AD-777 878

AN INTRODUCTION TO SEARCH THEORY

Mario Jorge Ferreira Braga

Naval Postgraduate School
Monterey, California

March 1974

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER AD 777878
4. TITLE (and Subtitle) An Introduction to Search Theory		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis; March 1974
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Mario Jorge Ferreira Braga		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE March 1974
		13. NUMBER OF PAGES 117
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Naval Postgraduate School Monterey, California 93940		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

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)

Search theory
Local detection function
Sweep Width

Uniformly optimal search plan
Search grid

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

This paper intends to be an introduction to search theory, and it is written primarily to assist staff officers in their duties.

For those who are in charge of planning and deciding about search, it is generally felt that is necessary to know what is behind the standard procedures found in tactical publications.

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE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

I

However, they are seldom familiar with the advanced mathematics required to read the sophisticated papers on the subject. This text is an attempt to conciliate those officers needs, with the reasonable amount of precision, which is required for planning, analyzing and evaluating search procedures.

An Introduction To Search Theory

by

Mario Jorge Ferreira Braga
Commander, Brazilian Navy
B.S., Brazilian Naval Academy, 1956

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 1974

Author _____

Approved by: _____
                                          Thesis Advisor

_____
                                          Second Reader

_____
Chairman, Department of Operations Research
       and Administrative Sciences

_____
                                          Academic Dean

1

ABSTRACT

This paper intends to be an introduction to search theory, and it is written primarily to assist staff officers in their duties.

For those who are in charge of planning and deciding about search, it is generally felt that is is necessary to know what is behind the standard procedures found in tactical publications. However, they are seldom familiar with the advanced mathematics required to read the sophisticated papers on the subject. This text is an attempt to conciliate those officers needs, with the reasonable amount of precision, which is required for planning, analyzing and evaluating search procedures.

TABLE OF CONTENTS

I. SETTING THE PROBLEM

A. SEARCH AS AN OPTIMIZATION PROBLEM

Almost all relevant search situations are characterized
by three basic features:

(a) uncertainty about the target's position.

(b) limitation and uncertainty about the capability of
detection.

(c) scarcity of search resources.

In (a) what is meant is that target's position is described
by a probability distribution inside a "search area" (denoted
"S.A." from now on).

In (b) the point is that all sensors are limited in their
ranges and even inside the proper range there still exists a
positive probability of missing the target, due to factors
as operator's error, propagation conditions, and so on.

Finally in (c), the problem is that the resource's amount
which is available to perform a search is a finite one. This
fact introduces a "budget constraint" in the problem.

Two basic measures of effectiveness (designated as "M.O.E.'s"
from no on) can be used in search problems. The first is the
probability of finding the target, in a given time, by using
the available amount of resources. The second, is the expected
time to find the target. As will be seen, most of the pro-
cedures which maximize the former, minimize the latter. For
the time being, this paper will deal only with the first.

At this point, an example seems appropriate in order to introduce some important concepts, and to illustrate those already mentioned. Consider the S.A. in Figure 1.

It is obvious that for the target there are only two possibilities: Either it is located in cell I, or in cell II. Defining the events E_1= {Target is in cell I} and E_2= {Target is in cell II}, the target's location distribution is then:

$$P(E_1) = P_I = 0.8$$

$$P(E_2) = P_{II} = 0.2$$

Where $\sum_{j=I}^{II} P_j = P_I + P_{II} = 1$

CELL I	CELL II
$P_I = 0.8$	$P_{II} = 0.2$

Figure I-1

Suppose that the sensor which is in use works by discrete glimpses, and define the events D_1= {Target is detected in cell I} and D_2= {Target is detected in cell II}. Let

$$P(D_1/E_1) = B_1$$

$$P(D_2/E_2) = B_2$$

These two probabilities are assumed known and usually equal, and characterize the efficiency of the sensor. Finally, in order to take in account the scarcity of searching resources,

5

and

$$P(D_1) = P(D_1/E_1) \, P(E_1)$$

And for cell II, by identical reasoning

$$P(D_2) = P(D_2/E_2) \, P(E_2)$$

Let D= {target is detected}. Obviously, D is the union of
D_1 and D_2, and further more it is an union of two disjoint
events, for, if the target is detected, it occurs either in
cell I, or in cell II, but by no means in both. Therefore,
from the axioms of probability theory:

$$P(D_1 \cup D_2) = P(D) = P(D_1) + P(D_2)$$

or

$$P(D) = P(D_1/E_1) \, P(E_1) + P(D_2/E_2) \, P(E_2)$$

In order to gain more insight in the problem, suppose the
following numbers are assigned:

$$P_I = 0.8$$
$$P_{II} = 0.2$$
$$B_1 = B_2 = 0.6$$
$$G = 2 \text{ Glimpses}$$

Now, the problem is to know where to spend the available
glimpses in order to have the larger possible value for P(D),
the probability of detection. The place for the first glimpse
is obvious. Cell I having the largest probability is the
natural candidate. The general rule is, evidently, look for

6

suppose that only G glimpses are permitted. Recalling from probability theory, the definition of conditional probability, it can be stated that:

$$P(D_1/E_1) = \frac{P(D_1 \cap E_1)}{P(E_1)}$$

Where $D_1 r E_1$ stands for the intersection of the events D_1 and E_1. But if this intersection is carefully examined, it is not difficult to see that it is equivalent to the event D_1. In order to prove it, notice that D_1 is included in E_1, in the sense that, if detection occurs in cell I, evidently the target is there. Then D_1 implies E_1. On the other hand, it is possible that the detection fails to find the target in Cell I, and nevertheless it is there. Then E_1 does not imply D_1. Figure 2 shows a Venn diagram which illustrates this fact.
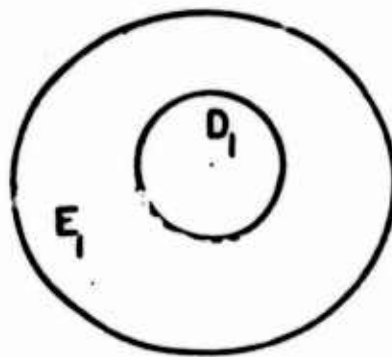

Figure I-2

Due to the above considerations, (*) becomes:

$$P(D_1/E_1) = \frac{P(D_1)}{P(E_1)}$$

7

the target at the place where it is most likely to be. The best place for the second glimpse, however, is not so obvious. There are two possibilities: Cell I again, or cell II. Therefore, by computing $P(D)$ under these two possible courses of action, it is possible to know which is better. Observe, however, after the sensor had looked in cell I, then P_I and P_{II} no longer are equal to $P(E_1)$ and $P(E_2)$. For, if cell I was already scanned and the target was not found there, it is less likely for this cell to contain it, and more likely for cell II. In the extreme case, when using a perfect sensor ($B_1 = B_2 = 1$), if the target is not found in cell I, then it is not there, and thus $P(E_1) = 0$ and $P(E_2) = 1$, whatever values they had before the glimpse. Therefore, after the first glimpse, it is necessary to update $P(E_1)$ and $P(E_2)$ in order to compute $P(D)$. Arguing in the same way as before, it can be stated:

$$P(E_1/\bar{D}) = \frac{P(E_1 n \bar{D})}{P(\bar{D})}$$

Where $\bar{D}$ is the complement of D. But, $P(E_1 n \bar{D})$ the probability of no detection even though the target is for sure in cell I, is equal to:

$$P(E_1 n \bar{D}) = P(E_1 n \bar{D}_1)$$

Because the only way for the event $E_1 n \bar{D}$ to occur is that no detection happens in cell I.

8

Also, the event $\bar{D}$ is equivalent to the event $(\bar{D}_1 n E_1) \cup (\bar{D}_1 n E_2)$, because if no detection occurs, either it does not occur in cell I with the target there $(\bar{D}_1 n E_1)$, or in cell II with the target there $(\bar{D}_2 n E_2)$. These events are of course, disjoint. After these considerations it is possible to write:

$$P(E_1/\bar{D}) = \frac{P(E_1 n \bar{D}_1)}{P(E_1 n \bar{D}_1) + P(E_2 n \bar{D}_2)}$$

Further, it is true that

$$P(E_1 n \bar{D}_1) = P(E_1)P(\bar{D}_1/E_1)$$

And

$$P(E_2 n \bar{D}_2) = P(E_2)P(\bar{D}_2/E_2)$$

Therefore

$$P(E_1/\bar{D}) = \frac{P(E_1)P(\bar{D}_1/E_1)}{P(E_1)P(\bar{D}_1/E_1) + P(E_2)P(\bar{D}_2/E_2)}$$

Finally, from the very definition of the event D_1, it follows that:

$$P(\bar{D}_1/E_1) = 1 - P(D_1/E_1)$$

And similarly

$$P(\bar{D}_2/E_2) = 1 - P(D_2/E_2)$$

9

Hence

$$P(E_1/\bar{D}) = \frac{P(E_1)\ [1-P(D_1/E_1)]}{P(E_1)\ [1-P(D_1/E_1)] + P(E_2)\ [1-P(D_2/E_2)]}$$

and, by the same reasons:

$$P(E_2/\bar{D}) = \frac{P(E_2)\ [1-P(D_2/E_2)]}{P(E_1)\ [1-P(D_1/E_1)] + P(E_2)\ [1-P(D_2/E_2)]}$$

The above results are a particular case of the well known theorem of Bayes, from probability theory. It is interesting to notice that the mentioned theorem can be generalized to any number of cells, and even to continuous distribution. The probabilities $P(E_1/\bar{D})$ and $P(E_2/\bar{D})$ are called posterior probabilities, in opposition to P_I and P_{II} which are called prior probabilities. In section I-2, this subject will be treated in a more detailed way.

Now, there is a way to know where the second glimpse must be placed. For, by using the numbers of the problem:

$$P(E_1/\bar{D}) = \frac{(0.8)\ [1-0.6]}{(0.8)\ [1-0.6] + (0.2)\ [1-0]} = 0.616$$

And

$$P(E_2/\bar{D}) = \frac{0.2}{(0.8)\ [1-0.6] + (0.2)\ [1-0]} = 0.384$$

Therefore, even after the first glimpse, the probability of cell I is still higher and the second must be placed there, in accordance to the general rule stated before.

The probability of detection, P, using this course of action, is now:

$$P = (0.616)(0.6) = 0.3696$$

Finally, to compute the probability of detection on either glimpse, $P(D)$, by using the first option, we can proceed as follows:

$$P(D) = P \text{ (Detection on 1}^{th} \text{ glimpse) +}$$
$$P \text{ (no detection on 1}^{th} \text{ glimpse and detection on the 2}^{th})$$

Or

$$P(D) = P_I B_1 + P(E_1/\bar{D}) [1-P_I B_1] B_1$$

$$= (0.8) [(0.8) + (0.616)(0.52)] \simeq 0.672$$

Until now, many important factors have been neglected for the sake of simplicity. For example, just two cells were used, no considerations were made about the size of the search area, the sensor was supposed to work by separate glimpses, and no attention was paid to the pattern of the track that the searching vehicle eventually describes inside of the search area. In order to improve the model, consider the following modifications in this example:

Suppose there is a target in the S.A. pictured in Figure 3, whose position is described by the following distribution: Probability of cell j contain the target $= P_j$, $j \epsilon$ {I, II, II, IV}

$$\sum_{j=I}^{IV} P_j = P_I + P_{II} + P_{III} + P_{IV} = 1$$

11

+----------------+----------------+
(w)	
CELL I	CELL II
+----------------+----------------+
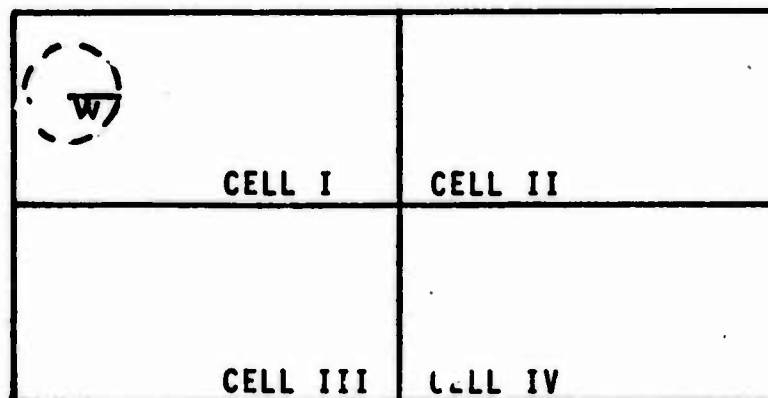| | |
CELL III	CELL IV
+----------------+----------------+

Figure I-3

Furthermore, suppose that the sensor's range is "W" and that within "W" it detects the target with probability 1. Let "v" be the speed of the platform where the sensor is, and suppose there is a fixed amount of time, say "T", for conducting the search. If the area that can be swept in the time "T" is smaller than "S.A.", then it is clear that the time must be shared in some way among the four cells. Of course, there are infinite ways of doing that, but the main interest is the one that maximize the probability of finding the target. Denoting by Z_j a measure of effort spent in cell "j", and by "a" the total amount of effort available, (in this case the fraction of the "S.A." that can be covered, given "w", "v", and "T"), it is possible to see that the "M.O.E." is a function of Z only. Now, the problem can be seen as:

$$\text{MAX } P \ (Z_I, \ Z_{II}, \ Z_{III}, \ Z_{IV})$$

$$\text{S.t} \hspace{4cm} (I-I)$$

$$\sum_{j=I}^{IV} Z_j \ A_j \leq a$$

Where $P(Z_I, \ Z_{II}, \ Z_{III}, \ Z_{IV}) = P(\bar{Z})$ is the probability of finding the target. Z_j is the DENSITY OF EFFORT in cell j, i.e., the ratio between the effort placed in cell j, and the area of cell j. If A_j is the area of cell j and t_j the time spent there, it can be written as:

$$Z_j = \frac{t_j}{A_j}$$

The reason the density of effort was chosen as measure, is that in some problems there are cells with different areas. Another point is that, although it has been said that the probability of finding the target is a function of the distribution of effort among the cells, and that the probability of detection is a function of the density of effort that has been placed there, it was not specified any analytical expression for it. Coming back to the example given, in order to develop an approach to this problem:

Let $B_j(Z_j)$ be the probability of finding the target in cell j, given that the target is there, and a density of effort Z_j was applied there. From elementary probability theory

$$P(\bar{Z}) = \sum_{j=I}^{IV} B_j(Z_j) \ P_j \hspace{3cm} (I-II)$$

13

Then, as P_j is already known, the problem now is to find B_j (Z_j). Incidentally, B_j (Z_j) which stands for the probability of finding the target, given it is there, and by the use of a density of effort Z_j is called a LOCAL DETECTION FUNCTION, or as it will be used now, L.D.F. Suppose now, that the search is in one of the cells and assuming that:

(a) If the target is in that cell, the probability distribution of its position is uniform.

(b) If passing by the target at a distance equal or smaller than "$\frac{W}{2}$" it will be detected with probability 1.

Under these circumstances, it seems reasonable to perform a systematic search in the cell. It can be done by using parallel and non overlapping sweeps, as shown in Figure I.[4]
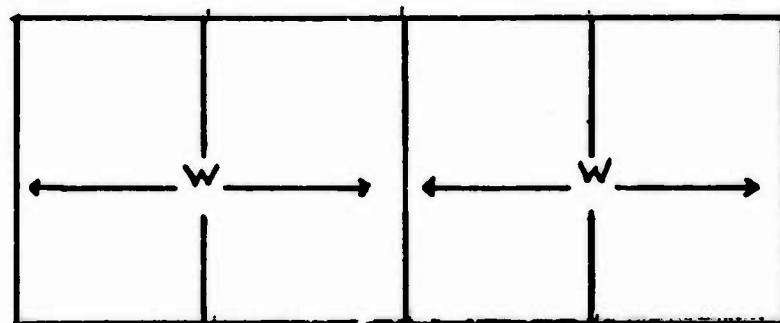

Figure I-4

Evidently, if A_j is the area of the cell, the probability of detection after the searching vehicle has traveled a track length "L" is $\frac{LW}{A_j}$, or using $\frac{L}{A_j}$ as the density effort Z_j. (We will now skip the subscripts on $B_j(Z_j)$):

$$B(Z) = \begin{cases} WZ & \text{if} \quad LW < A_j \\ 1 & \text{if} \quad LW \geq A_j \end{cases} \qquad \text{(I-III)}$$

This is a very appealing L.D.F., because of its simplicity.
However, it is not a simple matter to place the paths exactly
as in Figure 4.

Even if it is assumed that they are indeed parallel one
to another, errors in navigation will cause the paths to over-
lap in some places, and leave gaps in others. Thus, the
probability of finding the target will be less than (I-III).
Therefore, unless the navigation is very well made, the use
of formula (I-III) will overestimate the detection capability.

On the other hand, an entire asystematic procedure can be
used to place the path within the cell. Instead of parallel
paths, random ones. It can be formalized by adding one more
assumption to the other ones:

(c) The path is composed of segments with random lengths
and directions, mutually independent.
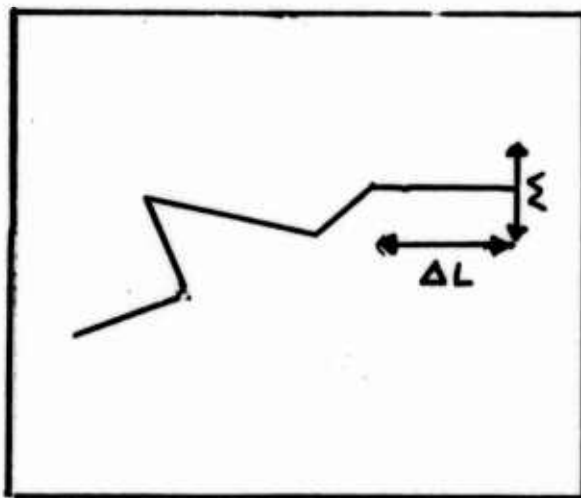The situation is pictured in Figure 5

Figure I-5

The probability of detection along segment "ΔL" is

$$f = \frac{W\Delta L}{A} \qquad (I-IV)$$

Where A is the area of the cell. Note that the ratio $\frac{\Delta L}{A}$ is the effort density "ΔZ" and therefore (I-IV) can be written as

$$f = W\Delta Z \qquad (I-V)$$

Let (B(Z) be the probability the target has been detected by the time the search vehicle has traveled a track length "L", and therefore placed a density of effort "Z" inside the cell. Then

$$B(Z + \Delta Z) = B(Z) + [1-B(Z)]W\,\Delta Z \qquad (I-VI)$$

This is because, either the target is detected during the length "L" (and B(Z) stands for this possibility), or the search fails to detect by the track length "L" (and succeeds in the next increment ΔL) and the probability of this event is [1-B(Z)]WΔZ (multiplication is valid here because of assumption (c)).

Now rearranging (I-VI)

$$B(Z + \Delta Z) - B(Z) = [1-B(Z)]\, W\Delta Z$$

or

$$\frac{B(Z + \Delta Z) - B(Z)}{\Delta Z} = [1-B(Z)]W$$

16

And in the limit

$$\frac{B(Z+\Delta Z) - B(Z)}{\Delta Z} = B(Z) = [1-B(Z)]W$$

Solving this differential equation

$$B(Z) = 1 - EXP[-WZ] \qquad\qquad (I-VII)$$

This is called RANDOM SEARCH FORMULA, a very important one,
first derived by B. O. Koopman in [1].

Plotting the two L.D.F's known already, it can be seen
that (I-VII) gives a lower bound of effectiveness of the search
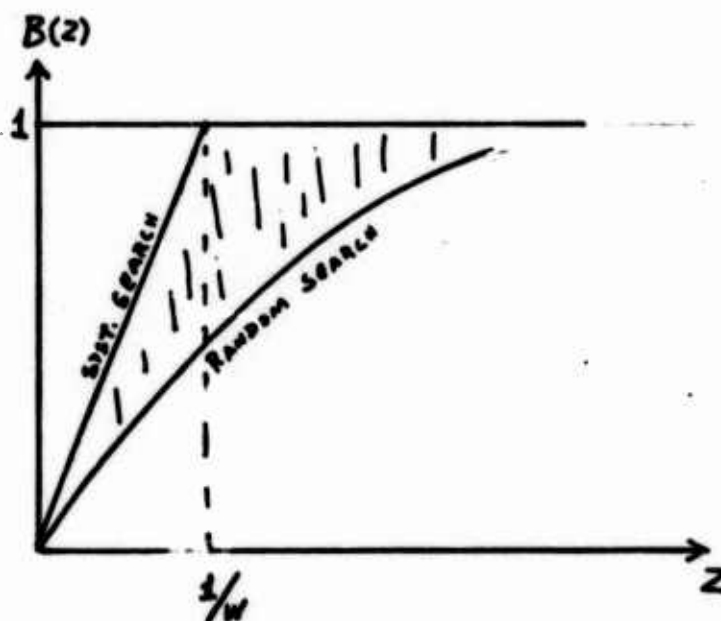inside the cell. The graph is in Figure I-6.

