B-6}$$

<u>Proof</u>. In Theorem 3.1, p. 14, of reference [k], Kullback shows that for $1 \leq j \leq J$ and $1 \leq m \leq M$,

$$\int_{-\infty}^{\infty} \log\left[\frac{\phi(m,x)}{w_{m,j}(x)}\right] \phi(m,x)\, d\mu(x) > 0.$$

Thus, expression (B-6) follows from the non-negativity of the elements of $\{\lambda_j : 1 \le j \le J\}$, which proves the theorem.

Let $\bar{D}(m_1, m_2, \phi, \lambda_1, \ldots, \lambda_g)$ be defined by

$$\bar{D}(m_1, m_2, \phi, \lambda_1, \ldots, \lambda_J) =$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{j=1}^{J} \lambda_j \log\left[\frac{\phi(m_1, x_1)\,\phi(m_2, x_2)}{w_{m_1,j}(x_1)\, w_{m_2,j}(x_2)}\right] \phi(m_1, x_1)\,\phi(m_2, x_2)\, d\mu(x_1)\, d\mu(x_2)$$

for $1 \le m_1 \le M$ and $1 \le m_2 \le M$,

where

(i) $0 \le \lambda_j \le 1$ for $1 \le j \le J$ and $\Sigma_{j=1}^{J} \lambda_j = 1$ and

(ii) for each m, $\phi(m, \cdot)$ is a density with respect to the measure μ.

Then let us prove the following theorem.

THEOREM B-2. For $1 \le m_1 \le M$ and $1 \le m_2 \le M_1$,

$$\bar{D}(m_1, m_2, \phi, \lambda_1, \ldots, \lambda_J) = D(m_1, \phi, \lambda_1, \ldots, \phi_J) + D(m_2, \phi, \lambda_1, \ldots, \phi_J).$$

Proof. We have

$$\bar{D}(m_1, m_2, \phi, \lambda_1, \ldots, \lambda_J)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{j=1}^{J} \lambda_j \log\left[\frac{\phi(m_1, x_1)\,\phi(m_2, x_2)}{w_{m_1,j}(x_1)\, w_{m_2,j}(x_2)}\right] \phi(m_1, x_1)\,\phi(m_2, x_2)\, d\mu(x_1)\, d\mu(x_2)$$

$$= \int_{-\infty}^{\infty} \sum_{j=1}^{J} \lambda_j \log\left[\frac{\phi(m_1, x_1)}{w_{m_1,j}(x_1)}\right] \phi(m_1, x_1)\, d\mu(x_1)$$

$$+ \int_{-\infty}^{\infty} \sum_{j=1}^{J} \lambda_j \log\left[\frac{\phi(m_2, x_2)}{w_{m_2,j}(x_2)}\right] \phi(m_2, x_2)\, d\mu(x_2)$$

$$= D(m_1, \phi, \lambda_1, \ldots, \lambda_J) + D(m_2, \phi, \lambda_1, \ldots, \lambda_J).$$

Thus, Theorems B-1 and B-2 show that D, as defined in (B-2), is non-negative and additive in the sense discussed above. Both of these results are, of course, consequences of the nature of the log function used in the definitions of D and $\bar{D}$. Special cases of Theorems B-1 and B-2 are presented in a variety of sources such as Kullback reference [k], Lindley reference [f], DeGroot reference [y], and Box and Hill reference [z].

Formulation as a Two-Person, Zero-Sum Game

The concepts of game theory were first introduced into the study of sequential experimental design by Chernoff in reference [x]. Here we generalize those notions to a wider class of games between the experimenter and Nature to include the selection procedure described by Lindley in reference [f].

Game formulation. At each stage of the experiment, we assume that the experimenter is interested in maximizing a quantity of the form (B-2). The number of points in the parameter space, alternatively referred to as states of Nature, as well as the number of experiments available to the experimenter, is finite, and we formulate this problem as a two-person, zero-sum game as follows.

Let us assume that before a given trial in an experiment, the experimenter must decide which of M random variables $\{\tilde{Y}_m : 1 \leq m \leq M\}$ he will sample. He also believes that depending upon the true state of Nature the payoff to him will vary according to his choice of random variable. Now let us say that the experimenter has decided that for $1 \leq j \leq J$ and $1 \leq m \leq M$, if the true state of Nature is "j" and he chooses random variable $\tilde{Y}_m$, the payoff to him will be $\rho(j, m)$ given by

$$\rho(j, m) \qquad \int_{-\infty}^{\infty} \log\left[\frac{\phi(m, x)}{w_{m,j}(x)}\right] \phi(m, x)\, d\mu(x),$$

where

(i) $w_{m,j}(\cdot)$ is the density with respect to μ of the random variable $\tilde{Y}_m$, given that j is the true state of Nature, and

(ii) for each $m \in \{1, \ldots, M\}$, $\phi(m, \cdot)$ is a density with respect to the measure μ.

