OPERATIONS RESEARCH
Vol. 54, No. 6, November–December 2006, pp. 1193–1200
ISSN 0030-364X | EISSN 1526-5463 | 06 | 5406 | 1193

DOI 10.1287/opre.1060.0332

Piled-Slab Searches

Alan Washburn
Operations Research Department, Naval Postgraduate School, 1411 Cunningham Road, Monterey, California 93942, awashburn@nps.edu

This paper deals with the conflict between simplicity and optimality in searching for a stationary target whose location is distributed in two dimensions, thus continuing an analysis that was begun in World War II. The search is assumed to be of the "piled-slab" type, where each slab consists of a uniform search of some simple region. The measure of simplicity is the number of regions (smaller is simpler). If each of a fixed number of elliptical regions is searched randomly, we find the optimal region size and the optimal division of effort between regions. Rectangular regions are also considered, as are problems where the regional searches are according to the inverse-cube law, instead of random search. There is a strong tendency for optimal inverse-cube law searches to consist of a single slab. We also consider problems where the amount of effort for each region is optimized myopically, with no consideration for the search of future regions.

Subject classifications: search and surveillance; probability applications.
Area of review: Military.
History: Received November 2004; revision received May 2005; accepted August 2005.

1. Introduction

There is sometimes a conflict between efficiency and simplicity when searching for a stationary lost object. Simplicity demands a search pattern that treats all parts of the searched area equally, whereas efficiency may require an uneven distribution of effort. In cases where the prior probability distribution of the target's location is unimodal about some "fix" or "datum," and where the environment is homogeneous, both goals can be approximately achieved by searching uniformly over an increasing sequence of simple regions $R_1, R_2, \ldots, R_n$, where the datum is in R_1 and each region contains its predecessor. In this way, points near the datum are searched more intensely than points farther away, even though all searches are uniform. Koopman (1946), for example, records the details of such a "piled-slab" scheme for finding a target that is lost according to a bivariate normal distribution. Our intention here is to optimize and generalize that analysis.

Another approach to simplicity is to require that all searches be in rectangles that do *not* overlap. Such an approach makes sense for prior distributions of target location that are more or less uniform over mutually exclusive regions. Dicenza (1980) describes an integer programming method for determining a collection of such rectangles that maximizes detection probability. Because our focus here is on priors that are unimodal, we will concentrate on nested regions, rather than nonoverlapping regions.

Random search is a special case with an important property that is not shared by the general search problem. Random search is covered in §2, while the general case is covered in §3. It sometimes happens that a search is completed without finding the target, but then extended in time for reasons that have emerged since the search began. The implied mode of optimization is myopic, where each slab is optimized without regard to the possibility of others in the future. A myopic plan will not in general be globally optimal, although the defect from optimality is usually slight. Myopic optimality may require "inversions" where a given search covers a smaller region than one completed earlier. This subject is taken up in §4. Section 5 contains the summary.

Computations are illustrated in the Excel workbook *PiledSlab.xls*, which can be downloaded from http://diana.cs.nps.navy.mil/∼arwashburn/.

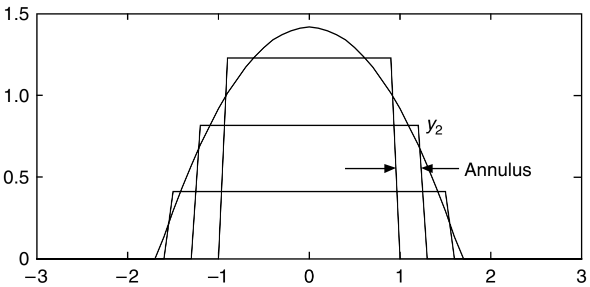
2. The Special Case of Random Search

Search of any region will be most efficient when the examined parts cover the entire region without overlap, but this kind of "exhaustive" search is seldom achieved in practice on account of unintended overlap among the examined parts. Random search is a limiting case where the examined parts are assumed to overlap at random more or less as they do in the fall of confetti. If an amount of covered area z (the amount of "effort") is cut into confetti and strewn uniformly at random over an area A, the fraction of A covered by confetti will be $1-\exp(-z/A)$ (Koopman 1956, pp. 519–521). The same random search formula holds even if the confetti is not uniformly distributed, as long as the ratio z/A is interpreted as the effort density in the vicinity of the target. The effort z is often assumed to be WL, where L is track length and W is sweep width, but this assumption is not required. Random search is often assumed to be a skeptical and practical approximation to what happens in reality, given the difficulties of navigation, the tendency of target locations to drift slightly as search proceeds, and the often indefinite nature of detection ranges. The important characteristic that distinguishes random search from the general case is that two independent random searches with effort amounts z_1 and z_2 are exactly equivalent to a single random search with total effort $z_1 + z_2$. In other words, the total amount of effort can be partitioned arbitrarily without changing the overall result.