Figure I-6

The real L.D.F. is generally somewhere in the dashed region
between the two curves. In order to improve the insight about
how precisely either model describes the situation, assume
that the error in the paths placement in Figure IV is normally

17

distributed with mean zero and standard deviation σ_N, and gap effects mentioned before will take place, even though the paths remain parallel. It can be shown (see [1] and [2]), that, if the rectangle is large compared to W, the probability of detecting the target depends only on the ratio σ_N/W. Figure 7 which was taken from (2), shows this dependence.
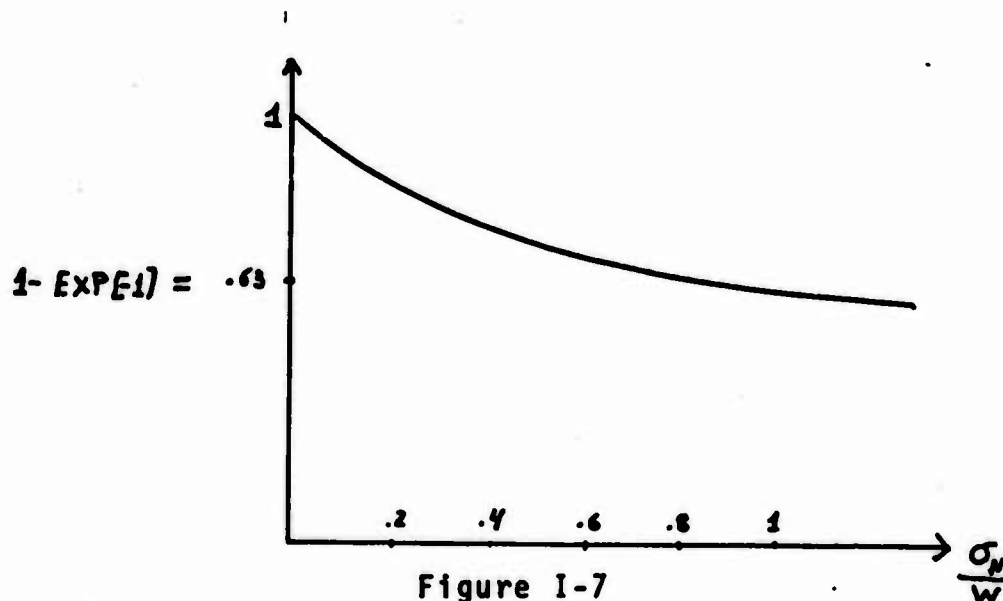
Figure I-7

Observe that, for small ratios $\frac{\sigma_N}{W}$, the probability approaches one and therefore the systematic search model applies, however, for the values of $\frac{\sigma_N}{W}$ about 0.2 the probability approaches very rapidly 1-EXP[-1] which is the probability that would be found if WL $= A_j$ is substituted in the formula (I-VIII).

For an aircraft, it is normally assumed (see [3]) σ_N equal to 20 miles and "W" at 5.000 feet of altitude, equal to 40 miles, which gives $\frac{\sigma_N}{W} = .5$ and the random search model applies.

18

The formula (I-VII), being a very conservative estimate of effectiveness, is far more useful than (I-III), and it will be used troughout this paper as the L.D.F. unless otherwise specified.

Now the maximization problem (I-I) can be written in a more precise way because, by substituting (I-VII) in (I-I) a closed expression for the probability of detecting the target is obtained.

$$P(Z) = \sum_{j=I}^{IV} P_j [1-EXP(-WZ_j)] \qquad (I-VIII)$$

and (I-1) becomes

$$MAX \sum_{j=I}^{IV} P_j [1-EXP(-WZ_j)$$

$$s.t. \qquad (I-IX)$$

$$\sum_{j=I}^{IV} A_j Z_j \leq a$$

To substitute some numbers in the example given, for sake of further classification, let: a) the probability distribution be:

$$P_I = P_{II} = P_{III} = P_{IV} = \frac{1}{4}$$

b) the searching vehicle speed by:

$$V = 150 \text{ knots}$$

c) the sensor's range be:

$$W = 15 \text{ miles}$$

19

d) the four cells have the same dimensions, say 30 X 50 miles, and f) suppose that the maximum time to be spent in the search is no more than two hours.

It seems reasonable to split evenly the effort among the cells and it is indifferent which one it is chosen to start. The density of effort in each cell is:

$$Z_j = \frac{L}{A_j} = \frac{Vt}{1500} = \frac{100 \times \frac{1}{2}}{1500} = \frac{1}{30}$$

Where $t = \frac{1}{2}$ was used, because there are only two hours to be spent in the search and $\frac{1}{2}$ hour in each cell. Observe that, due to the fact that, the probabilities and the areas are the same for all the cells and the effort is being split evenly among them, the density of effort is the same in the whole search area. So, it can be computed as:

$$Z_j = \frac{Vt}{A} = \frac{100X2}{7.000} = \frac{1}{30}$$

Where the total time (t=2 hours) was used and the total area (4 X 30 X 50 = 6000 sq. mi.). By using this density of effort:

$$P(\bar{Z}) = \frac{1}{4} \sum_{j=I}^{IV} [1-EXP(-\frac{15}{30})] = 1 - EXP[-\frac{1}{2}] = 0.394$$

Now, suppose that the probability distribution is given by

$$P_j = \begin{cases} 0.4 & ij \quad j = I \\ 0.3 & ij \quad j = II \\ 0.2 & ij \quad j = III \\ 0.1 & ij \quad j = IV \end{cases}$$

20

And recall that the maximum effort is already known. The available effort is $\frac{6000}{30}$ = 200. Then, (I-IX) becomes:

$$\text{MAX} \quad \sum_{j=I}^{IV} P_j \ [1\text{-EXP}(-WZ_j)$$

s.t.

$$\sum_{j=I}^{IV} A_j \ z_j \leq 200$$

The next point will be to consider several ways to allocating the effort or several functions Z_j, and see what happens to the probability of finding the target:

A. To divide evenly among the cells or

$$Z_j = \frac{1}{30} \ , \ j\epsilon \ \{I, \ II, \ III, \ IV\}$$

And

$$P(\bar{Z}) = \sum_{j=I}^{IV} P_j \ [1\text{-EXP}(1 \ \frac{15}{30})] =$$

$$0.\underline{4} \ [1\text{-EXP}(- \frac{15}{30})] + 0.3 \ [1\text{-EXP}(- \frac{15}{30})] +$$

$$0.2 \ [1\text{-EXP}(- \frac{15}{30})] + 0.1[1\text{-EXP}(- \frac{15}{30})] =$$

$$1 \ -\text{EXP}(- \frac{1}{2}) = 0.39\underline{4}$$

B. To spend all effort in cell I or

$$Z_j = \frac{Vt}{A_j} = \begin{cases} \frac{100 \ X \ 2}{1500} = \frac{2}{15} & \text{if} \quad j=I \\ \\ 0 & \text{otherwise} \end{cases}$$

21

and

$$P(\hat{Z}) = 0.\underline{4} \left[1 - EXP\left(-\frac{2 \times 15}{15}\right)\right] = 0.3\underline{4}6$$

C. To split evenly the effort between the cells I and II or:

$$Z_j = \frac{Vt}{A_j} = \begin{cases} \frac{100 \times 1}{1500} = \frac{1}{15} & \text{if } j\epsilon \{I, II\} \\ 0 & \text{otherwise} \end{cases}$$

and

$$P(\bar{Z}) = 0.\underline{4} \left[1 - EXP\left(-\frac{15}{15}\right)\right] + 0.3 \left[1 - EXP\left(-\frac{15}{15}\right)\right].$$

$$= 0.\underline{44}31$$

D. To divide the effort proportionally to the probability of each cell or

$$Z_j = \frac{Vt}{A_j} = \begin{cases} \frac{1.66 \times 48}{1500} = 0.0533 & \text{if } j = I \\ \frac{1.66 \times 36}{1500} = 0.0399 & \text{if } j = II \\ \frac{1.66 \times 24}{1500} = 0.0266 & \text{if } j = III \\ \frac{1.66 \times 12}{1500} = 0.0133 & \text{if } j = IV \end{cases}$$

Where 1.66 is the speed in miles per minute and 48, 36, 24, 12 the minutes spent in cells I to IV respectively. The probability is:

$$P(\bar{Z}) = (0.\underline{4}) \left[1 - EXP(-0.8)\right] + (0.3) \left[1 - EXP(-0.6)\right] +$$
$$(0.2) \left[1 - EXP(-0.4)\right] + (0.1) \left[1 - EXP(-0.2)\right] =$$
$$0.4402$$

22

E.   To divide evenly the effort among the cells I, II and III or

$$Z_j = \frac{Vt}{A_j} = \begin{cases} \frac{40 \times 1.66}{1500} = 0.0442 & \text{if } j \in \{I, II, III\} \\ 0 & \text{otherwise} \end{cases}$$

and by similar calculations

$$P(\bar{Z}) = 0.4356$$

F.   To divide the effort proportionally to probability, among cells I, II and III; do not place any effort in cell IV or

$$Z_j = \frac{Vt}{A_j} = \begin{cases} \frac{53 \times 1.66}{1500} = 0.058 & \text{if } j = I \\ \frac{40 \times 1.66}{1500} = 0.0442 & \text{if } j = II \\ \frac{27 \times 1.66}{1500} = 0.030 & \text{if } j = III \\ 0 & \text{if } j = IV. \end{cases}$$

and the probability is

$$P(\bar{z}) = 0.450$$

The table in Figure 8 summarizes the results reached till now, by the use of six courses of action.

23

ALLOCATION	PROBABILITY OF FINDING THE TARGET
A	$P(A) = 0.3940$
B	$P(B) = 0.3460$
C	$P(C) = 0.4431$
D	$P(D) = 0.4402$
E	$P(E) = 0.4356$
F	$P(F) = 0.4500$

Figure I-8

It can be seen that for different allocations, the probability varies in a wide range, (the range could be wider if some less obvious allocation was chosen, say, to spend all time in cell IV), and that common sense by itself is unable to provide any guidance, even in the simple example we are using. For, the best result achieved was by placing no effort in cell IV, which does not seem very wise at a first sight. Furthermore, there is not any procedure to know if we can do any better, than the one done with allocation F, because there are infinitely many possible allocations and we have only checked six.

In the next section, a specific procedure will be considered, which works very well in most cases.

Another point seen in the example, is that the same amount of effort (200) has been used in all six allocations.

24

It can be said that all of them "cost" the same, but "F"
gives a better "return" than the others. In fact, it can be
defined

$$\sum_{j=1}^{J} A_j \, Z_j = C \, (\bar{Z})$$

As the cost of allocation Z_j, and $C \, (\bar{Z})$ will be used to
denote the function and "C" to denote the values it takes.
Then, generalizing (I-1), the search problem can be set as:

$$\text{MAX } P(\bar{Z})$$
$$\text{s.t.}$$
$$C \, (\bar{Z}) \leq C$$

(I-X)

Or putting it in words, on allocation is desired, that is
a solution for the optimization under constraint problem
(I-X).

B. POSTERIOR TARGET's LOCATION DISTRIBUTION

Returning to the example given, it can be recalled that
P_j was defined as the probability that the target is located
in cell j. However, these probabilities are computed before
any search effort is placed in the S.A. Once, some effort is
placed, those values no longer hold, for after each increment
of effort placed in a cell without success, the probability
that the target is in the searched cell decreases, and, by
consequence, the other cells probabilities will be increased.

Then, $\tilde{P}_j$ can be said as the probability associated with cell j after some effort was placed in the S.A. "P_j" is called the prior probability and "$\tilde{P}_j$" the posterior probability, and the corresponding distributions prior and posterior target location distribution, respectively.

From probability theory, Bayes' theorem is recalled, (for demonstration see [4]), and by its use posterior probability is computed as follows: Suppose a search has been made using allocation and that we did not find the target. Then:

$$\tilde{P}_k = \frac{P_k \, [1-B(Z_k)]}{\sum\limits_{j=I}^{J} P_j \, [1-B(Z_j)]} \qquad\qquad (I-XI)$$

Where the numerator stands for the probability that the target is located in cell "K", and is not detected, and the denominator the probability of no detection in any cell.

For example, suppose a search was made according to allocation "B" and failed to find the target after the two hours established. The posterior distribution is computed:

$$\tilde{P}_k = \frac{P_k \, [EXP(-Z_k W)]}{\sum\limits_{j=I}^{IV} P_j \, [EXP(-Z_j W)]}$$

Then:

$$\tilde{P}_I = \frac{(0.4) \, [EXP(-\frac{2 \times 15}{15})]}{(0.4) \, [EXP(-\frac{2 \times 15}{15})] + 0.6} = 0.08$$

26

$$\tilde{P}_{II} = \frac{0.3}{(0.\underline{4})\,[EXP(-\frac{2X15}{15})] + 0.6} = 0.46$$

$$\tilde{P}_{III} = \frac{0.2}{(0.\underline{4})\,[EXP(-\frac{2X15}{15})] + 0.6} = 0.310$$

$$\tilde{P}_{IV} = \frac{0.1}{(0.4)\,[EXP(-\frac{2X15}{15}) + 0.6} = 0.153$$

The table below displays the effect on the target position distribution in two hours of search in cell I

BEFORE	AFTER
$P_I = 0.\underline{4}$	$\tilde{P}_I = 0.080$
$P_{II} = 0.3$	$\tilde{P}_{II} = 0.\underline{4}60$
$P_{III} = 0.2$	$\tilde{P}_{III} = 0.310$
$P_{IV} = 0.1$	$\tilde{P}_{IV} = 0.150$

Almost all the probability mass which was in cell I in the very beginning, was transfered to the other cells. There-fore, if there is more time left to continue the search, it is reasonable to transfer the effort from cell I, because it is almost impossible the target is located there. Indeed, a very intuitive way of allocated effort is to place it where the probability of success is higher. If the search starts in the cell with highest prior probability, and the criterion of placing the next increment of effort in the cell with highest posterior probability is used, an allocation is defined. As will be seen later, this allocation is in most cases, (including

the one dealt with), a solution to the maximization problem (I-IX). Applying this procedure to the example given: The search starts in cell "I" and stay there until another cell shows a higher posterior probability than cell "I." Due to the equality of the cells areas it is obvious that the next step will be cell "II." To compute for how long the search will continue in cell "I" only, consider that after placing a density of effort ε_j in cell "I," the posterior probability is:

$$\tilde{P}_I = \frac{0.\underline{4} \, [EXP(-15\varepsilon_1)]}{(0.\underline{4}) \, [EXP(-15\varepsilon_1)] + 0.3 + 0.2 + 0.1}$$

And cell II

$$\tilde{P}_{II} = \frac{0.3}{(0.\underline{4}) \, [EXP(-15\varepsilon_1)] + 0.3 + 0.2 + 0.1}$$

We want ε_1 such that

$$\tilde{P}_I = \tilde{P}_{II}$$

Therefore

$$0.\underline{4} \, EXP \, [-15\varepsilon_1] = 0.3$$

or

$$\varepsilon_1 = - \frac{\ln \left(\frac{0.3}{0.\underline{4}}\right)}{15} = 0.01918$$

From the definition of effort density we know:

$$z_j = \frac{Vt}{A_j}$$

28

Then,

$$t = \frac{A_j Z_j}{V}$$

In this case

$$\theta_1 = \frac{A_1 \epsilon_1}{V} = \frac{30 \times 50 \times 0.01918}{100} = 0.28770 \text{ hours}$$

The posterior distribution, after ϵ_1:

$$\tilde{P}_I = \frac{0.\underline{4} \text{ EXP}(-15\epsilon_1)}{(0.\underline{4}) [\text{EXP}(-15\epsilon_1)] + 0.6} = \frac{3}{9}$$

$$\tilde{P}_{II} = \frac{0.3}{0.\underline{4} \text{ EXP}(-15\epsilon_1) + 0.6} = \frac{3}{9}$$

$$\tilde{P}_{III} = \frac{0.2}{0.\underline{4} \text{ EXP}(-15\epsilon_1) + 0.6} = \frac{2}{9}$$

$$\tilde{P}_{IV} = \frac{0.1}{0.\underline{4} \text{ EXP}(-15\epsilon_1) + 0.6} = \frac{1}{9}$$

Now, in accordance to the procedure chosen, the effort must be split evenly between cells I and II until another cell (in this example cell III) shows a higher posterior probability. Then, for computing ϵ_2:

$$\tilde{\tilde{P}}_I = \tilde{\tilde{P}}_{III} = \frac{\frac{0.3}{0.9} \text{ EXP } [-15\epsilon_2]}{2(\frac{0.3}{0.9}) \text{ EXP } [-15\epsilon_2] + \frac{2}{9} + \frac{1}{9}}$$

Where $\tilde{\tilde{P}}_j$ stands for the posterior probability of cell "j" after ϵ_2 has been placed. Equating $\tilde{\tilde{P}}_I$ or $\tilde{\tilde{P}}_{II}$ to $\tilde{\tilde{P}}_{III}$,

$$(\frac{0.3}{0.9}) \text{ EXP } [-15\epsilon_2] = \frac{2}{9}$$

because, as al

because, as always, the denominators are identical. Then,

$$\epsilon_2 = - \frac{\ln \left(\frac{2}{3}\right)}{15} = 0.027$$

And the time increment

$$\Theta_2 = \frac{(A_1 + A_2) \epsilon_2}{V} = \frac{2 \times 30 \times 50 \times 0.027}{100} = 0.81 \text{ hours}$$

To be spent half in cell I and half in cell II. The new posterior distribution after the density ϵ_2 is:

$$\tilde{\tilde{P}}_I = \tilde{\tilde{P}}_{II} = \tilde{\tilde{P}}_{III} = \frac{\left(\frac{0.3}{0.9}\right) \text{ EXP } [-15\epsilon_2]}{2 \left(\frac{0.3}{0.9}\right) \text{ EXP } [-15\epsilon_2] + \frac{2}{9} + \frac{1}{9}} = \frac{2}{7}$$

$$\tilde{\tilde{P}}_{IV} = \frac{\frac{1}{9}}{2\left(\frac{0.3}{0.9}\right) \text{ EXP } [-15\epsilon_2] + \frac{2}{9} + \frac{1}{9}} = \frac{1}{7}$$

In the same way we compute ϵ_3, which is the effort density to be placed in cells I, II and III before moving to cell IV.

$$\tilde{\tilde{\tilde{P}}}_I = \tilde{\tilde{\tilde{P}}}_{II} = \tilde{\tilde{\tilde{P}}}_{III} = \frac{\left(\frac{2}{7}\right) \text{ EXP } [-15\epsilon_3]}{3(\frac{2}{7} \text{ EXP } [-15\epsilon_3] + \frac{1}{7}}$$

$$\left(\frac{2}{7}\right) \text{ EXP } [-15\epsilon_3] = \frac{1}{7}$$

and

$$\epsilon_3 = - \frac{\ln \left(\frac{1}{2}\right)}{15} = 0.04621$$

Observe that, in the example given, there is a limit of two hours to perform the search. Then, all θ_1 and θ_2 can. be used, but only a fraction of θ_3, that will be called θ_3'. It is obvious that:

$$\theta_3' = 2 - (\theta_1 + \theta_2) = 2 - (0.28770 - 0.81000)$$

$$= 0.90230 \text{ hours}$$

So, the time to be spent in each cell will be

$$t_1 = \theta_1 + \frac{\theta_2}{2} + \frac{\theta_3'}{3} = 0.99347 \text{ hours}$$

$$t_2 = \frac{\theta_2}{2} + \frac{\theta_3'}{3} = 0.70576$$

$$t_3 = \frac{\theta_3'}{3} = 0.30076$$

$$t_4 = 0$$

Where "t_j" stand for the total time in cell "j". Observe also that $t_j = Z_j$ in the example given, due to the numbers chosen. Hence the probability of finding the target can be computed as

$$P(\bar{Z}) = (0.4) \ [1-EXP(-t_1)] + (0.3) \ [1-EXP(-t_2)] +$$

$$(0.2) \ [1-EXP(-t_3)] + (0.1) \ [1-EXP(-t_4)]$$

$$= 0.4553$$

31 .