However, let us also assume that, in addition to choice of a particular state of Nature j and a particular random variable $\tilde{Y}_m$, called pure strategies for each player, the players may choose mixed strategies. A mixed strategy for Nature is a probability function λ defined on $\{1, \ldots, J\}$ and denoted by $\{\lambda_j : 1 \leq j \leq J\}$ where λ_j represents the probability with which Nature chooses parameter j at the step in question for $1 \leq j \leq J$. A mixed strategy for the experimenter is a probability function γ defined on $\{1, \ldots, M\}$ and denoted by $\{\gamma_m : 1 \leq m \leq M\}$, where γ_m represents the probability with which the experimenter chooses the random variable $\tilde{Y}_m$ for $1 \leq m \leq M$. However, in some cases, the rules of the game may specify that either Nature or the experimenter or both may choose only from among pure strategies.

Thus, if Nature chooses mixed strategy λ, the experimenter would obviously like to choose the random variable $\tilde{Y}_m$ to maximize his expected payoff. If Nature chooses mixed strategy λ, then the expected payoff to the experimenter for his choice of random variable $\tilde{Y}_m$ is given by

$$\Gamma(m) = \sum_{j=1}^{J} \lambda_j \, \rho(j, m)$$

$$= D(m, \phi, \lambda_1, \ldots, \lambda_J) \qquad \text{for } 1 \leq m \leq M.$$

Thus, if the experimenter may assume that he knows Nature's strategy at any step, his best strategy is to choose the random variable $\tilde{Y}_{m^*}$, where m^* maximizes Γ over the set $\{1, \ldots, M\}$. In this case, the maximum payoff to the experimenter called the value of this game is given by

$$v = \max_{m \in \{1, \ldots, M\}} \Gamma(m)$$

$$= \max_{m \in \{1, \ldots, M\}} D(m, \phi(m, \cdot), \lambda_1, \ldots, \lambda_J).$$

If, on the other hand, we may assume that Nature chooses a strategy from the class of mixed strategies G in such a way as to minimize the maximum payoff, then the value of the game to the experimenter is given by

$$v^* = \max_{m \in \{1, \ldots, m\}} \; \min_{(\lambda_1, \ldots, \lambda_J) \in G} \; \sum_{j=1}^{J} \lambda_j \, \rho(j, m)$$

$$= \max_{m \in \{1, \ldots, m\}} \; \min_{(\lambda_1, \ldots, \lambda_J) \in G} D(m, \phi(m, \cdot), \lambda_1, \ldots, \lambda_J).$$

<u>Chernoff's procedure A</u>. As above, let i^*_{n-1} be the mode of the posterior distribution of the parameters at step n-1. Then in reference [x] Chernoff has suggested that the experimenter choose the random variable $\tilde{Y}_m$ at the n^{th} step to maximize the function Γ_1^n defined on the set $\{1, \ldots, M\}$ as follows:

$$\Gamma_1^n(m) = \inf_{(\lambda_1, \ldots, \lambda_J) \in \Lambda(i_{n-1})} \sum_{j=1}^{J} \lambda_j I(i_{n-1}^*, j, m) \qquad \text{for } 1 \le m \le M,$$

where

(i) the function I is defined in (B-1) and

(ii) $\Lambda(\cdot)$ is defined in expression (B-5) above.

Chernoff has shown that this is equivalent to a choice on the part of the experimenter of a pure strategy to maximize his payoff in a game with payoff function ρ_1 given by

$$\rho_1(j, m) = \int_{-\infty}^{\infty} \log \left[\frac{w_{m, i_{n-1}^*}(x)}{w_{m, j}(x)} \right] w_{m, i_{n-1}^*}(x)\, d\mu(x) \qquad \text{for } 1 \le j \le J \text{ and } 1 \le m \le M,$$

where it is assumed that Nature is free to choose a mixed strategy from among all mixed strategies giving zero weight to the mode of the posterior distribution of the parameters at step n-1, and the experimenter is free to choose his strategy only from among his pure strategies. Here the value of the game at step n is given by

$$v_1 = \max_{m \in \{1, \ldots, m\}} \; \inf_{(\lambda_1, \ldots, \lambda_J) \in \Lambda(i_{n-1}^*)} \sum_{j=1}^{J} \lambda_j I(i_{n-1}^*, j, m).$$