Even though search is conceptually reduced to confetti casting, there is still the strategic question of how the density of confetti should be distributed over space. The optimal distribution is obvious if the prior distribution of the target's location (hereafter the "prior") is uniform within some area A, but it is not obvious under the more practical assumption that the prior is elliptical normal. Koopman (1980, pp. 91–137) describes circumstances where this should be the case. The optimal distribution of effort for an elliptical normal prior was worked out in World War II by Koopman and others "too numerous to list" (Koopman 1946, p. 1 or Koopman 1980, pp. 157–160). The resulting search density has the shape of an inverted cup, being constant on equiprobability ellipses and decreasing continuously as one moves away from the origin (Figure 1). Because such a search is both strategically optimal and locally random, we will refer to it as SOLR. In this section, we will maintain the same constraint on total search effort, but insist that n nested regions of constant density be employed, thus approximating the inverted cup with a solid composed of n piled slabs. The resulting detection probability will, of course, be smaller than the SOLR probability.

Let $A_0 = 0$, and for $i > 0$, let A_i be the area of region R_i. We assume in this section that there is a one-to-one map between regions and areas, so that specification of A_i determines R_i, as well as vice versa. Let y_1 be the total effort density in R_1, and for $i > 1$, let y_i be the total effort density in the annulus between R_{i-1} and R_i (Figure 2). The total

Figure 2. A plot of search effort density vs. distance from the datum in a radially symmetric problem.

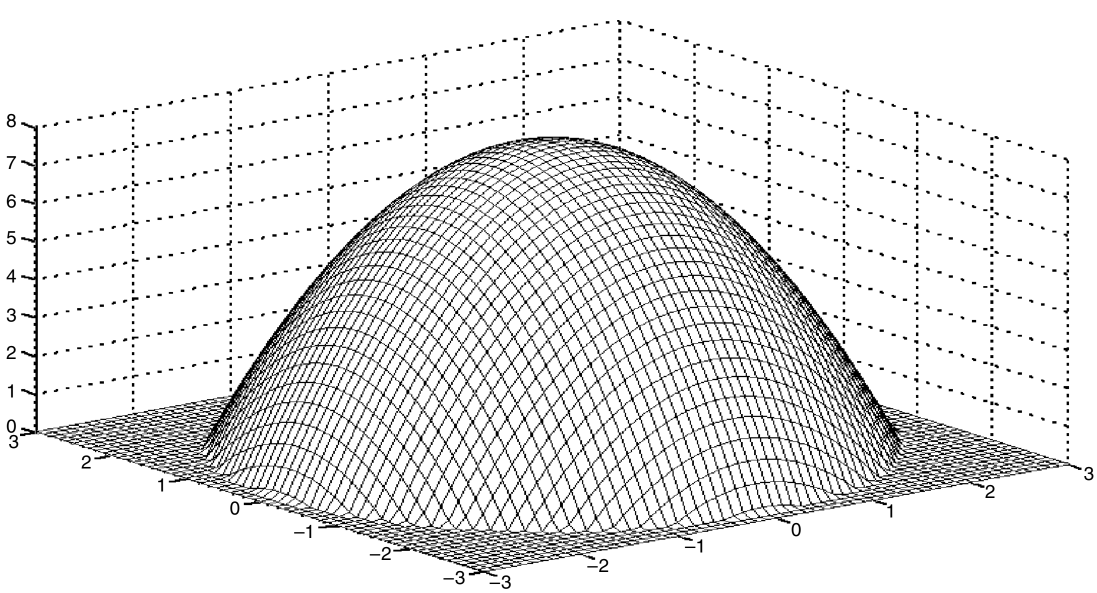


Notes. The continuous ideal density is approximated by three piled slabs. Density y_2 in the second annulus is partly from the second slab and partly from the third (widest) slab.

effort density y_i is partially applied in searches subsequent to the ith, but because the search is random and the target is assumed to be stationary, the conditional nondetection probability if the target is in the annulus is simply $\exp(-y_i)$. Let $\delta_i = A_i - A_{i-1}$ be the area of the ith annulus. The total effort applied to the annulus (in all n searches) is $y_i \delta_i$. The total effort applied in all searches to all regions is therefore $T(\mathbf{A}, \mathbf{y}) \equiv \sum_{i=1}^{n} y_i \delta_i$, where the boldface type indicates vectors of areas and amounts of effort, respectively.