Which is better than any other of six allocation tried in the last section. In fact, as said before, it is the best that can be done under these circumstances. Finally, it is interesting to observe that using the procedure above, if for any reason there is a need to stop before all the initial effort available is gone, the M.O.E. has still been maximized, under the new and smaller constraint. On the other hand, if at the end of the operation more effort is available, then the procedure chosen need no change; the search continues by placing the effort where the posterior probability is highest. Therefore, the procedure has a sort of independence in relation the constraint. This point is a very important feature in real world's operations (a change in the weather, as example, can determine the end of search) and characterizes what will be defined later as an UNIFORMLY OPTIMAL SEARCH PLAN.

C. SWEEP WIDTH

In the preceding section, a quantity "W" equal to 15 miles, was used as a distance within which the sensor will detect a target with probability 1. In the real world however, things are not so simple. From a rigorous point of view, the detection ability of a sensor can only be described by a random variable which associates to each distance a probability of detection, for a given type of target, a given condition of propagation (light, sound, infrared, radar waves, etc.) a given level of personnel's aptitude and a multitude of other, perhaps less important, factors. It should be obvious that it is not

32

and

$$\theta_3 = \frac{(A_1 + A_2 + A_3)\,\varepsilon_3}{V}$$

$$= \frac{3 \times 30 \times 50 \times 0.04621}{100} = 2.079\underline{4}5 \text{ hours}$$

To be spent one third in each one of the cells I, II, and III

The posterior distribution after ε_3 is:

$$\tilde{\tilde{\tilde{P}}}_I = \tilde{\tilde{\tilde{P}}}_{II} = \tilde{\tilde{\tilde{P}}}_{III} = \tilde{\tilde{\tilde{P}}}_{IV} = \frac{(\frac{2}{7}) \text{ EXP } [-15\varepsilon_3]}{\underline{4}\,(\frac{2}{7}) \text{ EXP } [-15\,\varepsilon_3]} = \frac{1}{\underline{4}}$$

Figure 9 displays the variation of the cell's probabilities, due to the application of ε_1, ε_2, and ε_3

Before the
Search

After ε_1

After ε_2

After ε_3

Figure I-9

easy to obtain the distribution of such a random variable.
In fact, in some attempts to derive the probability of detection for search radar, considering a given distance, the variance of the sample was so big, that the corresponding confidence internal for the estimate was wider than would reasonably be acceptable. The way normally used to overcome such difficulties, is by means of a lateral range curve, from which a SWEEP WIDTH is derived.

Consider Figure 10 where the vertical lines represent the target's paths. The distance between the sensor and the intersection of a vertical line with the horizontal axis, is the lateral range of the corresponding path.
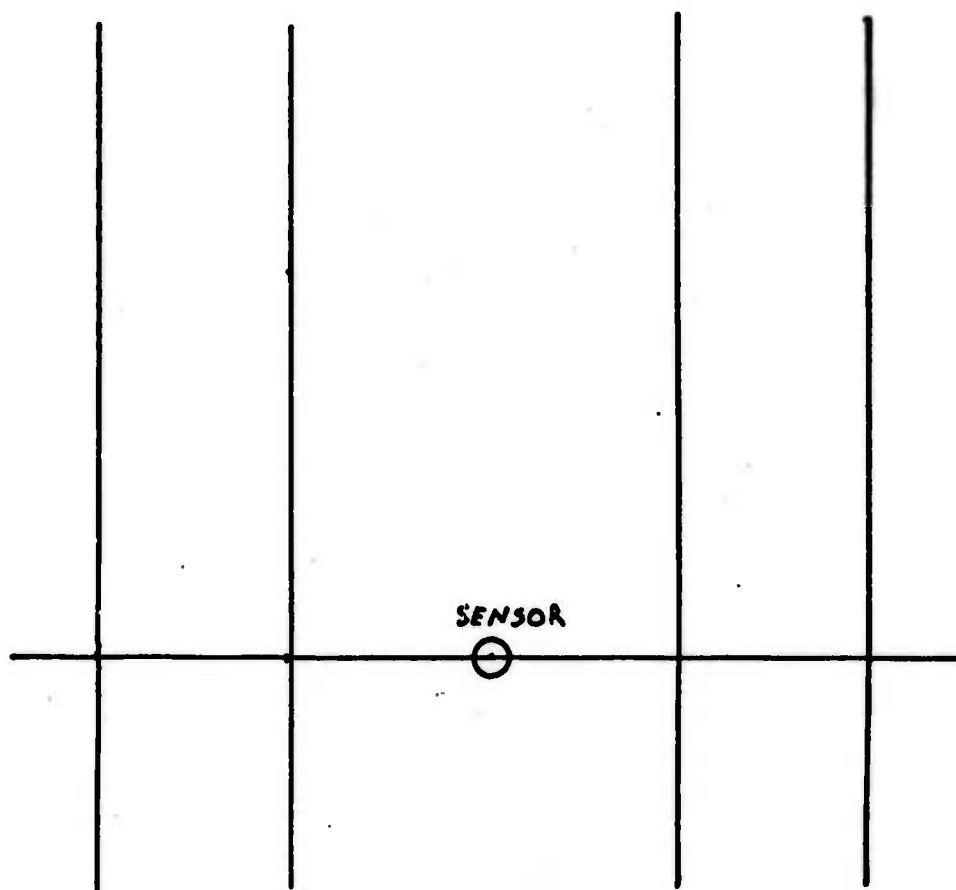
SENSOR

Figure I-10

If several targets are passed through each path and the percentage of those detected computed, no matter in what point of the path it happens, this percentage can be taken as the probability of detection of a target for which the closest point of approach is that lateral range. By plotting those probabilities against distance, the lateral range curve for the sensor in question is obtained. An example of lateral range curve for a radar, Figure I-11 shows:
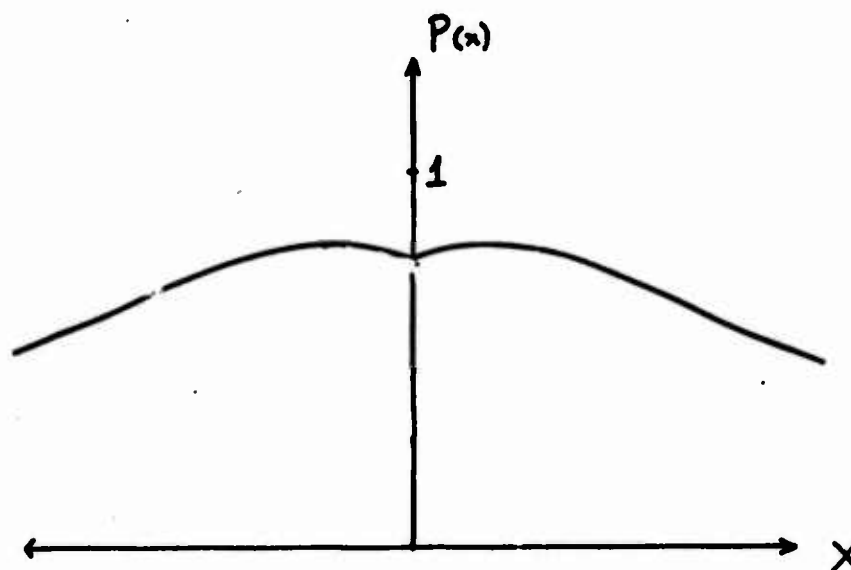

Figure I-11

Observe that the lateral range curve does not represent a probability distribution, observe also that the probability drops near the origin. This is due to the sea return.

35

The SWEEP WIDTH, which will be denoted by W is the area below the lateral range curve. The "W" used in sections I and II corresponds to a lateral range curve as the one in Figure 12, where $R = \frac{W}{2}$ is called the range of a definite range law of detection, a simplification very much used in real situations.
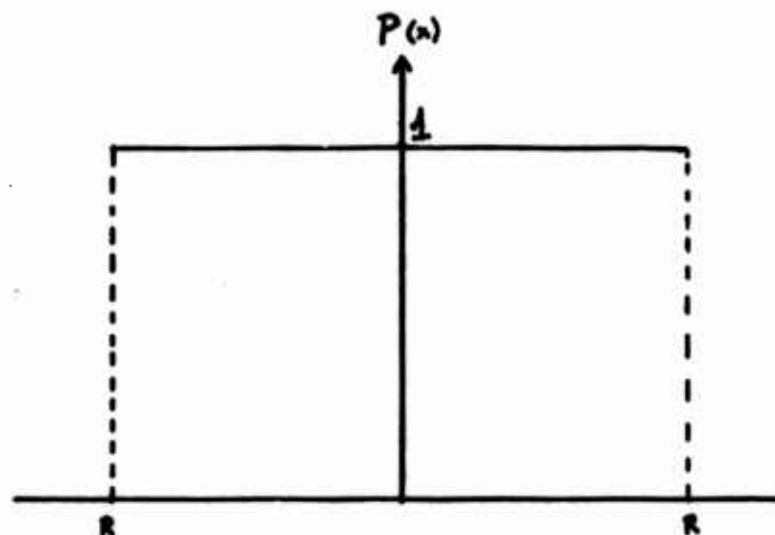

Figure I-12

Although in chapter II this subject will be discussed again, it is interesting to remark now, that in some situations, (chiefly in under-water search), it is necessary to treat the sweep width as a random variable.

D. MOVING TARGETS

Until now, all the derivations and concepts presented in this paper, were developed under the assumption that the target remains stationary during the search. If it is wanted to consider moving targets, the following remarks are important:

(a) The problem of effort allocation under these circumstances, is a very difficult one, and for most of the cases, has not been solved yet. The reason for this is that the prior probability distribution for the target's location is modified continuously in time, in an unknown way, by the target's movements. A case where the problem is tractable is where the target motion is deterministic, and with known parameters. In this case, the target "drags" the probability distribution in a known way, and as will be seen later, it is possible to solve the effort allocation problem.

(b) If, however, the search is to be made at random, then, the preceding discussion for the random search applies, whatever prior distribution is assigned to the target's location, and for any kind of target's motion. The rational is that, if the sensor's movements are at random then any fraction of the search area is equally likely to be scanned at a given time, and evidently it applies to the place where that target is at the considered time.

37

E. THE CONTINUOUS CASE

Suppose that instead of cells the S.A. is thought as a continuous set of points. Then, instead of cells' probabilities we will deal with points' probabilities, or more exactly, with a continuous bidimensional probability distribution.

Observe that in this case, most of our definitions should be reformulated. An allocation, for example, will be a function which assigns to each point of the S.A. an effort density. In a similar way L.D.F., and so on, must be redefined. This will be done later, but for the time being two points must be emphasized:

First, for theoretical purposes, the continuous approach is better, not only because it is closer to the real situation (after all a S.A. is continuous set of points), but also because of the mathematical tractability of continuous function, which permits us to derive general results and theorems.

Second, for practical purposes, the continuum of the S.A. is approximated by a large number of small cells and then the theory developed through the continuous model approximately applies.

Hence, all the intuitive discussion of I-A to I-D are entirely justified and the more sophisticated discussion in following pages will have the sole purpose of improving foundations.

II. __BASIC THEORY__

Under this heading, what is aimed is not only to formalize from a mathematical point of view the intuitive notions introduced in I, but also to provide a general overview of the main results of search theory.

A. SEARCH SPACE

As was mentioned before, the target's location is described by a probability distribution. More precisely, it is said that this location is a random variable.

The range of this random variable, (or better, the essential range, see [4]), which is some subset of the -plane, is called the search space.

Observe that the search space can be infinite, (as will be shown for the case of the bidimensional normal distribution, for example), and this is not a convenient feature, at least in practical situations. Because of this, search area is defined as:

B. SEARCH AREA

Is the subset of the search space where, search operations will be conducted.

As was shown in definition II-A, the search space, being any subset of the plane, can in fact be very pathological. Consider, as an example, the search space represented in Figure II-1.
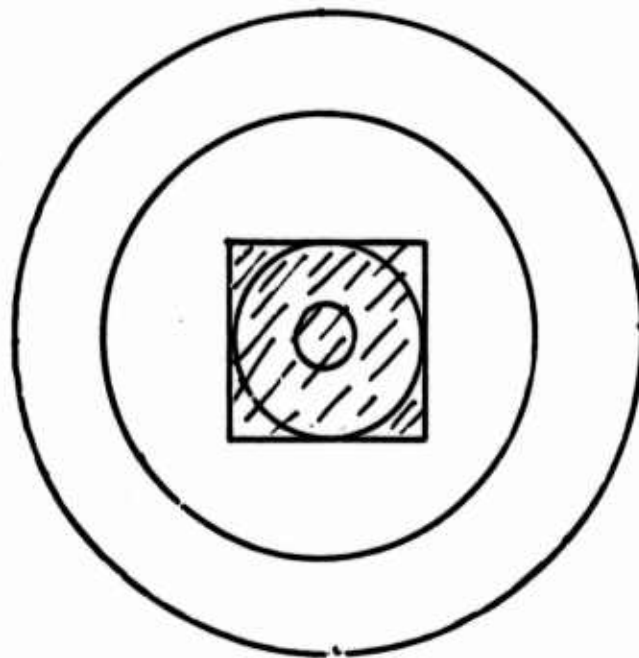
Figure II-1

The concentric circles stand for isoprobability curves.
It is a case of a infinite search space. If the dashed
area encloses a satisfactory amount of probability for
the purposes sought, the operations can be confined to that
same area. The dashed portion of Figure 1 is then an
example of search area.

Finally, in accordance with the model used, the search
area can be dealt with as a continuum, or it can be divided
in an arbitrary number of cells, assigning to each one the
probability mass it encloses. Furthermore if cell "k"
encloses probability P_k it is assumed that this amount of
probability is uniformly distributed in the cell.

C. LOCAL DETECTION FUNCTION

Let "x" be a point in the search area and "Z" the effort
density placed there. The function which to each pair (x, Z)
assigns the probability of detecting a target located in x is

called a local detection function. If points are replaced
by cells and for sake of notational consistency "x" by "j"
in the preceding lines, the definition given holds for the
descrete case.

Note that "x" is a bidimensional vector, $\bar{x} = (x_1, x_2)$ and
"Z" a non-negative real number. The L.D.F. will be referred
to as $B(\bar{x}, Z)$. This notation takes in account the possibility
(very remote in real world problems) of different forms of
L.D.F. in different cells.

D. L.D.F.s AND DETECTION RATES

In I two L.D.F.s were derived one for systematic and
the other for random search. Here, a more precise approach
will be used to obtain a general formulation from which the
above mentioned L.D.F.s are particular cases.

Consider a detection system which operation consists of
a series of instantaneous observations, for example, measuring
a distance to a submarine by a step scanning sonar. It is
called a discrete system. The mentioned observations will
be called glimpse and let "g" denote the probability of
detection on a glimpse. If "g" is a constant, the glimpses
are independent and if "N" denotes the random variable "number
of glimpses until detection occurs". Then:

$$P(N=n) = g(1-g)^{n-1} \qquad \text{(II-I)}$$

$$P(N \leq n) = 1 - (1-g)^n \qquad \text{(II-II)}$$

41

Because "N" is evidently a geometric distributed random variable and (II-I) and (II-II) are respectively its probability mass function and cumulative distribution function. Suppose now that the system used is a continuous one, i,e, it does not operate only at discrete points in the time, but instead it may reach detection at any moment. Rather than define and instantaneous probability as done before, let γ be a "detection rate" defined in such a way that $\gamma\Delta t$ is the probability of detecting the target during an interval of length Δt, given it has not been detected before. In this case let "T" be a random variable which stands for the time at which detection occurs. In analogy with (II-II), it is intended to find:

$$P\ (t) = P\ (T \leq t) \qquad\qquad (II-III)$$

consider

$$Q(t) = 1-P(t) = P(T \geq t) \qquad\qquad (II-IV)$$

the "survival function" of "T". It is clear that the relation

$$Q(t+\Delta t) = Q(t)\ (1-\gamma\Delta t) \qquad\qquad (II-V)$$

holds, by arguing as follows: $Q(t+\Delta t)$ is the probability of the event (detection does not occur by the time t+ Δt), which is logically equivalent to the intersection of the events A=(detection does not occur by time t) and B=(detection does not occur during the interval (t+ Δt)).

But by definition:

$$P(A) = Q(t)$$

and recalling that the interval (t + Δt) has length "Δt", and that "$\gamma\Delta t$" stands for the conditional probability that detection

42

has occured during the interval, given it has not happened
before:

$$P(B/A) = 1-\gamma\Delta t$$

from probability theory:

$$P(A\cap B) = P(A)\ P(B/A), \text{ hence} \qquad\qquad (II-V)$$

A simple manipulation of (II-V) yields:

$$\frac{Q(t+\Delta t) - Q(t)}{\Delta t} = - \gamma Q(t)$$

or in the limit when $\Delta t \to 0$

$$Q'(t) = - \gamma Q(t)$$

whose solution is:

$$Q(t) = Q(0)\ EXP\ [-\gamma t]$$

Since

$$Q(0) = 1$$

Obviously

$$P(t) = 1 - EXP\ [-\gamma t] \qquad\qquad (II-VI)$$

Observe, however, that both (II-II) and (II-VI) have been
derived under the assumption that "g" and γ remain costant
all time. If its variation with time is to be taken in
account, the expressions can be modified by writing:

$$P(N\leq n) = 1 -\prod_{i=1}^{n}\ (1-g_i) \qquad\qquad (II-VII)$$

for the discrete case, or

$$P(t) = 1 - EXP\ [- \int_{0}^{t} \gamma(t)\ dt] \qquad\qquad (II-VIII)$$

for the continuous, where "g_i" stands for the detection
probability on the ith glimpse, and $\gamma(t)$ is a time dependent
detection rate. Although there are many reasons by which g
and or "γ" do not remain constant in time (a human factor

43

reason is the fact that the device's operator "learns" while
he is watching the display), only the distance variation will
be considered as a source for the time dependence of "g" or
"γ". Thus, if by denoting r_i as the distance to the target
at the time of ith glimpse, and "r (t)" the distance at time
"t"

$$g_i = g(r_i)$$

and

$$\gamma(t) = \gamma[r(t)]$$

By substituting in (II-VII) and (II-VIII)

$$P(N \leq r) = 1 - \prod_{i=1}^{r} (1-g(r_i)) \qquad (II-IX)$$

and

$$P(t) = 1 - EXP [- \int_0^t \gamma [r(t)] dt] \qquad (II-X)$$

notice that the probability of detection during the interval
$(t_0, t_0 +dt)$ can be computed if it is a product of probability
of no detection before t_0 (which is given by 1 minus formula
(II-X), times probability of a detection during "dt" (which
is $\gamma [r(t_0)] dt$), or:

$$\gamma [r(t_0)] dt \ EXP [- \int_0^{t_0} \gamma[r(t)] dt]$$

but, on the other hand, it is also equal to the difference

$$P(t_0 + dt) - P(t_0) \simeq P'(t_0) dt$$

thus

$$P'(t_0^-) dt = \gamma [r(t_0)] dt \ EXP [- \int_0^{t_0} \gamma [r(t)] dt]$$

and

$$P'(t_0) = \gamma [r(t_0)] \ EXP [- \int_0^{t_0} \gamma [r(t)] dt] \qquad (II-XI)$$

44

$P'(t_0)$ is the value at t_0 of the density function of random variable "time until detection".

In Figure II-2, which is taken part from [1], part from [5] five examples of $\gamma[r(t)]$ are shown:

A) case I applies when (r) reaches a finite maximum at distance zero:

B) case II is when the maximum is infinite,

C) case III shows the effect of sea return shifting the maximum away from the origin.

D) case IV is when the probability is one when $r < r_0$.