In terms of the discriminator function D defined in expression (B-2), Chernoff states that at the n^{th} step the experimenter should sample the random variable $\tilde{Y}_m$ to maximize the function $D(\cdot, \eta(i_{n-1}^*, \cdot\ \cdot), \bar{\omega}(1), \ldots, \bar{\omega}(j))$ over the set $\{1, \ldots, M\}$. Thus, the value of the game may also be written as

$$v_1 = \max_{m \in \{1, \ldots, m\}} D(m, \eta(i_{n-1}^*, \cdot, \cdot), \bar{\omega}(1), \ldots, \bar{\omega}(J)).$$

As has been stated previously, if it is assumed that Nature may choose only from among her pure strategies, then the value of the game is given by

$$v_2 = \max_{m \in \{1, \ldots, m\}} D(m, \eta(i_{n-1}^*, \cdot, \cdot), \omega(i_{n-1}^{**}, 1), \ldots, \omega(i_{n-1}^{**}, J)).$$

It is a simple matter to show that

$$v_2 \geq v_1 .$$

Under mild restrictions, Chernoff has shown the following in reference [x]. (The stopping rule of procedure A is not directly relevant to our discussion and need not be defined here.)

LEMMA 1*. Let the stopping rule for procedure A be disregarded. Let τ be the smallest integer such that $i_n^* = T$ for $n \geq \tau$. Then there exist $b_1 > 0$ and $b_2 > 0$ such that

$$\Pr\{\tau \geq n\} \leq b_1 e^{-b_2 n} \quad \text{for } n \geq \tau .$$

While in reference [x] Chernoff has proved that procedure A has certain desirable asymptotic properties, he has also pointed out that procedure A may lead to "initial bungling," since "At first it is desirable to apply experiments which are informative for a broad range of parameter values. Maximizing the Kullback-Liebler information number may give experiments which are efficient only when θ is close to the estimated value."

<u>Lindley's procedure.</u> Lindley, reference [f], has suggested an alternative approach to the sequential-experimental design problem, which, as he points out, applies Shannon's definition of the information content of a probability distribution to the discussion of the notion of information in an experiment. This is closely related to the approach taken in formulating surveillance Policy II in Chapter III as we shall see below.

Lindley defines the amount of information provided by an experiment as the expected change in the entropy of the posterior probabilities of the parameters as a result of performing the experiment.

For example, if α^{n-1} is the posterior distribution of the parameters $\{1, \ldots, J\}$ at the $(n-1)^{st}$ step, then for $1 \leq m \leq M$ the Shannon information content in the selection of random variable $\tilde{Y}_m$ at the n^{th} step is $s(m)$ given as follows:

---

* Implicit in the proof of this lemma presented in reference [x] is the fact that for each $m \in \{1, \ldots, m\}$ and for any pair of distinct parameter values i and j, there exists $\beta \in \mathcal{X}$ such that $\mu(\beta) \neq 0$ and $\int_\beta w_{m,i}(x)\, d\mu x \neq \int_\beta w_{m,j}(x)\, d\mu(x)$.

$$s(m) = \int_{-\infty}^{\infty} \left[\sum_{j=1}^{J} f_{m,j}^{n}(x) \log\left[f_{m,j}^{n}(x)\right] - \sum_{j=1}^{J} \alpha^{n-1}(j) \log\left[\alpha^{n-1}(j)\right] \right] \left[\sum_{j=1}^{J} \alpha^{n-1}(j)\, w_{m,j}(x) \right] d\mu(x), \qquad \text{(B-7)}$$

where

$$f_{m,j}^{n}(x) = \frac{w_{m,j}(x)\, \alpha^{n-1}(j)}{\sum_{i=1}^{J} w_{m,i}(x)\, \alpha^{n-1}(i)} \qquad \text{for } -\infty < x < \infty,\ 1 \leq j \leq J.$$

Using Chernoff's results in reference [aa], if E is defined by

$$E(m, x) = \sum_{j=1}^{J} \alpha^{n-1}(j)\, w_{m,j}(x) \qquad \text{for } 1 \leq m \leq M \text{ and } -\infty < x < \infty,$$

then it is easily shown that

$$s(m) = D(m, E(m, \cdot), \alpha^{n-1}(1), \ldots, \alpha^{n-1}(J)), \qquad \text{for } 1 \leq m \leq M.$$

Lindley in reference [f] suggests that the experimenter choose the random variable $\tilde{Y}_m$ at the n^{th} step to maximize the function s defined above.

The following discussion relates Lindley's procedure to the maximum information-gain policy formulated in Chapter III.