Let $Q(A_0) \equiv 1$, and for $i > 0$, let $Q(A_i)$ be the exclusion probability that the target is outside of region R_i, so that $Q(A_{i-1}) - Q(A_i)$ is the probability that the target is in the ith annulus. Then, because the target must either be in one of the annuli or outside of R_n, the probability of not detecting the target is $N(\mathbf{A}, \mathbf{y}) \equiv Q(A_n) + \sum_{i=1}^{n} (Q(A_{i-1}) - Q(A_i)) \exp(-y_i)$. This is to be minimized subject to the

Figure 1. The inverted SOLR cup has the greatest search effort density at the origin (datum), decreasing smoothly to zero at some radius that depends on the total amount of effort available.



constraint $T(\mathbf{A},\mathbf{y}) \leqslant z$, where z is the total amount of effort available. Let the resulting minimal miss probability be $H_n(z)$.

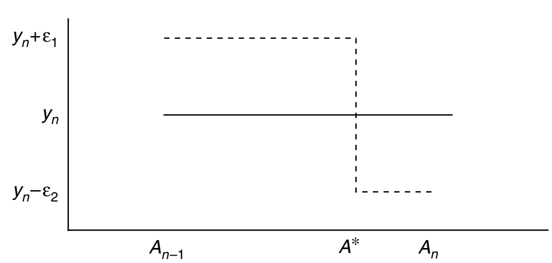
If the prior distribution of target location is uniform, then $Q(A)$ is linear in A, and there is nothing to be gained by a multislab search. The following lemma states the converse.

LEMMA 1. *Assume that $Q(A)$ is not a linear function of A over any interval of positive length. Then, for $z>0$, $H_n(z)$ is a strictly decreasing function of n, and every optimal solution is such that δ_i and y_i are strictly positive for $i=1,\ldots,n$.*

PROOF. There must be some $A>0$ for which $Q(A)$ is smaller than one, because otherwise $Q(A)$ would be linear (constant at one, in fact) on $[0,\infty)$. Therefore, it cannot be optimal to make $\delta_i = y_i = 0$ for all i because the alternative of making $\delta_1 = A$ and $y_1 = z/A$ results in a miss probability of $(1-Q(A))\exp(-z)+Q(A)$, which is smaller than one. If either δ_i or y_i is zero, there is an equivalent solution where n is decremented by one. As long as $H_n(z)$ is a strictly decreasing function of n, it follows that $\delta_i>0$ and $y_i>0$ for all i in every optimal solution. It remains only to show that $H_n(z)$ is strictly decreasing in n.

Let $y_n>0$ be the amount of effort in the outermost region of an arbitrary search plan, and consider an alternative solution where the allocation of effort is changed in the outermost annulus, which is divided into two parts by an intermediate region with area A^*. The effort density is increased by ε_1 in the inner annulus with area A^*-A_{n-1}, and decreased by ε_2 in the outer annulus with area $A_n - A^*$ (see Figure 3). Because $y_n>0$, this is possible for small perturbations. There is no change in the total amount of effort required as long as $\varepsilon_1(A^*-A_{n-1}) = \varepsilon_2(A_n-A^*)$. The net increase in the miss probability is $\exp(-y_n)\{a\exp(-\varepsilon_1)+b\exp(\varepsilon_2)\}$, where $a\equiv Q(A_{n-1})-Q(A^*)$ and $b\equiv Q(A^*)-Q(A_n)$. If $a/(A^*-A_{n-1}) > b/(A_n-A^*)$, the increase can be made negative by choosing $\varepsilon_1>0$, or it can be made negative by choosing $\varepsilon_1<0$ if the opposite inequality holds. That is, except in the case where $a/(A^*-A_{n-1}) = b/(A_n-A^*)$ for all A^* between A_{n-1} and A_n, any solution can be improved by introducing an additional region. Because the exception is eliminated by the hypothesis, this establishes the lemma. □

Figure 3. Revision of effort density from solid to dashed in the outermost annulus.