E) case V is when the so called definite law of detection occurs. In this case, detection is sure to occur as soon as the target reaches the dashed region and is impossible outside of it.
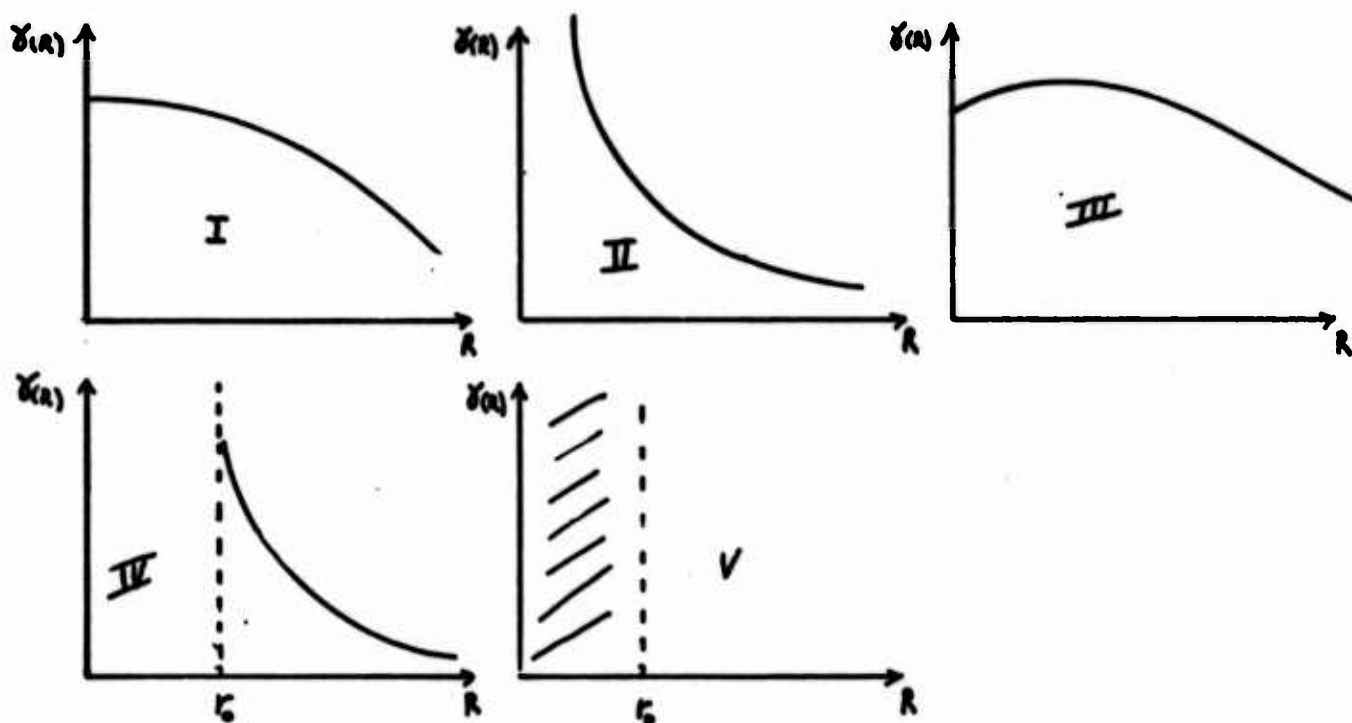

Figure II-2

Another important form of detection rate used for the case of visual detection is "the inverse cube law of sighting", whose expression is:

$$\gamma(r) = \frac{Kh}{(h^2 + r^2)^{3/2}} \qquad \text{(II-XII)}$$

Where "h" is the observer's height, "k" a constant which depends on fixed factors as contrast of wake against ocean, observer's ability, metereological conditions, area of the target's wakes etc......, and "r" is the distance. The above formula is derived under the assumptions that the observer detects the target by seeing its wake, and " $\gamma(r)$ " is proportional to the solid angle subtended at the point of observation by the wake.

When "r" is much large that "h" (II-XII) becomes:

$$\gamma(r) = \frac{Kh}{r^3} \qquad \text{(II-XIII)}$$

The details concerning the derivation of (II-12) are in [1].

E. LATERAL RANGE CURVES AND SWEEP WIDTH.

Although some nice expressions for $\gamma(r)$ were derived in last section, nothing was said yet about how to deal with r, the distance between the target and the sensor.

It is apparent, however, that it is very difficult to keep track of the distance variation when performing a search. For, if there is relative motion between target and sensor, at each instant, r, will assume a different value, and therefore P(t) will depend in the law of variation of r, as function of time. This is extremely inconvenient, because even for very simple types of motion, this problem can be an extraordinarily difficult one.

46

In order to avoid the mentioned difficulty some theory must be developed. Suppose there is a need to compute the probability of detection along a given track. A rectangular (x_1, x_2) coordinate system, will be adopted, with the sensor at the origin $(0,0)$. Notice that a relative motion approach is being used. The scalar distance from the target to the sensor, is a function of the time and is denoted ad r (t). if the track that interest is the one in Figure 3, it is easy to see that $r(t) = \sqrt{x_1^2\ (t) + x_2^2\ (t)}$.
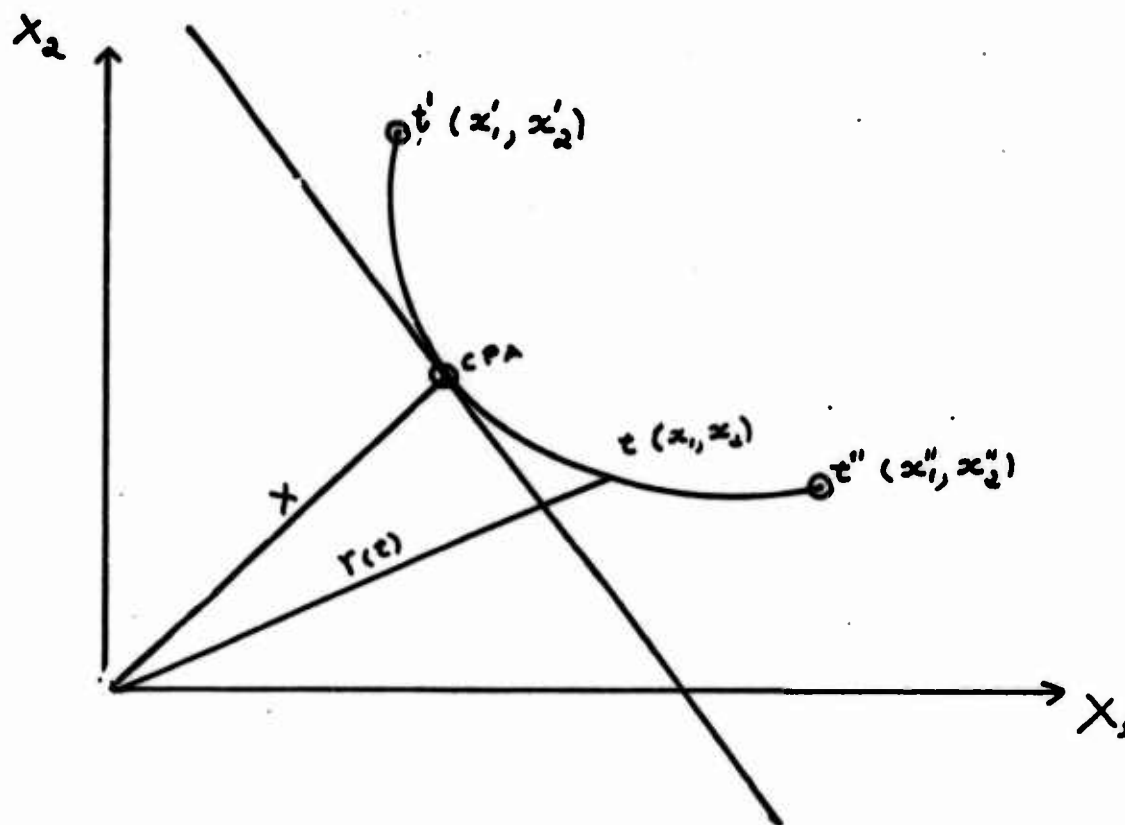

Figure II-3

47

At this point it is convenient to define a quantity "F", which is a characteristic of each possible path, and which is known as the sighting potential along the considered track. In order to make clear, the dependence of the sighting potential on the track it will be denoted by F(.), where the dot is to be replaced by some indicator of which track is being considered at the moment. Example: Suppose the paths are numbered. Then F(1) refers to path 1, and so on. By difinition:

$$F (\cdot) = \int_{t'}^{t''} \gamma \left[\sqrt{x_1^2 (t) + x_2^2 (t)} \right] dt$$

For the case of the continuous looking, and

$$F (\cdot) = - \sum_{i-1}^{n} \ln \left[1 - g \left(\sqrt{x_1^2 (t) + x_2^2 (t)} \right) \right]$$

for discrete model. By the use of the F(.) concept, equations (II-IX) and (II-X) may be united into

$$P(\cdot) = 1 - EXP [-F(\cdot)]$$

Which gives the probability of detection along the considered track, no matter if a continuous looking is being used, or separate glimpses. The verification is a matter of trivial algebra. It is important to state the property of additivity which the sighting potentials have. Suppose the probability of detection along at least one of two independents tracks (say track 1 and track 2) is to be computed, then:

$$P(1) = 1 - EXP [-F(1)]$$

$$P(2) = 1 - EXP [-F(2)]$$

48

The only way to fail in detecting is to fail in both paths or

$$1 - P(1 \text{ or } 2) = [1-P(1)] \ [1-P(2)]$$

but

$$[1-P(1)] \ [1-P(2)] = (EXP[-F(1)]) \ (EXP \ [-F(2)]) = $$
$$EXP \ [-(F(1) + F(2))]$$

then $P(1 \text{ or } 2) = 1 - EXP \ [-(F(1) + F(2))]$

what means that if F(1) and F(2) are known, just by adding them, the sighting potential along the total path is obtained.

The additivity applies to any number of paths and can be useful to deal with complicated tracks, by decomposing them in a series of simpler ones, whose potential can be easily computed and just by adding the sighting potentials.

Consider now the point labelled "CPA" in Figure 3. The initials stand for closest point of approach, and the distance x from the CPA to the sensor, is the target' lateral range. Observe that $\gamma(r)$ is a function of the distance and therefore, for all practical purposes, it is zero, beyond a certain distance. Thus it only makes sense to consider the variation of $\gamma(r)$ for values of r between x and r_0 such that $\gamma(r)=0$ if $r>r_0$. Suppose that we replace the actual path in figure 3, by its tangent at the CPA. The straight line being as long as we want, all distances between x and r_0 can be considered. On the other hand, the linearity of the path makes the problem much simpler. For example to compute F(.), instead of an integral along a path, a common Rieman integral is enough Finally, observe that when the relative velocity between sensor and target is constant, and the most common example of this

49

case is a sensor moving at constant velocity and a stationary target, then the relative track is actually a straight line.

Consider now Figure 4, where a straight line path is represented.

Figure II-4
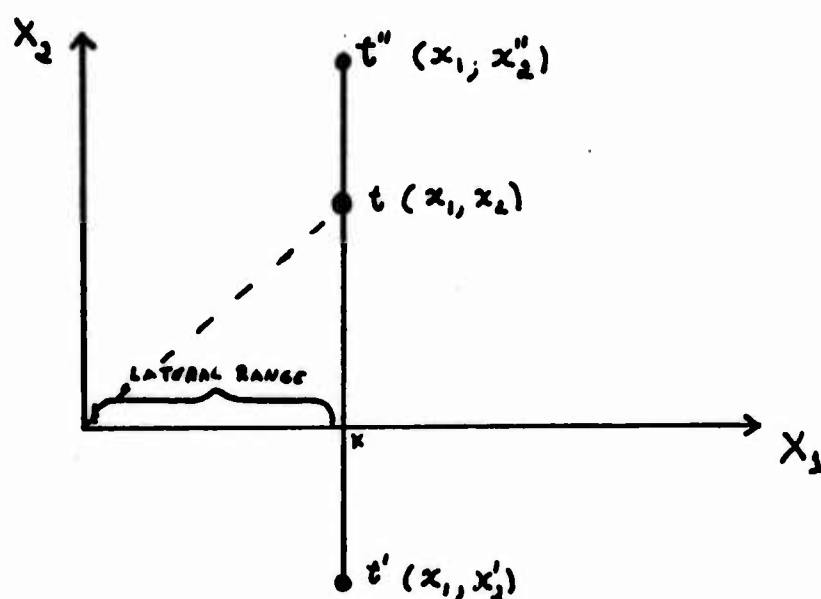
as it is already known:

$$\Upsilon(t) = \sqrt{x_1^2(t) + x_2^2(t)}$$

but now $x_1(t)$ is a constant, "x" the lateral range and as seen in figure 4, is the distance of the closest approach between the target and the sensor. Then,

$$r(t) = \sqrt{x^2 + x_2^2(t)}$$

furthermore as the direction of the relative velocity vector is parallel to the "x2" axis (the axis were chosen in this way),

50

$$r(t) = \sqrt{x^2 + x^2 t^2}$$

$$F(\cdot) = \int_{t'}^{t''} \gamma \left(\sqrt{x^2 + v^2 t^2} \right) dt$$

$$F(\cdot) = - \sum_{i-1}^{n} \ln \left[1-g \left(\sqrt{x^2 + v^2 t^2} \right) \right]$$

For the continuous looking case and separate glimpses case respectively.

Observe that now, depending on the form of the detection rate, both the integral and the sum above looks like mathematically tractable. In fact for many cases they have been evaluated in simple mathematical form. The interested reader can see [1] for a complete discussion. For the sake of further simplification, let the case where the path is an infinitely long straight line be examined. In this case $t' \to -\infty$ and $t'' \to \infty$ and the sum, which is used in the definition of the separate glimpses cases, becomes a series. In both cases however, $F(.)$ is now a function of the lateral range only. Thus it is more appropriate to write $F(x)$ instead of $F(.)$. The probability of detection along such a path, which by the same reasons is also a function of "x" only, can be computed

$$P(x) = 1 - EXP [-F(x)]$$

which defines a function called a lateral range curve. Therefore, even though we cannot truly handle general paths, we can approximate them by use of $P(x)$.

Hence forward, in this paper only the continuous looking case will be dealt with. The interested reader can consult

51

[1] or [5] for results and discussion concerned to the separate glimpses case. As it has been shown before

$$F(\cdot) = \int_{t'}^{t''} \gamma \left(\sqrt{x^2 + v^2 t^2} \right) dt$$

when $t' \to -\infty$ and $t'' \to \infty$:

$$F(x) = \int_{-\infty}^{\infty} \gamma \left(\sqrt{x^2 + v^2 t^2} \right) dt$$

or, considering that

$$vt = x_2$$

it is obtained

$$F(x) = \frac{1}{v} \int_{-\infty}^{\infty} \gamma \left(\sqrt{x^2 + x_2^2} \right) d x_2$$

and

$$P(x) = 1 - EXP \left[-\frac{1}{v} \int_{-\infty}^{\infty} \gamma \left(\sqrt{x^2 + x_2^2} \right) d x_2 \right]$$

Note that $P(x)$ is by no means a density function.

EXAMPLE I

LATERAL RANGE CURVE FOR THE DEFINITE RANGE LAW:

In this case:

$$\gamma(r) = \begin{cases} \infty & \text{if } r \leq r_0 \\ 0 & \text{otherwise} \end{cases}$$

recalling from Figure II-5

Therefore:

$$F(x) = \begin{cases} \infty & \text{if } x \leq r_0 \\ 0 & \text{if } x > r_0 \end{cases}$$

52

And

$$P(x) \begin{cases} 1 & \text{if } x \le r_0 \\ 0 & \text{if } x > r_0 \end{cases}$$
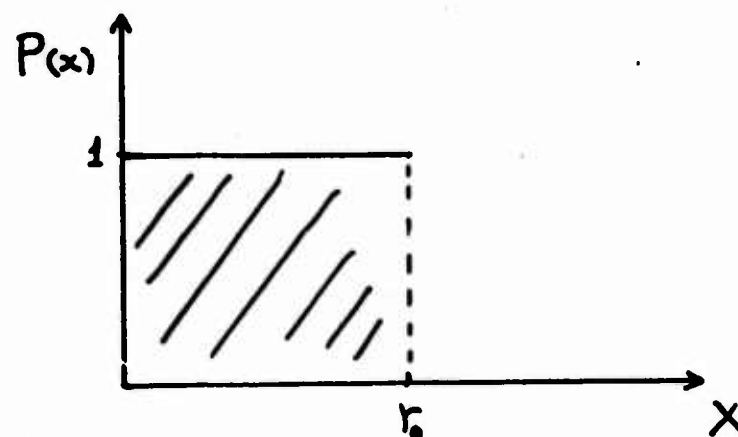
and its graph:

Figure II-5

EXAMPLE II

LATERAL RANGE CURVE FOR THE INVERSE CUBE LAW.

The detecting rate is

$$\gamma(r) = \frac{K}{r^3}$$

or

$$\gamma(\sqrt{x^2 + v^2 t^2}) = \frac{K}{(x^2 + v^2 t^2)^{3/2}}$$

and

$$F(x) = \frac{K}{V} \int_{-\infty}^{\infty} \frac{dx_2}{(x + x_2^2)^{3/2}} = \frac{2K}{Vx^2}$$

so that

$$P(x) = 1 - EXP\left[-\frac{2K}{Vx^2}\right]$$
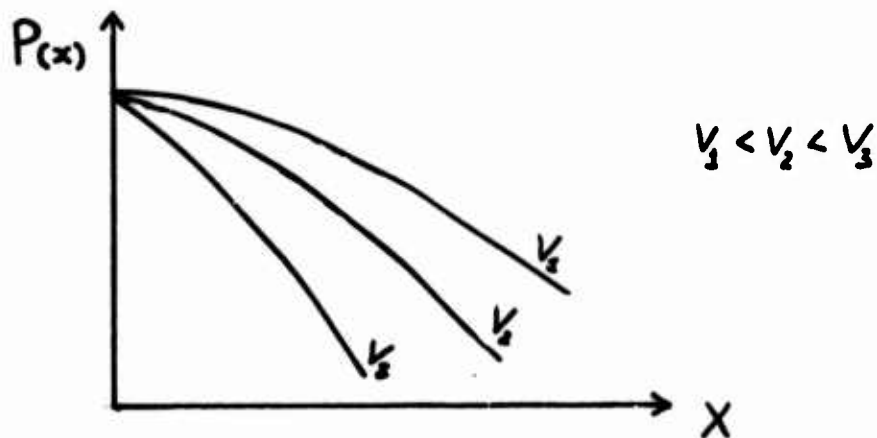
53

and its graph

$$V_1 < V_2 < V_3$$

Figure II-6

Observe that in this case P(x) depends on the speed, and
assuming all other factors remain constant, the probability
of detection at any lateral range decreases as the speed
increases.

As it has been seen before in two example, it is not very
difficult to obtain a compact mathematical form for $P(x)$.
Furthermore the parameters involved ("r_o" in example I and
"k" in example II) can be estimated by performing the experi-
ment described in chapter I. Now, in a search situation, the
target lateral range is not known for sure. For, if the
opposite is true, then no longer is necessary to search. To
take this uncertainty in account, a random variable X, is used
to described the lateral range.

Suppose that it is desired to compute the probability,
P, of detecting a target in a given search no matter how long
it takes. Recall that $p(x)$ was defined as the probability of
detecting a target, given that its lateral range is x, and

54

observe that the detection can occur at any point on the infinite straight line path used to define X, and thus, by any time. Further, if X is a random variable, and its density is denoted by $f_X(x)$, then, from probability theory:

$$P = \int_0^\infty P(x) \, f_X(x) \, dx \qquad \text{(II-XIV)}$$

Which is a very useful expression, as the next examples will show.