In the context of Chapter III, α^{n-1} indicates the current target location probability distribution P_B and $\tilde{Y}_m$ indicates the outcome obtained if the m^{th} cell is searched. We define $\tilde{Y}_m$ so that $\tilde{Y}_m = 1$ if a response is obtained and $\tilde{Y}_m = 0$ if no response is obtained. The measure μ in equation (B-7) assigns weight 1 to each of the sets $\{0\}$ and $\{1\}$ and $\mu(A - \{0, 1\}) = 0$ for any measurable set A.

The quantity $w_{m,j}(r)$ is the probability of obtaining an outcome r given that cell m is searched and the target is in cell j. Here, $r = 1$ indicates a response and $r = 0$ indicates no response. In the notation of Chapter III,

$$w_{m,j}(r) = Q(r, j, m).$$

The probability $f^n_{m,j}(r)$ is the probability that the target is in cell j given that cell m is searched and response r is obtained. In the notation of Chapter III,

$$f^n_{m,j}(r) = p_A(r, j, m).$$

If entropy of a discrete distribution P on J cells is denoted H[P], i.e.,

$$H[P] = - \sum_{j=1}^{J} P(j) \ln P(j),$$

and if as in Chapter III

$$U(m) = \sum_{j=1}^{J} \sum_{r=0}^{1} P_B(j)\, Q(r, j, k)\, H[P_A(r, \cdot, k)],$$

then

$$\begin{aligned} s(m) &= \sum_{r=0}^{1} - \left[H[p_A(r, \cdot, m)]\right] \left[\sum_{j=1}^{J} P_B(j)\, Q(r, j, m)\right] \\ &\quad + \sum_{r=0}^{1} \left[H[P_B]\right] \left[\sum_{j=1}^{J} P_B(j)\, Q(r, j, m)\right] \\ &= -U(m) + H[P_B]. \end{aligned} \qquad \text{(B-8)}$$

Equation (B-8) indicates that finding the m which maximizes s(m) (the Lindley approach) is equivalent to finding the m which minimizes U(m) (Chapter III approach).

Like the Chernoff procedure A, Lindley's procedure for choosing a random variable to sample at the n^{th} step may also be considered within the context of game theory. Once again, we think of Nature and the experimenter as playing a game with a particular payoff function. In this case the payoff function is slightly different that the one assumed by Chernoff in his procedure A. For a choice by Nature of the parameter j and a choice by the experimenter of the random variable $\tilde{Y}_m$, Lindley assumes that the payoff to the experimenter at the n^{th} step is given by

$$\rho_2(j, m) = \int_{-\infty}^{\infty} \log \left[\frac{\sum_{i=1}^{J} w_{m,i}(x)\, \alpha^{n-1}(i)}{w_{m,j}(x)} \right] \left[\sum_{i=1}^{J} w_{m,i}(x)\, \alpha^{n-1}(i) \right] d\mu(x)$$

$$\text{for } 1 \leq j \leq J \text{ and } 1 \leq m \leq M. \qquad \text{(B-9)}$$

Thus, the payoff assumed by Lindley in his game with Nature is significantly different from that assumed by Chernoff. Also different is the strategy assumed for Nature. Lindley assumes that for her strategy at the n^{th} step, Nature chooses the mixed strategy α^{n-1}. That is, Nature chooses parameter j with probability $\alpha^{n-1}(j)$ at the n^{th} step.

Consequently, assuming that Nature plays mixed strategy α^{n-1} at step n, if the experimenter wishes to maximize his expected payoff, he must choose the random variable $\tilde{Y}_{m^*}$ such that m^* maximizes the function Γ_2^n defined as follows:

$$\Gamma_2^n(m) = \sum_{j=1}^{J} \alpha^{n-1}(j)\, \rho_2(j, m)$$

$$s(m) \quad D(m, E, \alpha^{n-1}(1), \ldots, \alpha^{n-1}(J)), \qquad \text{for } 1 \leq m \leq M.$$

Thus, the value of the game or the maximum information which the experimenter can derive from a sample at the n^{th} trial is given by

$$v_3 = \max_{m \in \{1, \ldots, m\}} D(m, E, \alpha^{n-1}(1), \ldots, \alpha^{n-1}(J)).$$

Large-sample results similar to the ones obtained for Chernoff's procedure A in reference [x] have not yet been obtained for Lindley's procedure. However, it is fairly obvious that since Lindley's procedure uses all the information about the parameters available to the experimenter at every stage, it will not be subject to "initial bungling" to the same extent as Chernoff's procedure A. However, conversely, due to its heavy reliance upon all information regarding the parameters at each step rather than only the most likely as in procedure A, its large-sample properties may not be as dramatic as those attending Chernoff's procedure.