2.1. Bivariate Normal Prior, Elliptical Regions

Consider first the circular bivariate normal distribution where the standard deviation in each direction is σ. This case has also been investigated by Stone (1975, pp. 179–188), who found necessary and sufficient conditions for a search to be optimal under the restriction that the search density in each region must be an integral multiple of some arbitrary quantum (whereas here the arbitrary parameter is the number of regions).

Because the equiprobability contours are circles, the natural class of regions is concentric circles with increasing radii r_i. The exclusion probability that the target is located outside of the ith circle is $Q(A_i)=\exp(-0.5(r_i/\sigma)^2)$ (e.g., Washburn 2002, pp. 1–9), and the area of the circle is of course $A_i=\pi r_i^2$. Let $x_i \equiv 0.5(r_i/\sigma)^2$, in which case $A_i = x_i(2\pi\sigma^2)$ and

$$N(\mathbf{A},\mathbf{y}) = \sum_{i=1}^{n}\big\{(\exp(-x_{i-1})-\exp(-x_i))\exp(-y_i)\big\} + \exp(-x_n), \tag{1}$$

$$T(\mathbf{A},\mathbf{y}) = (2\pi\sigma^2)\sum_{i=1}^{n} y_i(x_i - x_{i-1}). \tag{2}$$

It is convenient to incorporate the constraint into the objective function by appending a Lagrangian term, in which case the function to be minimized is

$$f(\mathbf{x},\mathbf{y};\lambda) = \sum_{i=1}^{n}\big\{(\exp(-x_{i-1})-\exp(-x_i))\exp(-y_i)+\lambda y_i(x_i-x_{i-1})\big\} + \exp(-x_n), \tag{3}$$

where $0=x_0\leqslant x_1\leqslant\cdots\leqslant x_n$ and $y_i\geqslant 0$. The factor $2\pi\sigma^2$ has been incorporated into λ in this dimensionless version of the problem. If $\lambda\geqslant 1$, it is optimal not to search, and the minimal value is one. There is no minimum if $\lambda=0$, so we assume that $0<\lambda<1$. The Slab Theorem gives the optimal solution, but first we need another lemma, which we state without proof:

LEMMA 2. *Let $g(x)\equiv \ln((\exp(x)-1)/x)$ for $x>0$, with $g(0)\equiv 0$. Then, $g(x)$ is a continuous, monotonically increasing, unbounded function on $[0,\infty)$. Furthermore, $g(x)-x$ is monotonically decreasing and unbounded (below) on the same range.*

SLAB THEOREM. *For some positive parameter μ, the only optimal solution of (3) is when $x_i = i\mu$ and $y_i=(n-i+1)\mu$.*

PROOF. Let $\delta_i = x_i - x_{i-1}$ for $i=1,\ldots,n$. According to Lemma 1, the globally optimal solution must be in the interior, where $\delta_i>0$ and $y_i>0$ for all i. It is therefore necessary that the derivatives with respect to y_i all be zero. Those derivatives are

$$d/dy_i f(\mathbf{x},\mathbf{y};\lambda) = -\exp(-y_i)\big(\exp(-x_{i-1})-\exp(-x_i)\big) + \lambda\delta_i, \qquad i=1,\ldots,n. \tag{4}$$

Equating the derivatives to zero and solving for $x_i + y_i$ (this involves dividing by δ_i, but doing so is permissible), we obtain

$$x_i + y_i = g(\delta_i) - \ln(\lambda), \quad i = 1, \ldots, n, \tag{5}$$

where $g()$ is the function introduced in Lemma 2. Equating the derivatives with respect to x_i to zero results in the same set of equations, except that δ_i in (5) is replaced by the forward difference $y_i - y_{i+1}$, with $y_{n+1} = 0$. Because $g()$ is monotonic increasing, the backward differences of $\mathbf{x}$ must therefore equal the forward differences of $\mathbf{y}$. It follows that $x_i = y_{n-i+1}$, $i = 1, \ldots, n$. We can therefore eliminate $\mathbf{y}$ from consideration, and rewrite (5) as

$$x_i + x_{n-i+1} = g(\delta_i) - \ln(\lambda), \quad i = 1, \ldots, n. \tag{6}$$