EXAMPLE 3

Assume that X is uniformly distributed in the interval (0,D). This is the case of a barrier with length 2D, in the center of which the sensor is located. If a target tries to trespass such a barrier, and is equally likely to cross it at any point, then the distribution of X is the one mentioned. Or:

$$f_X(x) = \begin{cases} \dfrac{1}{D} & \text{if } 0 \le x \le D \\[2mm] o & \text{otherwise} \end{cases}$$

Where "D" stands for a distance such that:

$$P(x) = 0 \quad \text{if} \quad x > D \qquad \text{(A)}$$

what means that the detection system is not effective at ranges larger than D. In this case (II-XIV) is:

$$P = \frac{1}{D} \int_0^D P(x) \, dx$$

or what amount to the same due to equation (A):

$$P = \frac{1}{D} \int_0^\infty P(x) \, dx$$

55

But, the integral in the formula above is just half of the area below the lateral range curve. By definition, this area is called "sweep width", or in symbols

$$W = 2 \int_0^\infty P(x) \, dx = \int_{-\infty}^\infty P(x) \, dx$$

so the formula for the probability of detection becomes

$$P = \frac{W}{2D}$$

Observe that the formula above justifies the use that way made of "w" in I. For, if 2D is distance such that P=1 the definition of "w" used in the mentioned example is reafirmed. Figure 7 below pictures the situation.
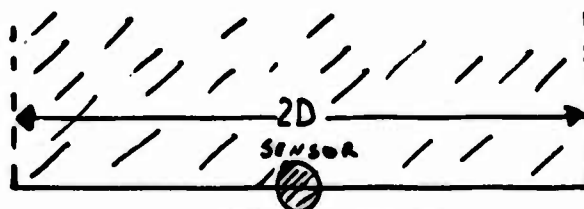
Figure II-7

Observe that the dashed area is swept with efficiency w/2D and that is the motive of the name sweep width. Under the same assumption about the lateral range distribution, two expressions for "w", will be derived:

SWEEP WIDTH FOR DEFINITE RANGE LAW

In this case, recalling from example 1:

$$P(x) = \begin{cases} 1 & \text{if} \quad x \leq r_0 \\ 0 & \text{otherwise} \end{cases}$$

Then

$$W = 2 \int_0^{r_0} dx = 2 \, r_0$$

56

which is the sweep width used as example in I. It is also twice the dashed area in Figure 5.

EXAMPLE 5

SWEEP WIDTH FOR INVERSE CURVE LAW

In this case as established in example 2:

$$P(x) = 1 - EXP \left[- \frac{2K}{Vx^2}\right]$$

Then

$$W = 2 \int_0^\infty P(x) \ dx =$$

$$= 2 \int_0^\infty \left\{1 - EXP \left[- \frac{2K}{Vx^2}\right]\right\} \ dx$$

The result can be obtained by the use of integration by parts:

$$W = 2 \sqrt{\frac{2\pi k}{V}}$$

In real life situations, the track's lengths are not infinitely long. However, if they are long enough in the sense that the sensor's detection capability tends to zero beyond a certain point in the path, then the preceding discussion applies. Suppose the length of the track is:

$$Y = x_2'' - x_2'$$

57

such that

$$\gamma(r) = 0$$

If $r > \sqrt{x^2 + (x_1'')^2}$ or $r > \sqrt{x^2 + (x_2')^2}$

For any lateral range x,. For this kind of situation the time during what the target is exposed to detection is:

$$T = \frac{Y}{V}$$

Because after this time $\gamma[r(t)] = 0$, then, the detection rate along the path:

$$\lambda = \frac{P}{T}$$

but

$$P = \frac{W}{2D}$$

so

$$\lambda = \frac{W}{2DT}$$

or, substituting the value of "T"

$$\lambda = \frac{WV}{2DY}$$

but 2DY is the area swept by the sensor, then

$$\lambda = \frac{WV}{A} \qquad\qquad (II-XV)$$

58

is the average detection rate along the considered path. Now, suppose a sensor with velocity and sweep width known and a target whose location has a uniform distribution in the search area. If it is known how many paths are intended to be placed in the S.A. and the time spent in each one of them, then it is possible to compute the probability of finding the target by a suitable combination of each track's probability, which is a function of λ. Considering that the sum of the time spent in each path can be easily transformed in a total length and then in a density of effort, that probability of detection can be written in terms of density of effort, which is just the definition for L.D.F. Consider the following examples:

EXAMPLE 6

L.D.F. FOR RANDOM SEARCH

Suppose that:

(a) the target's location is uniformly distributed inside an area "A" which is extremely large compared to the sensor's influence area.

(b) The sensor's vehicle movement is such that any region of the area is equally likely to be covered.

(c) The tracks are mutually independent.

(d) The search vehicle's speed is always the same.

The preceding assumptions characterizes a random search. For this case the detection rate is a constant at any instant in time. For, due to the independence of the paths, the instantaneous probability of detection is the same for them all.

59

By a reasoning, which is similar in all aspects to the
one presented in II-E for the sighting potential's cases, the
probability of detection at or before time "t" is

$$P(t) = 1 - EXP [-\lambda t] = 1 - EXP [-\frac{WVt}{A}]$$

which leads to the L.D.F.

$$B(Z) = 1 - EXP [-WZ] \qquad (II-XVI)$$

that was already derived in I.

EXAPLE 7

L.D.F. FOR SYSTEMATIC SEARCH

The general case of systematic searching by parallel
sweeps is treated in [1]. Now, here will be derived the case
when "W" is given by a definite range law and stated, without
proof (which also can be found in [1]), the result for the
inverse cube law. For the first case recall that $W=2r_o$ and
that within such a distance the detection is sure to happen.
If "A" stands for the total search area, the portion of "A"
which was not yet swept after time "t" is

$$A(t) = \begin{cases} A - WVt & if \quad 0 \leq t \leq \frac{A}{VW} \\ o & if \quad \frac{A}{VW} < t < \infty \end{cases}$$

and by this time the detection rate is

$$\lambda(t) = \frac{WV}{A(t)}$$

60

therefore the probability of detection by time "t" is:

$$P(t) = \begin{cases} 1 - EXP \left[- \int_0^t \lambda(t)dt\right] = \frac{WV}{A} & \text{if} \quad o \leq t < \frac{A}{WV} \\ \\ 1 & \text{if} \quad \frac{A}{WV} \leq t \end{cases}$$

which leads to the L.D.F.

$$B(Z) = \begin{cases} Wz & \text{if} \quad o \leq t < \frac{A}{WV} \\ \\ 1 & \text{if} \quad \frac{A}{WV} < t \end{cases}$$

that agrees again with the one derived in I. For the second case, the inverse cube law, the sweep width seen in example 4 was found to be

$$W = 2 \sqrt{\frac{2\pi k}{V}}$$

Through some more elaborate arguments than the preceding ones, but still relying in the same general principles, it can be shown that the L.D.F. for the inverse cube law is

$$B(Z) = 2 \, \Phi \, (\frac{V\pi}{2} \, WZ) - 1 \tag{II-XVIII}$$

Where Φ stands for the cumulative distribution function of the standard normal distribution. Now it is possible to draw Figure 6, chapter I again, including the L.D.F. (II-XVIII). This is done in Figure II-8 below.
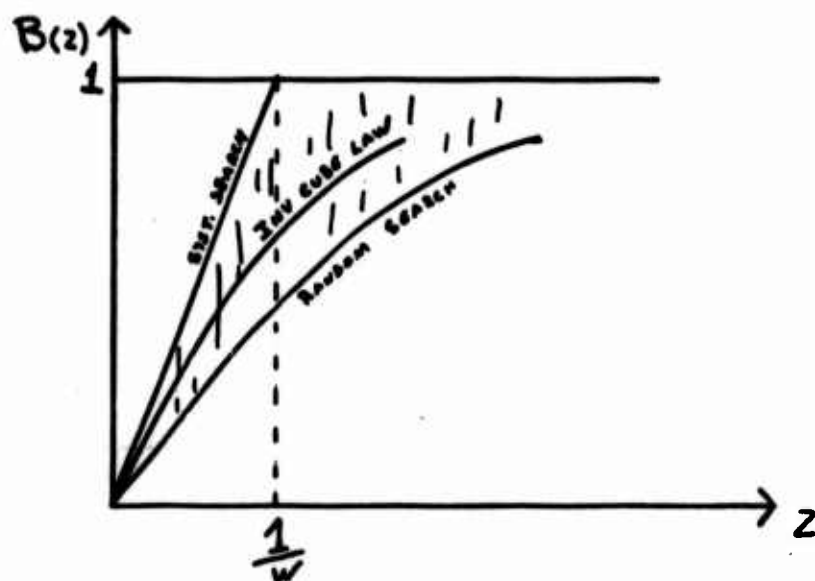
Figure II-8

As was said before in I-A, the best situation is systematic search with definite range law, and the worst situation is the random search case. The actual L.D.F., when performing a systematic search, must be situated between them.

F.   THE DISTRIBUTION OF THE TARGET'S LOCATION

The results established in the anterior sections depend in the uniform distribution of the target's location. However, it is not always the case that this type of distribution can be assumed to describe the target's position. The way to overcome this difficulty is by means of a search grid which is constructed as follows:

Suppose a bidimensional distribution is adopted to describe the target's location and that its C.D.F. is denoted by $F(x_1, x_2)$.
$$[X_1, X_2]$$

62

Then the probability that the target is located in cell R_k of Figure II-9 is

$$P_k = F(x_1', x_2') - F(\overset{..}{x}_1, x_2') - F(x_1', \overset{..}{x}_2) + F(x_1', x_2') \qquad (II\text{-}IXX)$$

$$[X_1 X_2'] \qquad [X_1 X_2] \qquad [X_1 X_2] \qquad [X_1, X_2]$$
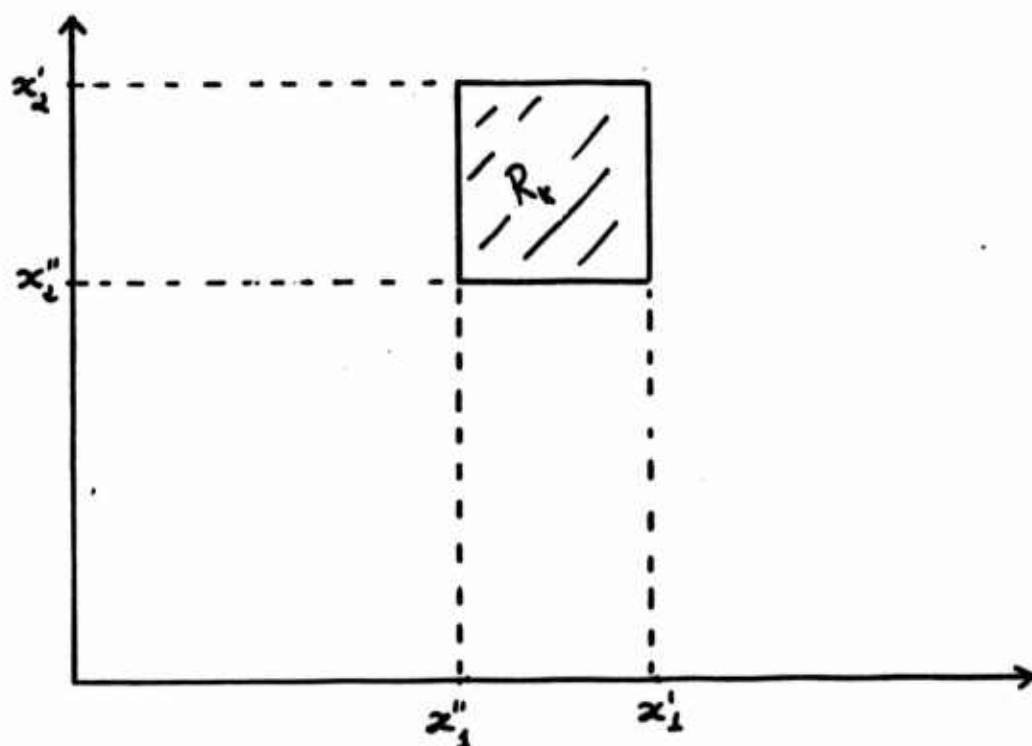

Figure II-9

Divide the area "A" under consideration into "J" non-overlapping cell in such a way that:

$$\sum_{j=1}^{J} P_j = 1$$

If we assume that for any "j" the probability "P_j" is uniformly distributed inside the cell "R_j", then for each cell all the preceding discussions apply. The mentioned

63

assumption leads to replace the original continuous distribution by a discrete one. If the size of the cells is not too big, the new model is not very different from the actual one. It is a very common practice to center the grid's coordinate system in the highest probability cell and set the axis parallel to the north/south orientation of the navigation chart. If the domain of the bidimensional distribution is infinite, the S.A. encloses the desired probability to be divided in the cells. Evidently in this case, the cells' probabilities do not quite add up to 1. Further details are found in [7] and [8]. Some examples of distributions most often encountered in search problems will be given below:

EXAMPLE 7

BIDIMENSIONAL UNIFORM DISTRIBUTION (INDEPENDENT X_1 AND X_2)

Although this is a very well known distribution, it is worth while to state its formal definition:
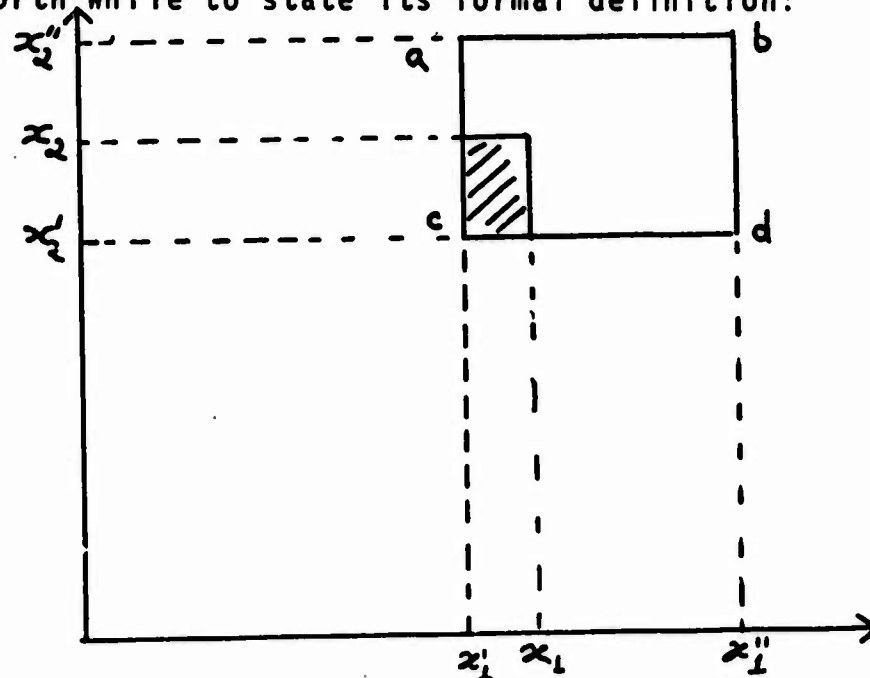

Figure II-10

64

Consider the rectangle [a,b,c,d] of Figure 10. A vector random variable $[X_1, X_2]$ has a bidimensional uniform distribution defined over [a,b,c,d,] if its density function is:

$$f(x_1,x_2) = \begin{cases} \dfrac{1}{(x_1'' - x_1')(x_2'' - x_2')} & \text{if } \begin{array}{l} x_1' < x_1 < x_1'' \\ x_2' < x_2 < x_2'' \end{array} \\ [X_1,X_2] \quad 0 \quad \text{otherwise} \end{cases}$$

The C.D.F. is

$$F(x_1,x_2) = \int_{x_1'}^{x_1} \int_{x_1'}^{x_2} f(x_1,x_2) \, dx_1 \, dx_2$$
$$[X_1,X_2] \qquad [X_1,X_2]$$

or be performing the intergration

$$F(x_1,x_2) = \begin{cases} 0 & \text{ij } x_1 \leq x_1' \text{ or } x_2 \leq x_2' \\ \dfrac{(x_1-x_1')(x_2-x_2')}{(x_1''-x_1')(x_2''-x_2')} & \text{if } \begin{cases} x_1' < x_1 < x_1'' \\ \text{and} \\ x_2' < x_2 < x_2'' \end{cases} \\ 1 & \text{if } \begin{cases} x_1 > x_1'' \\ \text{and} \\ x_2 > x_2'' \end{cases} \end{cases}$$
$$[X_1 X_2]$$

Observe that in Figure 10, the C.D.F. argument is represented by the dashed region and its value is given by the ratio between the region's area and the total area, as the name of "uniform" intuitively suggests. Another point is that the marginal densities for X_1, and X_2 are:

$$f_{X_1}(x_1) = \int_{x_2'}^{x_2''} f(x,M) \, du = \dfrac{1}{(x_1'' - x_1')}$$
$$[X_1,X_2]$$

65

and

$$f_{X_1}(x_1) = \int_{x_1'}^{x_1''} f(M, x_2)\, du = \frac{1}{(x_2'' - x_2')}$$

if x_1 and x_2 are inside the proper range and "0" otherwise.

Then

$$f(x_1, x_2) = f_{X_1}(x_1)\, f_{X_2}(x_2)$$
$$[X_1, X_2]$$

and also

$$F(x_1, x_2) = F_{X_1}(x_1)\, F_{X_2}(x_2)$$
$$[X_1, X_1]$$

all considering the proper range, of course. The above results characterize the independence of the random variables X_1 and X_2.

EXAMPLE 8

THE BIVARIATE NORMAL DISTRIBUTION

Th's is a very useful distribution in real life problems, for it is a well known fact that error distributions are in general normal. Navigation error can be decomposed in two components: Latitude and longitude, and if both are univariate normal their composition is bivariate normal. Example: if a ship in distress asks for help and gives her last fix, it is reasonable to use this distribution to describe her position. By using x_1 to denote the east/west direction, x_2 for the north/south. M for mean 6 for standard deviation and ρ_{X_2} for coeficient of correlation, the bivariate normal density:

$$f_{[X_1,X_2]}(x_1,x_2) = \frac{1}{2\pi\sigma_{X_1}\sigma_{X_2}(1-\rho^2)^{\frac{1}{2}}} \; EXP \left\{ - \frac{1}{2(1-\rho^2)} \left[\left(\frac{x_1 - M_{X_1}}{\sigma_{X_1}}\right)^2 - \right. \right.$$

$$\left. \left. 2\rho \left(\frac{x_1 - M_{X_1}}{\sigma_{X_1}}\right) \left(\frac{x_2 - M_{X_2}}{\sigma_{X_2}}\right) + \left(\frac{x_2 - M_{X_2}}{\sigma_{X_2}}\right)^2 \right] \right\}$$

If however it is assumed that the north/south and east/west errors are independent (which is common practice) then $\rho = 0$. Furthermore if the distribution is centered in the datum then $M_{X_1} = M_{X_2} = 0$. With these simplifications the above formula becomes:

$$f_{[X_1,X_2]}(x_1,x_2) = \frac{1}{2\pi\sigma_{X_1}\sigma_{X_2}} \; EXP \left\{ - \frac{1}{2} \left[\left(\frac{x_1}{\sigma_{X_1}}\right)^2 + \left(\frac{x_2}{\sigma_{X_2}}\right)^2 \right] \right\}$$

Figure 11 shows, the shape of $f_{[X_1 X_2]}(x_1 x_2)$ and the contours of equal probability, which are generally ellipses with the main axis laying in the coordinate axis, which corresponds to the higher variance. The contours are circles when $\sigma_{X_1} = \sigma_{X_2}$
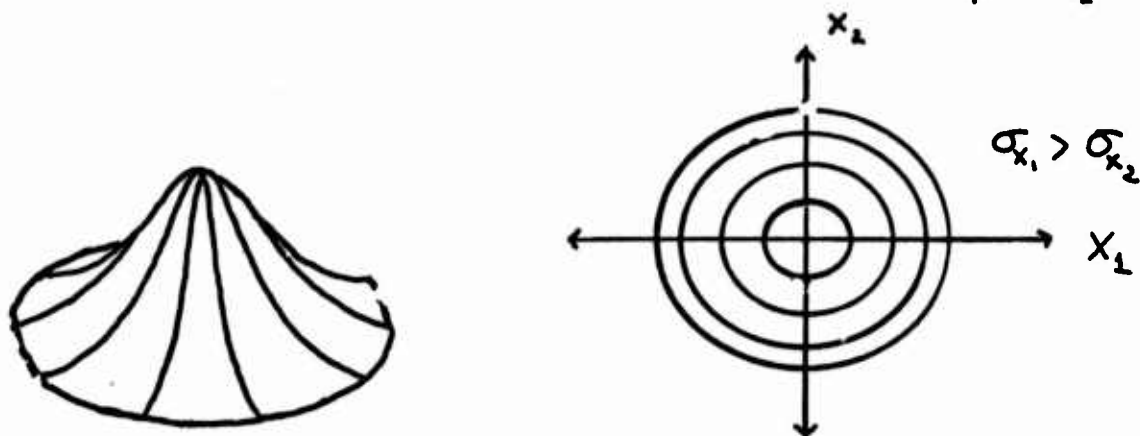

Figure II-11

67

Considering that the values of the standard normal CDF
are easily available in many tables and that, due to their
independence

$$F(x_1, x_2) = F_{X_1}(x_1) F_{X_2}(x_2)$$
$$[X_1, X_2]$$

it is very simple to compute any region's probability, in this
particular case. In reference to Figure II-9, formula (II-IXX)
becomes:

$$P_K = [F_{X_1}(x_1') - F_{X_2}(x_1'')] [F_{X_2}(x_2') - F_{X_2}(x_2'')]$$

But

$$F_X(x) = \phi \left(\frac{x-u}{\sigma_X}\right)$$

In this case M=0, then

$$F_X(x) = \phi \left(\frac{x}{\sigma_X}\right)$$

and

$$P_K = [\phi \left(\frac{x_1'}{\sigma_{X_1}}\right) - \phi \left(\frac{x_1''}{\sigma_{X_1}}\right)] [\phi \left(\frac{x_2'}{\sigma_{X_2}}\right) - \phi \left(\frac{x_2''}{\sigma_{X_2}}\right)] \quad (*)$$

Where ϕ, is the standard normal CDF. In order to derive the
dimensions of a search area for this case, the common usage is
to construct a rectangle, centered on the datum, whose sides'
length are usually expressed in terms of standard deviation,
and that encloses the desired amount of probability. This
can be done by replacing in (*) P_k by the desired probability,
and solving the equation backwards, taking advantage of the
symmetry of the normal distribution. Figure 12 picture the
case of a search area for a circular normal distribution, in
this case (*) becomes:

68

$$P_k = [\phi\left(\frac{x_1'}{\sigma}\right) - \phi\left(-\frac{x_1'}{\sigma}\right)]\,[\phi\left(\frac{x_2'}{\sigma}\right) - \phi\left(\frac{x_2'}{\sigma}\right)] =$$

$$[\phi\left(\frac{x_1'}{\sigma}\right) - \phi\left(-\frac{x'}{\sigma}\right)]^2$$
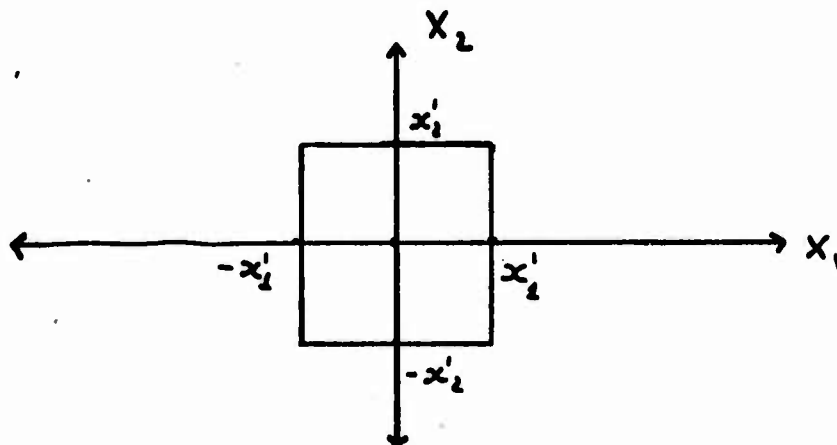
Figure II-12

And by consulting a table, the solution is straight forward.
As illustration, some values are tabled below:

Side Length	Probability Enclosed
1 Standard Deviation	0.466
2 Standard Deviation	0.911
3 Standard Deviation	0.994
4 Standard Deviation	$\simeq 1.0$

Finally there is another way of writing the density of
this distribution which sometimes is very helpfull. It is
the polar from and the formula that will be stated holds
for the circular case $(6_{x_1} = 6_{x_2})$

$$f(r,\theta) = \frac{r}{2\pi\sigma^2}\ \text{EXP}\ [-\frac{r^2}{2\sigma^2}]$$
$$[R_1\theta]$$

Where, as usual

$$x_1 = r \cos \theta$$
$$x_2 = r \text{ sen } \theta$$
$$r \geq 0$$
$$0 \leq \theta \leq 2\pi$$

EXAMPLE 9

MULTIPLE SCENARIOS DISTRIBUTIONS

An interesting problem which arises in searches on large scale, (for example the search for the H-bomb dropped accidentally in the Spanish cost in January 1966, see [7] for details) is the problem of multiple scenarios.

Suppose that from two different sources of information, two different datums are given to the analyst in charge of planning the search. Figure 13 pictures the situation.
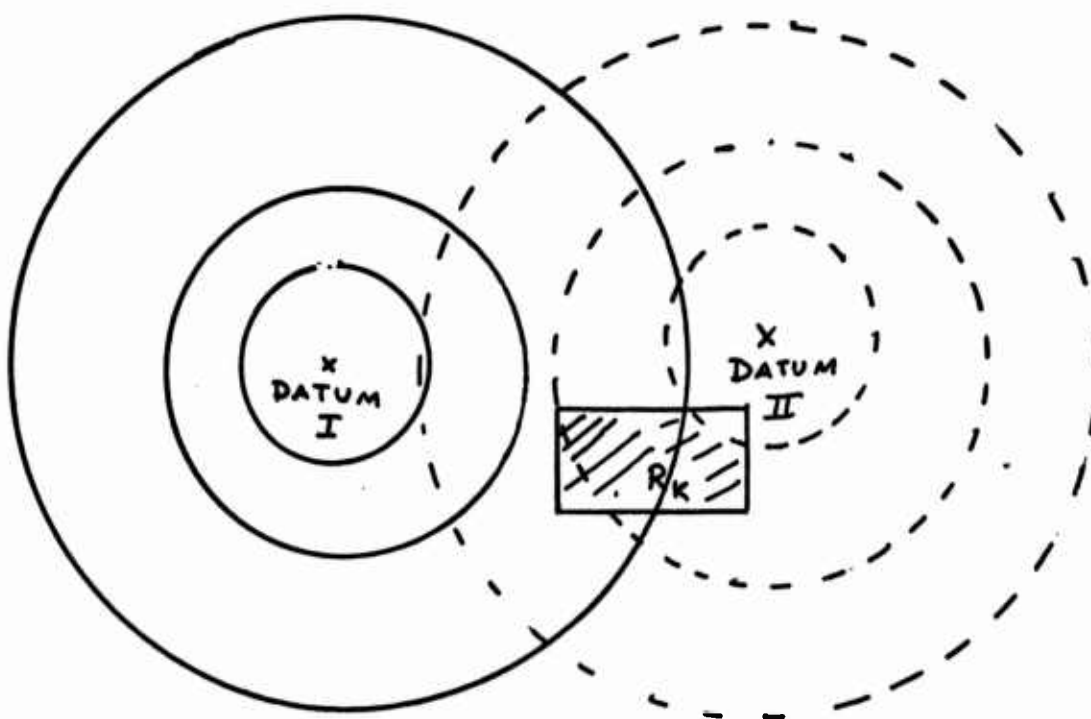
Figure II-13

Even assuming a circular normal distribution, the problem of where to center it still remains. The best approach, is to center a distribution on each datum and by weighting the reliability of each source by a number between 0 and 1, CRJ (which reads "credence in source "J"), obtain a distribution which then is used as the target's location distribution. Consider Figure 13 and suppose that the probability in the dashed cell R_k is to be computed. From probability theory is recalled that:

Prob (R_k) = Prob $(R_k$/DATUM I Is CORRECT) Prob (DATUM I is CORRECT) + Prob $(R_k$/DATUM II Is CORRECT) Prob (DATUM II Is CORRECT)

IF CRJ is taken as the probability of datum j being correct, then:

$P(R_k)$ = $P(R_k$/I) CRI + $P(R_k$/II) CRII

Care should be taken about the value of CRJ. This is a very subjective matter, and can have an enormous influence in the search.

EXAMPLE 10

OTHER DISTRIBUTIONS

Although the normal and the uniform are the most usual distributions in search problems, there are, besides the multiple scenarios situations, other cases where they do not apply, at least in the straightforward way presented here.

Suppose, for instance, that a position is available from a target, say normally distributed around a datum, and a radio direction finder bearing, with a uniformly distributed error of k degrees to each side. The situation can be visualized in Figure 14

Figure II-14

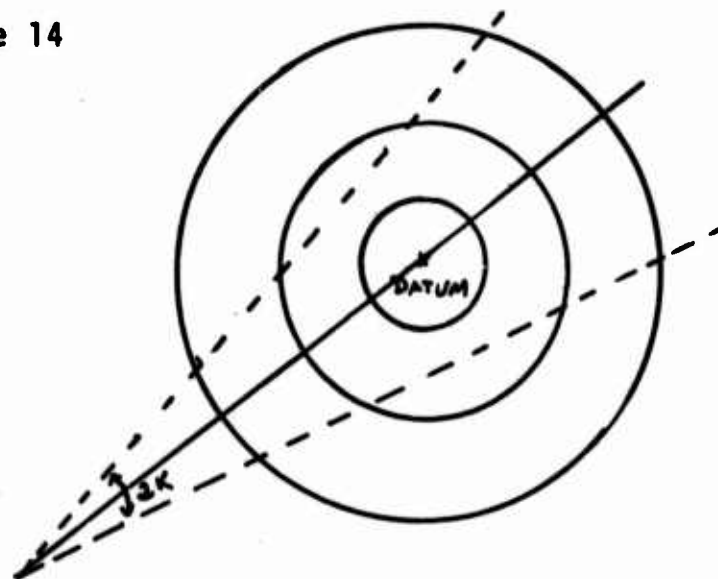It is not hard to imagine many other analogous situations. In any case, however, the procedure is the same, combining the probabilities by conditioning, to get a search grid. Sometimes, a computer is needed, either because of the big volume of operations or even due to the analytical impossibility of deriving the conditional distribution. In this last case, a monte carlo method must be used (see[7]).

G. SEARCH PLANS

In I the ideas of allocation, and of the cost of an allocation were discussed. Here, a more general idea of a search plan is introduced. Consider a function $M(j,\phi(t))$ such that:

(a) $M(j,\phi)$ is a allocation for each ϕ

(b) $M(j,\phi)$ is increasing in ϕ for $j=1....J$.

This function is called a search plan. The replacement of j by a vector (x_1, x_2), leads to the definition for the continuous case. Condition a) means that, for each cell (point) the function M assigns a density of effort z, which depends also in ϕ , the amount of effort available. Condition b) is the formal statement of the fact that it is impossible to take effort out of a cell (point), once it was placed there.

The definition of cost of an allocation carries over to the case of search plans. In fact, for the discrete case it is the same as defined in I, while for the continuous, the sum is replaced by an integral, an it becomes:

$$C(M(\underline{x},\phi) = \int_{\underline{x}} M(\underline{x},\phi)dx$$

Suppose now that ϕ is an increasing function of the time, $\phi=\phi(t)$. The problem which was solved in I, with the track length measuring the effort, and given by

$$Vt = L$$

where V remained constant, was an example of this situation.

73

In this case, $\phi(t)$ is simply Vt. If the search plans is such that

    a) $C(M(j,t)) = \phi(t)$

    b) Prob $(M(j,t)) = $ MAX $\{$Prob $(Z):C(Z) \leq \phi (t)\}$

Then M is a UNIFORMLY OPTIMAL SEARCH PLAN, denoted $M^{\#}$.

Observe that the above definition formalizes what was intuitively shown in I. For, if at any time t, the plan costs exactly the amount of effort by that time available, and reaches the maximum value of probability of detection which is attainable with that effort, then, besides of being the best possible plan for the constraint which is imposed, it is "constraint independent" in the sense that no constraint alteration will be able to alter the uniform optimum search plan.

In the next section it will be discussed how to construct uniformly optimal search plans.

H. RETURN FUNCTION AND UNIFORMLY OPTIMAL SEARCH PLANS

According to what was stated in I, the mathematical formulation of the resource allocation problem is:

    MAX $P(Z)$
    s.t.
    $C(Z) \leq C$

Which, in the context of uniformly optimal search plans, is equivalent to:

    MAX $P (M(j,\phi))$
    s.t.
    $C (M(j,\phi)) \leq \phi (t)$

74

Therefore, the problem is to find a search plan which by spending at each instant "t" all the effort which by that time is available, maximizes the probability of detection.

The discussion of the general solution for this type of problem depends strongly on mathematical programing methods, and is far beyond the scope of this paper. The interested reader can see [2], and [9], for applied and theoretical, respectively, details. However, it is possible to devise a solution for the mentioned problem, by using the concept of a return function, whose meaning can be intuitively explained, as following:

Suppose that an increment of effort, h, is applied in cell j. The corresponding increment in probability of detection, considering that an amount z was already applied, is:

$$P_j [B_j (Z + h) - B_j (Z)]$$

and the increment in cost is

$$hA_j$$

but

$$P_j [B_j (Z + h) - B_j (Z)] \simeq P_j B_j' (Z) h$$

therefore

$$\frac{P_j B_j' (Z) h}{A_j h} = \frac{P_j B_j' (Z)}{A_j}$$

is the ratio between the increase in effectiveness and the increase in cost, that results from placing a small increment of effort in cell j. This ratio is called the return function, and is denoted by $\lambda(z)$. But it is also a M.O.E. of what is being done, for the mentioned increment of effort will produce

75

better results where $\lambda(z)$ is maximum. In fact, (see [2] for proof) it can be formally stated: If M is the search plan such that the next increment of effort is applied in the cell where $\lambda(z)$ is highest, then $M=M^{\#}$ is the uniformly optimal plan. The preceding discussion which has been done for the discrete case, applies, with suitable transformations, to the continuous, as it will be shown in the examples.

EXAMPLE 11 (discrete case)

RETURN FUNCTION FOR THE RANDOM SEARCH L.D.F.

Let

$$B_j (Z) = 1 - EXP [- WZ] , j = 1, 2 ... J$$

then

$$B_j' (Z) = EXP [-WZ]$$

and

$$\lambda_j (Z) = \frac{P_j \; EXP \; [-WZ)}{A_j}$$

Suppose that all the cells have the same area. Then, the next increment of effort must be placed in the cell where $\lambda j(z)$ is highest. But this is exactly the procedure used in the heading I example, for Pj EXP[-wz] is the numerator of the posterior probability formula for cell j, and due to the equality of the denominators, only the numerators were taken in to considerations.

Therefore, for the considered L.D.F. both procedures are equivalent. Furthermore, this equivalence is a proof that the search plan used in the mentioned example, was uniformly optimal.

76

EXAMPLE 12

AN UNIFORMLY OPTIMAL PLAN FOR A CONTINUOUS SEARCH AREA CASE

Suppose the situation described by the following assumptions:

a) The target's location is described by a circular normal distribution, i,e:

$$f(r,\theta) = \frac{1}{2\pi\sigma^2} EXP\left[-\frac{r^2}{2\sigma^2}\right]$$
$$[R_1\theta]$$

for

$r \geq 0$

$0 \leq \theta \leq 2\pi$

b) The LDF for the case, which independs in (x_1, x_2), is:

$$B(Z) = 1 - EXP[-WZ]$$

and, therefore, assuming W=1

$$B'(Z) = EXP[-Z]$$

c) The available search effort is given by

$$\phi(t) = WVt$$

d) The return function, for the continuous case, is defined as:

$$\lambda(\underline{x},Z) = f(x_1,x_2) B'(\underline{x},Z)$$
$$[X_1,X_2]$$

or, in polar coordinates:

$$\lambda(r,Z) = f(r,\theta) B'((r_1\theta,Z)$$
$$[R,\theta]$$

77

for a given z.

To obtain an uniform optimal plan, it is necessary to find λ for each time t, namely $\lambda(t)$ and then solve this equation for

$$Z = M^{\#} ((r,\theta),t)$$

because $M^{\#}((r\theta),t)$) is the amount of effort that must be placed at point $(r_1\theta)$, by the time t.

Under these assumptions, it is possible to derive an expression for $\lambda(t)$ (see [2] for details) and the expression is:

$$\lambda(t) = \frac{1}{2\pi\sigma^2} EXP \left[- \left(\frac{WVt}{\pi\sigma^2}\right)^{\frac{1}{2}}\right]$$

by letting

$$\left(\frac{WV}{\pi\sigma^2}\right)^{\frac{1}{2}} = K$$

the search plan is:

$$M^{\#} ((r_1\theta)_1 t) = \left[\ln (f (r_1\theta)/\lambda(t))\right]^{+}$$

$$= \left[K\sqrt{t} - \frac{r^2}{2\sigma^2}\right]^{+}$$

where

$$[X]^{+} = \begin{array}{l} X \text{ if } X \geq 0 \\ 0 \text{ if } X < 0 \end{array}$$

which does not depend on θ. This is due to the fact that $f(r_1)$ is radially symmetric and also because B(Z) does not
$[r_1\theta]$
depend on position. Then, supressing θ, it is possible to write the plan in a more convemient way:

$$M^{\#}(r,t) = \begin{cases} K\sqrt{t} - \dfrac{r^2}{2\sigma^2} & \text{For } r \leq R(t) \\ \\ 0 & \text{For } r > R(t) \end{cases}$$

where

$$R(t) = 2\sigma^2 K\sqrt{t}, \ t \geq 0$$

Therefore, the optimal plan starts at the origin and spreads in such a way that by time t all the effort is placed inside a circle of radius R(t). This means that as the search progresses, the effort is accumulated around the origin. This makes sense because the origin has the highest prior target location density.

Such a plan can only be approximated in real life situations.

I. EXPECTED TIME FOR FINDING THE TARGET

In I, it was pointed out that the time taken to find a target, can be used as a M.O.E. for search plans. However, it should noticed that this time is a random variable T.

A reasonable approach is to minimize the expected value of the random variable T. Let μ be the expectation of T, and let P(t) be the probability of finding the target at, or before time t.

Notice that P(t) is the C.D.F. of T. By definition, the value of μ is:

$$\mu = \int_{-\infty}^{\infty} t\, P'(t)\, dt$$

but, due to the fact that T is a non negative random variable, it can be demonstrated (see [10]), that:

$$\mu = \int_{0}^{\infty} [1 - P(t)]\, dt$$

Now, consider the uniform optimal search plan $M^{\#}$. By the very definition of $M^{\#}$:

$$P[M^{\#}(\underline{x},t)] \geq P[M(\underline{x},t)]$$

where M is any other search plan. Therefore:

$$1 - P[M\#(\underline{x},t)] \leq 1 - P[M(\underline{x},t)]$$

and

$$\int_{0}^{\infty} \{1 - P[M^{\#}(\underline{x},t)]\}\, dt \leq \int_{0}^{\infty} \{1-P[M(\underline{x},t)]\}\, dt$$

or

$$\mu^{\#} \leq \mu$$

where $\mu^{\#}$ and μ are the mean times to find the target, for $M^{\#}$ and M, respectively.

Therefore, the uniformly optimal search plan minimizes the expectation of T, and under these circumstances it is imaterial what M.O.E. is used in the problem.

80

EXAMPLE 13

EXPECTED TIME FOR THE UNIFORM TARGET'S LOCATION DISTRIBUTION

Suppose that the following conditions holds:

(a) The target's location is uniformly distributed inside an area A.