Because $g()$ is monotone increasing, it follows from (6) and the commutativity of addition that $\delta_i = \delta_{n-i+1}$, $i = 1, \ldots, n$. Using this, and subtracting the ith equation of (6) from the $i+1$st equation, we obtain

$$\delta_{i+1} - \delta_{n+i-1} = \delta_{i+1} - \delta_i = g(\delta_{i+1}) - g(\delta_i), \quad i = 1, \ldots, n-1. \tag{7}$$

It follows from (7) that $g(\delta_{i+1}) - \delta_{i+1} = g(\delta_i) - \delta_i$, and therefore from the monotonicity of $g(x) - x$ that $\delta_i = \delta_{i+1}$, $i = 1, \ldots, n-1$. Because the successive differences of $\mathbf{x}$ must all be equal, the only alternative is that $x_i = i\mu$ and $y_i = (n-i+1)\mu$ for some parameter μ. From (6), it follows that $(n+1)\mu = g(\mu) - \ln(\lambda)$, which always has a unique, positive solution for μ as long as $0 < \lambda < 1$. □

COROLLARY. *Suppose that the distribution of the target location is bivariate normal with standard deviations σ_X and σ_Y in the two coordinates, that an n-slab search of expanding isoprobability ellipses is to be made, the search of each slab being uniformly random over the entire ellipse, and that the total area searchable for all ellipses (total search effort) is A. Let $z \equiv A/(2\pi\sigma_X\sigma_Y)$ be the total normalized search effort, let $\mu = \sqrt{2z/n(n+1)}$, and let $\rho \equiv \exp(-\mu)$. Then, the ith search should be made in an ellipse whose true area is $i\mu(2\pi\sigma_X\sigma_Y)$, and the total amount of search effort used in the search of that ellipse should be $i\mu^2(2\pi\sigma_X\sigma_Y)$. The detection probability resulting from this search is $H_n(z) = 1 - \rho^n(1 + n(1-\rho))$.*

PROOF. In the elliptical case, the containment variable x_i is $a_i/(2\pi\sigma_X\sigma_Y)$, where a_i is the area of the isoprobability ellipse R_i. If the parameter μ is as in the Slab Theorem, then the total normalized effort applied is $\sum_{i=1}^{n} y_i \cdot (x_i - x_{i-1}) = \mu^2 n(n+1)/2$. This must equal the similarly normalized area available, namely, $A/(2\pi\sigma_X\sigma_Y)$, which implies that μ must be as stated. The effort density in the search of the ith ellipse must be $y_i - y_{i+1}$, which is also μ, so the total true effort spent in searching that ellipse is the area of the ellipse times μ, or $i\mu^2(2\pi\sigma_X\sigma_Y)$, as claimed. The resulting miss probability is

$$\begin{aligned} 1 - H_n(z) &= \exp(-x_n) + \sum_{i=1}^{n} \{(\exp(-x_{i-1}) - \exp(-x_i))\exp(-y_i)\} \\ &= \rho^n + \sum_{i=1}^{n} (\rho^{i-1} - \rho^i)\rho^{n-i+1}. \end{aligned}$$

As was to be shown, the latter sum is $\rho^n(1 + n(1-\rho))$. □

Note that the optimized effort density (the ratio of effort to area) in every slab is μ, an operationally simplifying feature. Figure 2 shows the SOLR cup and the optimal three-slab approximation to it for a problem where $\sigma_X = \sigma_Y = z = 1$, in which case $\mu = 0.408$.

Although the n slabs can actually be searched in any order, there is something to be said for searching them in order of size, the smallest region being searched first. If that is done, it is easy to check that the detection probability at the end of searching slab i is as large as possible, subject to only searching i slabs with whatever amount of effort has been expended at that time. This is a kind of limited uniform optimality. It is possible to perform an SOLR search in such a manner that the detection probability is maximized at every time, from which it follows that SOLR search also minimizes the mean time to detection (Stone 1975, pp. 51–52). There is no comparable result for slab searches. Indeed, there is likely to be a conflict between the goals of simplicity and minimizing the conditional mean time to detection because the former requires a raster search of each slab that will not minimize the latter. Nonetheless, in the above limited sense, it is best to search the slabs in increasing order.

It can be shown that

$$\lim_{n\to\infty} 1 - H_n(z) = 1 - (1 + \sqrt{2z})\exp(-\sqrt{2z}),$$

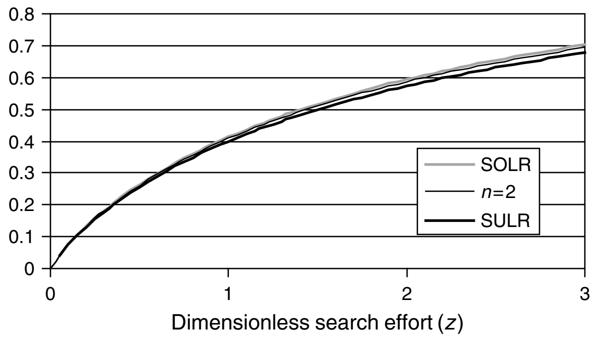
the detection probability for the SOLR case. Washburn (2002, pp. 5–6, 5–11) also defines the strategically uniform locally random (SULR) case corresponding to $n = 1$, for which $1 - H_1(z) = (1 - \exp(-\sqrt{z}))^2$. We have thus bridged the two extremes, showing how efficiency increases as the search is allowed to become more complicated. It turns out to increase very little (Figure 4). Even very simple searches with $n = 1$ or 2 achieve a detection probability that is remarkably close to the SOLR limit.

Although the global minimum must be interior, there are also local minima where either δ_i or y_i is zero. The reader may wish to investigate the tendency of a numerical optimization procedure to converge to one of these edge minima using page “Circles” of *PiledSlab.xls*.

2.2. Bivariate Normal Prior, Rectangular Regions

Next, we consider a sequence of regions that are squares if the bivariate normal is circular or, more generally, rectangles for which the width/length ratio is σ_X/σ_Y. The

Figure 4. Detection probability increases with n, but SOLR ($n = \infty$) is not much larger than SULR ($n = 1$).



problem is to determine an optimal sequence of n such rectangles, together with an amount of effort for each, subject to the constraint that the total amount of effort for all of them is A. Rectangles are simpler regions to search than ellipses, hence our interest.

If A_i is the area of the ith such rectangle, then the inclusion probability is the square of the probability that a normal random variable is contained within an interval, specifically,

$$1 - Q(A_i) = \left[2\Phi\left(0.5\sqrt{\frac{A_i}{\sigma_X \sigma_Y}}\right) - 1\right]^2.$$

Lemma 1 still applies because this function is not linear over any interval, so we can conclude that only interior minima of the miss probability can be global. As in the case of circular regions, there are numerous inferior local optima where either δ_i or y_i is zero for at least one value of i, any of which may prove attractive to a numerical optimization procedure, depending on the starting point. As a practical matter, the solution to the circular problem is a good starting point for the square problem, generally resulting in fast convergence to a nearby interior minimum. See the "Squares" sheet of *PiledSlab.xls*, which uses Excel's Solver to seek a minimum from a starting point entered by the user.

Rectangles can be efficiently searched using a lawnmower pattern that utilizes parallel tracks of equal length. This cannot be said for ellipses, so there is a good argument for searching squares if the sacrifice in detection probability is small. Table 1 compares the circular and square solutions of a problem where $n = 4$. The miss probability is 0.469551 when circles are optimized, or 0.476504 when squares are optimized. If the optimal circular areas are simply converted to squares with the same area, the miss probability rises slightly to 0.476509. There is little to be lost by searching squares instead of circles, and still less to be lost by forcing the analytic circular solution onto the square problem.

Table 1. The circular solution when $\sigma = 1$, $n = 4$, and $A = 10$, with the offset for the square case shown in parentheses.

Index i	A_i	$1 - Q(A_i)$	y_i
1	2.51 (−0.04)	0.33	1.60 (+0.011)
2	5.01 (−0.05)	0.55	1.20 (+0.005)
3	7.52 (−0.03)	0.70	0.80 (+0.001)
4	10.03 (+0.03)	0.80	0.40 (−0.001)

Note. Early square areas are slightly smaller, with slightly higher effort densities.