(b) The LDF is:

$$B(Z) = 1 - EXP [1WZ]$$

therefore, the CDF for T is:

$$F_T(t) = P(T \leq t) = P(t) = 1 - EXP [-\frac{WVt}{A}]$$

under these conditions, the expected time for finding the target is:

$$E[T] = \mu = \int_0^\infty [1 - P(t)[dt = \int_0^\infty EXP [-\frac{WVt}{A}]dt = \frac{A}{WV}$$

Applying this formula to chapter's I example, in the case $P_j = 1/4$ for all cells, and recalling that the total area has 6.000 square miles, the expected time is:

$$\mu = \frac{6.000}{15 \times 100} = 4 \text{ hours}$$

Observe that if in the CDF, t is replaced by μ, the probability of detection after the expected time is gone is:

$$P(\mu) = 1 - EXP [-\frac{WVA}{WVA}] = 1 - EXP [-1] = 0.633$$

The preceding computations were done under the assumption
that the search only stops if the target is located. However,
it is possible to compute the expected time to detect the
target, under the assumption that the search continues either
until the target is located, or until some pre-assigned value
of probability of detection, say P*, is reached. Let T*
be the time when:

$$P(t) = P*$$

from condition b)

$$P(T*) = P* = 1 - EXP\left[-\frac{WVT*}{A}\right]$$

and therefore

$$T* = \frac{A}{WV} \ln\left[\frac{1}{1-P*}\right]$$

Let θ be a random variable which describes the duration of
such a search. The C.D.F. for θ, can be derived considering
that:

$$P(\dot{\theta} \leq \theta \quad if \quad \theta < T*) = 1 - EXP\left[-\frac{WV\theta}{A}\right]$$

and

$$P(\dot{\theta} \leq \theta \quad if \quad \theta \geq T*) = 1$$

Then

$$F(\theta) = \begin{cases} 1 - EXP\left[-\dfrac{WV\theta}{A}\right] & \text{if } \theta < T^* \\ \\ 1 & \text{if } \theta \geq T^* \end{cases}$$

and the expected value of θ is:

$$E[\theta] = \int_0^\infty (1 - F(\theta))\, d\theta = \int_0^{T^*} EXP\left[-\frac{WV\theta}{A}\right] d\theta$$

$$+ \int_{T^*}^\infty (1-1)\, d\theta = \int_0^{T^*} EXP\left[-\frac{WV\theta}{A}\right] d\theta = \frac{P^*A}{WV}$$

Suppose that in the 6.000 square miles area of Heading' I example, a search is intended to be performed, which is to be stopped when a probability of detection P* is reached. In order to gain some insight in the order of magnitudes involved, T* and E[θ] will be computed for several values of P*, as it is shown in the table below:

P*	T* (Hours)	E[θ] (Hours)
0.500	2.772	2.000
0.600	3.665	2.400
0.700	4.811	2.800
0.800	6.437	3.200
0.900	9.210	3.600
0.910	9.624	3.640
0.920	10.100	3.680
0.930	10.612	3.720
0.940	11.230	3.760
0.950	11.980	3.800
0.960	12.872	3.840
0.970	14.020	3.880
0.980	15.648	3.920
0.990	18.420	3.960
0.992	19.312	3.968
0.994	20.444	3.976
0.996	22.084	3.984
0.998	24.856	3.992
1.000	T*→∞	4.000

It must be noticed the enormous disagreement between the time to stop and the expected time for the search, as P* is increased. Suppose that in a mine sweeping problem, the mines are randomly distributed in the area A and the sensor characteristics are, as before W = 15 mile and V = 100 knots. In such a context, P(t) can be interpreted

as the fraction of the total number of mines, which was detected by time t. The tremendous decreasing return's effect can be seen in the above table (column 1 and 2). For example, to increase the fraction of mines detected from 50% to 70%, the stopping time is increased in 2.039 hours, and to increase that fraction from 99.4% to 99.8%, the time increases in 4.412 hours.

Suppose now the same area A, sweep width W, and velocity V, as in example 13, i.e., 6000 square miles, 15 miles, 100 knots, respectively. However, the probability distribution is no longer uniform but given by

$$
P_j = \begin{cases} 0.\underline{4} & \text{if } j = I \\ 0.3 & \text{if } j = II \\ 0.2 & \text{if } j = III \\ 0.1 & \text{if } j = IV \end{cases}
$$

As in the last part of Heading's I example. Recall that a uniformly optimal search plan was derived there, which can be written as:

$$
M^{\#}(j,t) = \begin{cases} Z_I = \dfrac{Vt}{A_1} \\ \\ Z_{II} = Z_{III} = Z_{IV} = 0 \end{cases} \quad \text{IF } 0 < t \leq \theta_1
$$

85

$$M^{\#}(j,t) = \begin{cases} Z_I \doteq \dfrac{V(\Theta_1 + \frac{t}{2})}{A_1} \\[2em] Z_{II} = \dfrac{Vt}{2A_2} \qquad \text{IF } \Theta_1 < t \leq \Theta_1 + \Theta_2 \\[2em] Z_{III} = Z_{IV} = 0 \end{cases}$$

$$M^{\#}(j,t) = \begin{cases} Z_I = \dfrac{V(\Theta_1 + \frac{\Theta_2}{2} + \frac{t}{3})}{A_1} \\[2em] Z_{II} = \dfrac{V(\frac{\Theta_2}{2} + \frac{t}{3})}{A_2} \\[2em] \qquad\qquad \text{IF } \Theta_1 + \Theta_2 < t \leq \Theta_1 + \Theta_2 + \Theta_3 \\[2em] Z_{III} = \dfrac{Vt}{3A_3} \\[2em] Z_{IV} = 0 \end{cases}$$

$$M^{\#}(j,t) = \begin{cases} Z_I = \dfrac{V(\Theta_1 + \Theta_2 + \Theta_3 + \frac{t}{4})}{A} \\[2em] Z_{II} = \dfrac{V(\frac{\Theta_2}{2} + \frac{\Theta_3}{3} + \frac{t}{4})}{A_2} \\[2em] \qquad\qquad \text{IF } \Theta_1 + \Theta_2 + \Theta_3 < t \\[2em] Z_{III} = \dfrac{V(\frac{\Theta_3}{3} + \frac{t}{4})}{A_3} \\[2em] Z_{IV} = \dfrac{Vt}{4A_4} \end{cases}$$

86

Where

$$\theta_1 = 0.28770 \text{ hours}$$

$$\theta_2 = 0.8100 \text{ hours}$$

$$\theta_3 = 2.07945 \text{ hours}$$

The corresponding CDF is

$$P(t) = P_I \left(1 - \text{EXP}\left[-\frac{WVt}{A_1}\right]\right) \text{ if } 0 \le t < \theta_1$$

$$P(t) = \begin{cases} P_I \left(1 - \text{EXP}\left[-\frac{WV(\theta_1 + \frac{t}{2})}{A_1} + \right. \right. \\ \\ P_{II} \left(1 - \text{EXP}\left[-\frac{WVt}{2A_2}\right]\right) \end{cases} \text{ if } \theta_1 \le t < \theta_1 + \theta_2$$

$$P(t) = \begin{cases} P_I \left(1 - \text{EXP}\left[-\frac{WV(\theta_1 + \frac{\theta_2}{2} + \frac{t}{3})}{A_1}\right]\right) + \\ \\ P_{II} \left(1 - \text{EXP}\left[-\frac{WV(\frac{\theta_2}{2} + \frac{t}{3})}{A_2}\right]\right) + \\ \\ P_{III} \left(1 - \text{EXP}\left[-\frac{WVt}{3A_3}\right]\right) \end{cases} \text{ if } \theta_1 + \theta_2 \le t < \theta_1 + \theta_2 + \theta_3$$

$$P(t) = \begin{cases} P_I \left(1 - \text{EXP}\left[-\frac{WV(\theta_1 + \frac{\theta_2}{2} + \frac{\theta_3}{3} + \frac{t}{4})}{A_1} + \right. \right. \\ \\ P_{II} \left(1 - \text{EXP}\left[1 \frac{WV(\theta_{2/2} + \theta_{3/3} + t/4)}{}\right]\right) + \\ \\ P_{III} \left(1 - \text{EXP}\left[-\frac{WV(\theta_{3/3} + t/4)}{A_3}\right]\right) + \\ \\ P_{IV} \left(1 - \text{EXP}\left[-\frac{WVt}{4A_4}\right]\right) \end{cases} \text{ if } \theta_1 + \theta_2 + \theta_3 \le t$$

87

or replacing the values for A_j, Θ_j, W, V, and P_j:

$$0.\underline{4} \ (1 - \text{EXP} \ [1t] \quad 0 \leq t \leq 0.28770$$

$$P(t) = \begin{cases} \begin{cases} 0.4 \ (1 - \text{EXP} \ [-0.28770 + \frac{t}{2}] \ + \\ \\ 0.3 \ (1 - \text{EXP} \ [- \frac{t}{2})) \end{cases} \quad 0.28770 < t \leq 1.0977 \\ \\ \begin{cases} 0.\underline{4} \ (1 - \text{EXP} \ [-0.6927 + \frac{t}{3}]) \ + \\ \\ 0.3 \ (1 - \text{EXP} \ [-0.405 + \frac{t}{3}]) \ + \quad 1.0977 < t \leq 3.17715 \\ \\ 0.2 \ (1 - \text{EXP}[- \frac{t}{3}]) \end{cases} \\ \\ \begin{cases} 0.\underline{4} \ (1 - \text{EXP} \ [-1.3927 + \frac{t}{\underline{4}}]) \ + \\ \\ 0.3 \ (1 - \text{EXP} \ [-1.105 + \frac{t}{\underline{4}}]) \ + \\ \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad 3.77715 < t \\ 0.2 \ (1 - \text{EXP} \ [-0.7 + \frac{t}{\underline{4}}]) \ + \\ \\ 0.1 \ (1 - \text{EXP} \ [- \frac{t}{4}]) \end{cases} \end{cases}$$

As was mentioned before, in general:

$$\mu = \int_0^\infty [1 - P(t)] \ dt$$

In this case, however, due to the form of P(t), this expression becomes:

$$\mu = \int_0^{\Theta_1} [1 - P(t)] \ dt + \int_{\Theta_1}^{\Theta_1 + \Theta_2} [1 - P(t)] \ dt +$$

$$\int_{\Theta_1 + \Theta_2}^{\Theta_1 + \Theta_2 + \Theta_3} [1 - P(t)] \ dt + \int_{\Theta_1 + \Theta_2 + \Theta_3}^\infty [1 - P(t)] \ dt$$

88

or

$$I_1 = \int_0^{0.2877} 1-0.8[1-EXP(-t)]dt = 0.27342 \text{ hours}$$

$$I_2 = \int_{0.2877}^{1.0977} 1-0.\underline{4}[1-EXP(-0.2877-\frac{t}{2}]-0.3[1-EXP(-\frac{t}{2})] \, dt = 0.59\underline{4}1216$$

$$I_3 = \int_{1.0977}^{3.17715} 1-0.\underline{4}[1-EXP(-0.6927-\frac{t}{3})]-0.3[1-EXP(0.\underline{4}05-\frac{t}{3})] -$$

$$0.2[1-EXP(-\frac{t}{3})] \, dt = 0.5972783 \text{ hours}$$

$$I_{\underline{4}} = \int_{3.17715}^{\infty} 1-0.\underline{4}[1-EXP(-1.3927-\frac{t}{\underline{4}})] -0.3[1-EXP(-1.105 - \frac{t}{\underline{4}})] -$$

$$- 0.2[1-EXP(-0.7 - \frac{t}{\underline{4}})] - 0.1 [1-EXP(- \frac{t}{\underline{4}})]$$

$$= 0.5270036 \text{ hours}$$

Then

$$\mu = I_1 + I_2 + I_3 + I_{\underline{4}} = 1.9918235 \text{ hours}$$

89

The computation above is very cumbersome, and, in fact, for a larger number of cells it may be impracticable. In real life situations however, it is possible to program a computer to do the work. Anyway, the knowledge of the expected time to find the target, is very important in search planning, and this example shows how to do it for a grid type distribution.

If, as it was done in the preceding example, the search is to be stopped if the target is not found by time T*, when P(T*) = P*, a pre-assigned value, then, due to the form of P(t), becomes much more difficult from the computational point of view, to determine the values of T* and E[θ]. The general approach to the problem is: (a) given P*, determine in what interval T* is. Invert the corresponding expression for P(t) and obtain T*; (b) compute $\int_0^{T*}$ [1-P(t)] dt = E[θ]

As example, suppose first P* = 0.06 considering that P(0.2877) = 0.098; then T* belongs to the first interval, therefore:

$$T^* = - \ln [1- \frac{0.06}{0.4}] = 0.16252$$

or in general:

$$T^* = \frac{A_k}{VW} [\ln (1- \frac{P^*}{P_k})]^*$$

because in this case, due to the numbers which are being used,

$$\frac{A_j}{VW} = 1$$

for any j.

Then

$$E(\theta) = \int_0^{T^*} [1-P(t)]\, dt$$

or in the example

$$E(\theta) = \int_0^{0.16252} (1-0.\underline{4}\,[1 - EXP\,(-t)])\, dt = 0.157 \text{ hours}$$

However, if P* is chosen to be equal to 0.25, things are not so easy. Considering that

$$P(1.097) = 0.298$$

then T* is located in the second interval, and the equation to be solved is:

$$P(x) = 0.\underline{4}\,[1 - EXP\,(-0.28770 - \tfrac{x}{2})] + 0.3\,[1 - EXP(-\tfrac{x}{2})]$$

and

$$T^* = 0.2877 + x$$

which only can be solved by numerical methods. After the value of T* is known, the expected time is:

$$E(\theta) = \int_0^{0.28770} \{1-0.\underline{4}\,[1 - EXP(-x)]\}dx +$$

$$\int^{T^*} \{1-0.\underline{4}\,[1-EXP(-0.2877 - \tfrac{x}{2})] - $$

$$0.28 \quad \}$$

$$- 0.3\,[1-EXP\,(-\tfrac{x}{2})]\}\, dx$$

The above calculations, could be done on a computer.

EXAMPLE 15

EXPECTED TIME FOR FINDING A TARGET WHOSE POSITION HAS A CIRCULAR NORMAL DISTRIBUTION

In example 12, it was derived an $M^{\#}$ plan for this case. The probability of detection at or before time t under this plan is:

$$P(t) = 1 - (1 + K \sqrt{t}) \, EXP \, [- k \sqrt{t}]$$

whose derivation can be found in [2]. The expected time is then:

$$\int_0^\infty [1 - P(t)]dt = \int_0^\infty [1 + K \sqrt{t}] \, EXP \, [- k \sqrt{t}]dt$$

$$= \frac{\sigma \pi \sigma^2}{WV}$$

Again, fixing a probability P*, and stopping the search either if the target is found or if a time T* such that P(T*) = P* is reached, leads to another way of defining the expected time. In this case:

$$P* = P(T*) = 1 - (1 + k \sqrt{T*}) \, EXP \, [- k \sqrt{T*}]$$

equation which can be solved for T* by numerical methods. As before, the expected time for a search under these conditions is:

$$E(\theta) = \int_0^\infty [1 - F(\theta)] \, d\theta =$$

$$= \int_0^{T*} \{(1 + K\sqrt{t}) \, EXP \, [-K\sqrt{t}]\} \, dt =$$

$$= T* + \frac{8}{K^2} - 2KEXP(-K\sqrt{T*}) \, \{\frac{T*}{2K} + \frac{\sqrt{T*}}{K^2} + \frac{2}{K^3}\}$$

where θ, $F(\theta)$ and $E(\theta)$, have the same meaning as in example 13.

J. FALSE TARGETS

Suppose a detection device, a radar for example, and consider the problem of setting it in sensitivity control. The more sensitive the instrument is made the less likely it will miss a target. Then, for any given distance the instantaneous probability of detecting the target, given it is there, is an increasing function of the radar's sensitivity. On the other hand, as the sensitivity is made higher, the probability of a false alarm is also increased. For with higher sensitivity more noise is presented in the radar's scope, and since the operator is unable to differentiate between signal and noise, he will claim more frequently the presence of a target, where in fact what he detected was noise.

This is one of the several situations that arises during a search when it is necessary to distinguish between the target and false alarms. A member of the set composed of targets and false targets in called a "contact." Once there is a contact in the sensor's display, it takes some time to identify it as a target or as a false target, and even then the identification may be incorrect.

From the above considerations, it can be concluded that the presence of false targets should affect the search plan. In fact, under these circumstances, the search takes place in two phases: The first, called the broad search, is conducted using a sensor which is able to detect the contacts, but can not identify them. The second, the contact identification phase, is conducted by another type of sensor (in most cases, the human eye), which identifies the contacts.

At this point, it is necessary to make the following assumptions in order to build a model to handle the problem:

(a) The search area is divided into J cells, each one with area A_j, and having a false target density δ_j. The actual number N_j of false targets in A_j is Poisson distributed:

$$P(N_j = n) = \frac{(A_j \delta_j)^n EXP(-A_j \delta_j)}{n!}$$

Let T_j stand for mean time to identify a contact in cell j.

(b) As soon as a contact is detected, the broad search stops, and the contact is investigated until positive identification. If it is identified as a false target, the broad search starts again.

Under the preceding assumptions, it can be demonstrated, that it is not possible to find a uniformly optimal search plan. However, it is possible to devise a plan which minimizes the expected time to find the target (observe that in this case the two M.O.E. do not coincide).

The following results will be stated without proof
(complete discussion can be found in [2]):

(R.1) The return function, for this case is given by:

$$\lambda_j(Z) = \frac{P_j \, B_j'(Z)}{A_j[1 + \delta_j T_j B_j'(Z)]}$$

(R.2) The best search plan, M^*, is the one which allocates
the next increment of effort to the cell (or cells) where $\lambda(Z)$
is higher.

(R.3) The probability of finding the target is still:

$$P(Z) = \sum_{j=1}^{J} P_j \, B_j \, (Z_j)$$

However, when computing the probability at or before time T,
it is to be understood that this time refers to broad search
time, not including therefore the time spent in contact
investigation.

(R.4) The cost of an allocation is in this case.

$$C(Z) = \sum_{j=1}^{J} A_j \, Z_j + A_j \, \delta_j \, T_j \, B_j \, (Z_j)$$

Where Z_j refers to the amount of broad search effort allocated
to cell j. For sake of clarification, suppose that Z_j stands
for the time per unit area spent at cell j, then:

$$C(Z) = \sum_{j=1}^{J} A_j \, t_j + A_j \, \delta_j \, T_j \, B_j \, (t_j)$$

At this point, an example seems to be helpful.

EXAMPLE 16

Suppose that a helicopter is searching for a trawler X. whose location distribution is pictured in Figure 15 below:
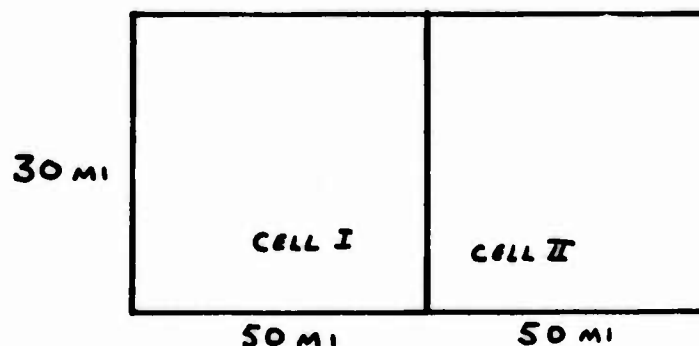

Figure II-15

Suppose further that there are several other trawlers in both areas, and the helicopter must approach each one it contacts in order to verify if she is the trawler X. If it is assumed that the trawlers are uniformly distributed in each area, the model presented applies, playing X the role of target, and all other trawlers the role of false targets.

Let the LDF·be:

$$B(Z) = B_j \left(\frac{t_j}{A_j}\right) = 1 - EXP \left[- \frac{t_j VW}{A_j}\right]$$

where t_j is the amount of time spent in cell j.
And the other parameters:

$$T_j = 12 \text{ minutes} = 0.2 \text{ hours}$$
$$V = 100 \text{ knots}$$
$$W = 15 \text{ miles}$$

And due to past experience in that fishing area, it is estimated that there are about 3,000 trawlers in cell I, and 75 in cell II then:

$$\delta_I = 2/sq. \; mi.$$
$$\delta_{II} = 0.05/ \; sq. \; mi.$$

In order to know where to start, it is necessary to compute the return function for $Z = 0$. In this case:

$$\lambda(\frac{t_j}{A_j}) = \frac{P_j \; WV \; EXP \; (- \frac{WVt_j}{A_j})}{A_j[1 + \delta_j T_j \; WV \; EXP \; (- \frac{WVt_j}{A_j})]}$$

and

$$\lambda_I(0) = \frac{0.55 \; X \; 1500}{1500 \; [1 + (0.2) \; (2) \; (1500)]} = \frac{0.55}{601} = 9 \; X \; 10^{-4} \; hours^{-1}$$

$$\lambda_{II}(0) = \frac{0.45 \; X \; 1500}{1500 \; [1 + (0.2) \; (0.05) \; (1500)]} = \frac{0.45}{151} = 3 \; X \; 10^{-2} \; hours^{-1}$$

Therefore, the plan starts in cell II, the one with smaller prior probability. This happens because of the big difference in false targets' density, over weighted the difference in probability. The effort will be placed only in cell II, until a time, Θ such that:

$$R_{II}(\Theta) = R_I(0)$$

or

$$\frac{0.45 \; EXP \; [-\Theta]}{1 + 15 \; EXP \; [-\Theta]} = 9 \; X \; 10^{-4}$$

97

Which solved for θ, gives

$$\theta = 3.9 \text{ hours}$$

after this time, the effort must be split evenly among the cells.

K. SEARCH PLANNING

The following are basic topics in search planning:

(a) Determination of the initial target location probability distribution.