Reber (1957) discusses an infinitely divisible search plan where the search density is in the shape of a smooth pyramid that approximates the inverted SOLR cup. Reber's plan by design has rectangular contours of constant search density. The associated miss probability is, of course, larger than the SOLR miss probability. It can even be larger than the miss probability in an optimized square search, in spite of the fact that the square search might be thought of as an attempt to approximate Reber's plan. The example given above is a case in point—Reber's miss probability is 0.476814, which is larger than 0.476504 (see page "Squares" of *PiledSlab.xls*).

3. General Independent Search of Expanding Regions

In this section, we consider a sequence of uniform, independent searches of a sequence of regions, each of which includes all of its predecessors. Although the individual searches are uniform, they are not necessarily random. It is not necessarily true that the total search effort in an annulus is a sufficient statistic for computing the miss probability, so we proceed differently. Let

$N_i \equiv$ event target not located in region i;
$D_i \equiv$ event target detected during the search of region i;
$E_i \equiv$ event target detected during the search of region i, or before.

Also let P_i be the probability of E_i. Then, because $E_i = D_i \cup E_{i-1} \cap D_i'$, where D_i' is the negation of D_i, it follows from the theorem of total probability that

$$P_i \equiv P(E_i) = P(D_i) + P_{i-1} P(D_i' \mid E_{i-1}). \tag{8}$$

Let $Q_i = P(N_i)$ and $F_i \equiv P(D_i \mid E_{i-1})$. Because the search of the ith region is uniform, and because E_{i-1} implies N_i', $F_i = P(D_i \mid N_i')$, the probability of detection given that the target is in the region searched. Because

$$P(D_i) = P(D_i \mid N_i')(1 - Q_i),$$

(8) is equivalent to

$$P_i = F_i(1 - Q_i) + P_{i-1}(1 - F_i), \quad i \geqslant 1. \tag{9}$$

This holds even when $i = 1$ if we define $P_0 = 0$ for convenience. If we have a way to compute the exclusion probabilities Q_i and the conditional detection probabilities F_i, Equation (9) can be used to determine P_1, P_2, etc. by recursion. Equation (9) is derived and employed for that purpose by Engel (1981) in the special case of rectangular regions.

3.1. Bivariate Normal Prior, Inverse-Cube Law

The inverse-cube law was developed in World War II as a model of detection of a ship's wake. It has convenient analytic properties (Koopman 1956, pp. 521–524; Washburn 2002, pp. 2–9 to 2–12), and still enjoys modern use as a detection model that lies between random search and exhaustive search in terms of effectiveness. IAMSAR (2003), for example, includes search-planning procedures for ideal search conditions (inverse-cube law) and "poor" search conditions (random search).

The inverse-cube law assumes that the area in question is covered by parallel tracks spaced S apart. The detection probability depends on the coverage ratio $c \equiv W/S$, where W is the sweep width. Specifically, if the coverage ratio in the ith search is c_i, then the detection probability, given that the target is actually in the ith area, is

$$F_i = 2\Phi\left(\sqrt{\frac{\pi}{2}}c_i\right) - 1. \tag{10}$$

Equivalently, c is the ratio of the area covered by the search (track length multiplied by W) to the area of the region searched (track length multiplied by S), an effort density. Because the area covered and track length are proportional, track length itself can be taken to be the measure of effort.

Consider the problem of optimally allocating a total amount of track length to a sequence of expanding squares. If the ith square has area A_i, then it contains the target with a probability that depends on the ratio of the square's side $\sqrt{A_i}$ to the standard deviation σ:

$$1 - Q_i = \left[2\Phi\left(\sqrt{A_i}/(2\sigma)\right) - 1\right]^2. \tag{11}$$

Using (9), we can now calculate P_i by recursion, as long as the squares and the coverage ratios are known. There remains the question of determining these optimally, subject to a constraint on total effort.

This optimization problem was approximately solved by Koopman (1956) and his colleagues (hereafter the Koopman plan) in World War II. The Koopman plan involves maintaining a constant track spacing in all searches, that spacing being given by the formula $S = 0.723\sqrt{\sigma W}$. For that spacing, the number of regions n is determined so that the overall distribution of searching effort closely approximates the optimal distribution in the SOLR search (see Koopman 1980, pp. 214–221, or Frost and Stone 2001 for details). In the Koopman plan, the number of regions n is determined by the sweep spacing.