(b) Evaluation of the search system capabilities

(c) The decision to begin search

(d) The decision to terminate search

Before examining each one of the mentioned topics, it should be noticed that the above list is by no means exhaustive. In fact, it is very likely that for each situation a multitude of other factors must be considered. However, as was mentioned in the very first sentence of this section, these are basic topics in the sense that they are worthy of consideration under any circumstances, and also that, the decisions concerning them have a strong influence in the discussions about any other factors.

In the determination of a prior distribution for the targets position, it is essential to keep in mind that although sub-jective judgement play a central role in this matter, this does not mean that it is just a guessing process. On the contrary,

by observing some rules in analyzing all the relevant information about the target, it is possible to obtain a distribution which from the viewpoint of the search model, is accurate enough.

In general, it is possible to think of two basic situations at the starting point of a search operation. Either no relevant information about the target's last position is available, or a datum is somehow known. In the first case, the uniform distribution is the appropriate one. The problem is then to determine the position and the dimensions of the search area. Usually, in this case a rectangle is constructed and adopted as S.A., for the sake of simplicity. The location and dimensions of this rectangle will rely strongly on the analysis of the available information, and mainly on expert's opinions. Before seeing an example, it is necessary to recall that the mentioned rectangle must contain the target with probability 1, and therefore, the less information available, the larger the S.A. will be, and consequently more effort must be placed there, in order to assure a determinated level of probability of detection.

EXAMPLE 1

Consider Figure I, where "A" and "B" are airports, and suppose that an airplane left "A" and did not arrive at "B". Its last position is unknown (from a very rigorous point of view, "A" is the aircraft last position, however, for practical purposes, this kind of reasoning is useless).
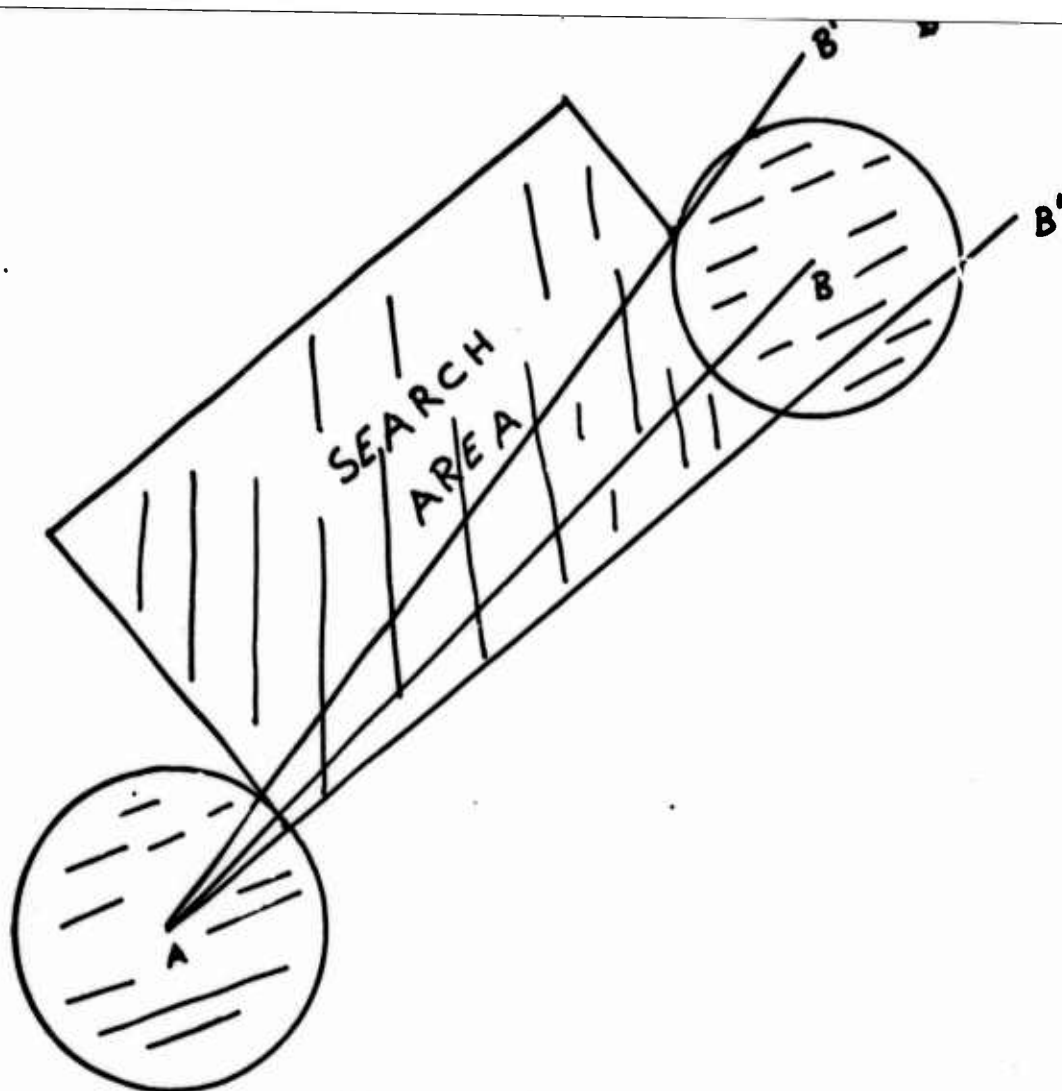
Figure II-16

Flight experts are asked to inform what is the maximum error of navigation for the aircraft in question. Meteorology people give the direction and intensity of winds in the area during the time of flight. From these considerations it is concluded that AB' and AB" are the extreme paths for the airplane. Further, the dashed regions centered at "A" and "B" are the approach control zones of the airport, and the plane has been seen leaving the "A" zone, and did not enter the "B" zone.

Flight control people assure that it is impossible to cross enter zone without being detected. Finally, considering that between "A" and "B", there is no place for an emergency landing, the conclusion is that the plane crashed somewhere, and again, by consulting flight experts, it is determined the maximum distance of landing from each point of the extreme courses, under the prevailing conditions of wind. The search area of Figure 16 is then a result of all the above processes. It should be noticed that at this point, all the work of the search planner is to obtain from the experts the necessary amount of information in order to draw the minimum search area which contains the target with probability one.

When a last position is known, the next question concerns how it was determined. The reasons is that there is always an error in navigation and the probability distribution for the target's location will reflect this error. In fact, when a last position is known, the target's position distribution can be viewed as the error in navigation distribution. There-fore, it is by the knowledge of how the mentioned last position was obtained that it is possible to obtain the form and the parameters of a corresponding probability distribution. In fact, for a large class of situations, a huge mass of data from past experience allows a search planner to obtain very accurate results in fitting a distribution to a specified situation. Reference [3] contains tables and curves that can be used in many cases, and which are results taken from past experience. Perhaps these distribution fitting procedures can be made more clear with the next examples.

101

Suppose a small boat asks for help, and gives her actual position. Evidently there is an error in this position, and this error is assumed normal. Reference [3] tabulates, in its chapter 6, a "total error in position," denoted "c", which stands for the radius of the circle which, with probability 0.5, contains the target. Assuming that the position was based in dead reckoning, four hours after last fix and that the average speed for that type of boat is 10 knots, the value for the "total error in position" is:

$$C = 46 \text{ miles}$$

From the properties of the bivariate normal distribution, already discussed in II.

$$\sigma_{X_1} = \sigma_{X_2} = \frac{C}{1.18}$$

or, in this case:

$$\sigma_{X_1} = \sigma_{X_2} = \frac{46}{1.18} = 38.9 \text{ miles}$$

Therefore, under the assumption of normality, if the search area is centered in the last position, the target's location distribution is a circular normal, with mean zero and standard deviation 38.9 miles. A search grid can be constructed if necessary, by the use of the results presented in II

It is interesting to mention, that the drift effect can be computed in cases like this, by making

$$C = \sqrt{d^2 + \lambda^2}$$

102

Where X is the former value for c (46 miles in this example), and d is a "drift error" which can be found, for many types of targets and meteorological and oceanographic conditions, in chapter 6 of reference [3].

Observe, however, that a drifting target is in fact a moving target, and that the described procedure is a practical approximation that allows the search planner to deal with a moving target in the same way as with a stationary one, just by increasing the variance of its location distribution. This approximation is only valid within the limits of the small velocity that characterizes drifting.

Finally, the assumption of normality comes from the generalized notion that errors in navigation are in fact sums of many errors (instrument reading errors, plotting errors, etc.), and thus, due to the central limit theorem (see reference [4], it is reasonable to talk of normality.

EXAMPLE 3

THE SEARCH FOR THE SSN THREASHER

This is a real life example, which is discussed in full detail in reference [7]. Also, the interested reader can find important material about under water search in reference [8].

The Thresher was lost in approximately 8,400 feet of water during her sea trials, after an overhaul, near the gulf of Maine. In establishing her location distribution, the analysts in charge, considered the following points:

(1) Last Known Position

At 0917 on April 10, 1967, the USS Skylark received a UQC (underwater telephone) message from Thresher reporting that she was in trouble for maintaining her trim. A few seconds after, Skylark heard breakup noises over the UQC. At this same time, Skylark determined her own position by using a loran alfa. Then, at this point, the following are the relevant facts.

_____ The Skylark position, determined by the loran alfa, which has a "total error in position" (see last example), of 1.5 miles.

_____ The nominal range of UQC, which is 5,000 yards. Therefore, by the same reasons discussed in example 2, the Skylarks' position can be described by a circular normal distribution centered at the loran' fix, and with standard deviation given by:

$$\sigma_{X_1} = \sigma_{X_2} = \frac{1.5}{1.18}$$

On the other hand, by the knowledge of the UQC range, it was assumed that the Thresher's position had a uniform distribution inside of a circle whose radius is 5.000 yards, and centered in each possible Skylark's position. This, leads to a compound distribution, as discussed in II-F for, in order to compute the probability in any cell in the search space, it is necessary to deal with a uniform density conditioned by a normal one.

104

(2) Horizontal and Vertical Motion Before and After
the Last Known Position

Although the various courses steered by the Thresher
before the accident was known, no factual information con-
cerning speeds before, and courses and speeds after the
accident were available. Also nothing was known for sure
about vertical motion. However, by using experts' opinions,
a maximum range from the last position, that the submarine
could be carried to was estimated. Notice that this maximum
range also condition the probability distribution.

(3) Oil Slick

A diesel oil slick was noticed nearby the Skylark's
position. However, noreliable information could be extracted
from the experts, because of the great uncertainty about
winds, currents, etc., in the area.

By using the information contained in (1) and (2), and
disregarding (3), the search grid was constructed as follows:

I. A search area of 10 X 10 miles was established. This
was done by computing the size of the 0.9 probability square,
for the compound distribution (see II-F example, for the case
of a circular normal), and adding the range predicted in (2),
by the experts.

II. The search grid was laid out by dividing the search
area in squares, each one mile on a side, and computing the
respective probabilities. Finally, the following observations
concerning to the construction of search grids are of interest:

105

(1)   Whenever possible the coordinate system for
the grid should be the north/south system of the mercator
chart for the are in question, and the units of measure the
same, i.e., for the x_1 axis longitude, for the x_2 axis latitude.

(2)   The choice of the dimension for each cell is
dictated by the ratio between the standard deviation of the
navigation error (σn) and the sweep width (w). This is due to
the fact that if $\sigma N/W$ is large, then it is very likely that
appreciable amounts of effort intended for one cell will be
placed outside of that cell. Reference [7] has a nomograph
that can be used for determining the correct dimensions.
Recall that values of σn, for several cases, can be obtained
by dividing the "C" factors tabulated in references [3] and
[7] by 1.18. In reference [3], "C" is called "total error in
position," and in reference [7], "circular probable error."

(3)   Although the calculations of each cell's
probability presented in II-F might seem, at first sight, very
straightforward, in fact they are usually not. For example,
in the search for the Scorpion (see reference [8], nine
scenarios were considered. Besides the obvious volume of
calculations, it is not common that the distribution of a
random variable, conditioned by many others, can be written
in a closed analytical form. In this case, the correct approach
is the use of a Monte Carlo method. A computer simulation for
generating a search grid is described in reference [7]. A

106

complete list of the program, which is written in FORTRAN IV,
and that can be used in a wide range of situations, can be
found in Appendix D, of that reference.

Item B

After the target's location probability distribution is
known, the next step for the search planner is to evaluate what
can be done with the resources he has. This is a very important
step, because it may be that he does not have means enough to
accomplish his mission. Suppose, for instance, that under the
situation described in I-A example, we want a search with
probability 0.95 of finding the target. As can be recalled,
this is impossible with the resources available, which did not
allow more than 0.45.

Although the very first consideration in evaluating a
sensor system, should be with the LDF to be employed, the
random search formula, being a conservative estimate of search
efficiency, should be employed.

Determination of the sweep width then becomes the first
step. The best way to obtain a value for a sensor's sweep
width is to integrate its lateral range function, as described
in II-E. However, in many cases, the lateral range curve is
not available and the following are special guidelines to
deal with such situations;

(1) For visual search, use the numbers of the table in
Figure 7-2 reference [3].

(2) For electronic devices, the following procedures,
listed in order of preference, are recommended:

107

I. When the minimum detection range is known:

$$W = (1.5) \text{ (minimum detection range)}$$

II. When average detection range is known:

$$W = (1.2) \text{ (average detection range)}$$

III. When maximum detection range is known:

$$W = (0.8) \text{ (maximum detection range)}$$

IV. When no detection range is known:

$$W = (0.4) \text{ (horizon range)}$$

Notice, however, that the above rules of thumb should be used carefully, and their results always checked by common sense. For instance, suppose that a search planner wants to estimate the sweep width for an airplane radar, operating at 3,000 feet of altitude in searching for a small boat. By rule IV, the sweep width is 25.6 miles, which, for a small boat is exaggerated.

The visual sweep width for the best meteorological conditions for an airplane at 3,000 feet, searching for a boat of less than 30 feet long is 7.3 miles (see Table 7-2 in reference [3]), and this is a good estimate of the radar sweep width in this case. However, if instead of a small boat, the plane were searching for a 10,000 ton ship, then the 25.6 miles of rule IV would be a good guess.

Anyway, reference [3] should be consulted whenever such rules are to be used.

(3) For the special case of underwater search, reference [7] states very interesting procedures for estimating sweep widths of sensors such as underwater cameras, magnetometers, sonars and so on. It should be mentioned that when performing an underwater search, the sweep width is often treated as a random variable. For instance, consider the simple case where for a given sensor the sweep width is w_1 if the target is just seated in the bottom, or w_2 if the target is buried in the mud. If p stands for the probability of being seated in the bottom, 1-p is the probability of being buried in the mud. The probability of detection is:

$$P(Z) = p(1-EXP[1w_1 Z]) + (1-p)(1-EXP(-w_2 Z))$$

observe that if the mean value

$$\bar{w} = pw_1 + (1-p) w_2$$

is used to compute P(Z), the expression

$$P(Z) = 1-EXP(-\bar{w}Z)$$

Leads to a different result than (*), which, as can be proved, is the correct one. Usually, in underwater search, random sweep widths are approximated by gamma distributions. A complete discussion can be found in reference [2].

A point of interest is the evaluation of the overall
sweep width when more than one sensor is looking for the same
target. In order to derive such a result, ler w_j be the sweep
width of the jth sensor, and M_j its maximum range. Clearly,
the ratio $\beta_j = \frac{W_j}{M_j}$
can be thought as a measure of the probability of detection
inside that maximum range. If n sensors are used in a group,
then the combined sweep width is given by:

$$W = W_1 + W_2 (1-\beta_1) + \cdots W_N(1-\beta_1)(1-\beta_2) \cdots (1-\beta_{N-1})$$

where $W_? \geq W_3 \geq \cdots \geq W_N$

A proof for this statement can be found in reference [7].
Another relevant point in determining sweep width has to do
with the targets' physical dimensions. This is an entry for
many tables (table 7-2, reference [3], for example, divides
the targets in seven categories, ranging from life rafts to
ships over 10,000 tons of displacement). In the Thresher's
case, the sweep width was det ermined assuming a 268 by 300
foot target which were the submarine's dimensions. However,
after the search, it was concluded that due to the impact on
the bottom and the consequent breakup, the light debris
(twisted metal, etc.) scattered by the submarine, together
with main parts of her hull, were in fact a target of 1,000
X 5,000 yards. Therefore, the sweep width adopted was smaller
than the real one. Reference [7], in its section 4.8, discusses
in detail this point and tabulates some of the results. Finally,
an approximated expected time for the search must be computed.

110

The word "approximate" was used because, at this point, the search planner does not know yet what search plan he will employ. In fact, the choice of a search plan will depend on this first rough estimate of search time. Anyway, by using the methods and formulas presented in II-I, it is possible to obtain a reasonable approximation to the expected search time. Notice that the possibility of false targets must be carefully considered at this point; if there is a considerable density of false targets in the area, the broad search time can be a very misleading measure of time to be spent.

ITEM C

Suppose that a search operation costs an amount θ per unit of time, and that the value of the target is Δ. If T^* is the expected time to find the target, then θ^*, the expected cost is:

$$\theta^* = \theta T^*$$

Therefore, in order to make a "profit", the search must begin only if:

$$\Delta > \theta T^*$$

Observe that T^* can be computed as was done in II-I, either under the assumption that the search continues until the target is found, or until a certain level of probability is reached. In the latter case, θ^* will stand for the cost to reach that amount of probability. At this point, an example seems helpful.

EXAMPLE 4

Suppose that after an exercise, a torpedo sank. If its value is 10,000 dollars, the cost for searching its 4,000 dollars/hour, and the available sensor has a speed of 4 knots and a sweep width of 0.5 miles, should a search planner order the search to begin if:

(a) the torpedo's location is uniformly distributed inside a 2 X 2 mile search area.

(b) the torpedo's location has a circular normal distribution with σ = 1 mile? For case (a), the expected time is:

$$T^* = \frac{A}{WV} = \frac{4}{2} = 2 \text{ hours}$$

therefore

$$\theta^* = 4.000 \text{ X } 2 = 8.000 \text{ dollars}$$

and

$$V = 10.000 > \theta^* = 8.000$$

Then the correct decision is to start the search. For case (b) the expected time is:

$$T = \frac{2\pi\sigma}{WV} = 3.14 \text{ hours}$$

Therefore

$$\theta^* = 4.000 \text{ X } 3.14 = 12560 \text{ dollars}$$

and

$$V = 10.000 < \theta^* = 12560$$

So the best decision is to abandon the torpedo.

112

In real life situations, the big problem is to assign values for θ and Δ. However, the above guidelines are useful whenever such a decision must be made.

ITEM D

The reasons for terminating search can be viewed in the same context as the reasons for beginning search.

In fact, if the first increment of effort is spent because the expected return exceeds the expected cost, then the logical time for concluding the search is when the expected return of additional increment of effort is smaller than its expected cost. Again, as was pointed out in item (C), the big problem in real life is to estimate value and costs. However, in the stopping case, it is possible to argue, aft er the probability of detection has r eached a huge value and the target still has not been found, that there is statistical evidence to indicate that some assumption is incorrect and therefore to justify termination of the search, or at least to reevaluate the initial assumptions before search continues. Using the same notation as in item (c), and for the case of the circular normal prior, search must stop by time given by:

$$t = \frac{(\frac{3\Delta}{\theta T^*} - 1)^2 T^*}{\sigma}$$

For the case of the uniform prior, where the returns are constant, the rule is to stop when the probability of detection reaches some fixed value, say 0.95.

LIST OF REFERENCES

1. Operations Evaluation Group Report No. 56, <u>Search and Screening</u>, by B. O. Koopman, 1946.

2. Stone, L. D., <u>Notes on Search Theory</u>, presented at the U.S. Naval Postgraduate School during the winter quarter of 1973.

3. <u>National Search and Rescue Manual</u>, C6-308, United States Coast Guard.

4. Zehna, P. W., <u>Probability Distributions and Statistics</u>, Allyn and Bacon, 1970.

5. Pollock, S. M., <u>Selected Methods and Models in Military Operations</u>.

6. <u>Parzen, E., Stochastic Processes</u>, Holden Day, 1962.

7. Stone, L. D., and Richards, H. R., <u>Manual for the Operations Analysis of Deep Ocean Search</u>, Daniel Wagner Associates, 1971.

8. Richardson, H. R., and Stone, L. D., <u>Operations Analysis During the Underwater Search for Scorpion</u>. Daniel Wagner Associates, 1968.

9. Hadley, G., <u>Non Linear and Dynamic Programming</u>, Addison-Wesley, 1964.

10. Feller, W., <u>An Introduction to Probability Theory and its Applications</u>.