When the Koopman plan falls short of being optimum, as it sometimes does, the reason is usually not that it does a poor job of approximating an SOLR search, but rather that the SOLR search is not a good thing to approximate when detections are according to the inverse-cube law. Recall that the optimal number of regions in the case of random search is theoretically infinite. That is not true in the case of the inverse-cube law. Employing many regions is still wise in the sense that it permits the overall search to be shaped so that the heaviest effort is put where the target is most likely to be. However, the successive searches are all (we assume) independent of each other, whereas any given inverse-cube law search represents an attempt to cover its region in an organized manner. Thus, each additional region introduces an element of disorganization into a search that might otherwise have been more organized. As a result, the optimal number of regions is surprisingly small. Table 2 shows some comparisons of the best-known solution with the Koopman plan. Except in problems where very high detection probabilities are possible, the optimal number of regions is one. The interested reader can do further testing of this claim using page "InvCube" of *PiledSlab.xls*. The penalty for using too many regions can be significant, particularly if one takes into account the wasted effort that occurs, as a practical matter, when switching from one region to another.

Table 2. The optimal (alternatively Koopman) number of regions and the optimal (alternatively Koopman) detection probability in six problems, where track length L and sweepwidth W are as shown, and the standard deviation is $\sigma = 1$.

L	W	Regions, n	Det. Prob., P_n
96.2	0.01	1 (4)	0.123 (0.112)
96.2	0.1	1 (4)	0.617 (0.565)
96.2	1.0	2 (4)	0.994 (0.985)
54.09	0.1	1 (3)	0.451 (0.417)
54.09	1.0	2 (3)	0.970 (0.950)
24.04	0.1	1 (2)	0.259 (0.246)
24.04	1.0	1 (2)	0.857 (0.829)

Note. An independent search according to the inverse-cube law is made in each region.

The fact that a single region is typically the optimal number of regions means that optimal searches (within the class being considered) are simple ones, a useful property. However, the same fact also means that the amount of searching effort needs to be known before the search begins. Suppose that the initial search of an optimal square region A_1 does not result in finding the target, and that the decision is subsequently made to continue searching for a second amount of track length. The optimal square A_2 is unlikely to be coincident with A_1, even though the optimal number of regions would still be one if the total amount of track length could be allocated optimally. In other words, there may be cause to regret having searched A_1 in the first place. Such problems are the subject of the next section.

4. Myopic Optimization and Inversion

So far, we have assumed that a given amount of search effort must be optimally distributed over a number of stages, each of which consists of the search of a region that includes its predecessors. It is sometimes the case that the amount of effort for each stage is not known beforehand. Unexpected search resources may arrive, or the importance of finding the target may increase to the point where a search is extended in time. At each stage of planning, a known amount of search effort is available, but even the existence of subsequent stages may be in doubt. We consider, then, the "myopic" search-planning problem where the objective is always to maximize the detection probability for the current search. Equation (9) is central, with P_{i-1} summarizing past decisions, and with current decisions affecting Q_i and F_i. NAVSAR (Navy 1990), a search and rescue tool sometimes used by the U.S. Navy in looking for downed aircraft, is organized in this manner. The myopic optimization problem is also considered in IAMSAR (2003, pp. 4–21 to 4–33).

There is a penalty for not knowing the future. Consider the second data line of Table 2, for example. If the total track length of 96.2 is split into four equal parts, each of which is myopically optimized, the detection probability is only 0.558, smaller than both the Koopman plan and the optimal search, each of which is nonmyopic in the sense that each knows the total effort available when the first search is undertaken. Search for a stationary target can be arranged so that myopic search is always globally optimal, provided the class of search strategies is sufficiently large (Washburn 2002, pp. 5–16). However, the class of piled-slab searches is not sufficiently large to permit this, particularly in searches where the optimal number of slabs tends to be one.

Casual use of (9) can lead to incorrect answers in myopic optimization. For example, suppose that the prior distribution is circular normal with $\sigma = 1$, and that an amount of effort $A_1 = 2\pi$ is available for a single random search. According to the corollary, the optimal search region R_1 has area $a_1 = 2\pi$, $\mu = z = 1$, and the optimal detection probability is $P_1 = H_1(1) = (1-\exp(-1))^2 = 0.3996$. Now suppose that a second search with amount of effort $A_2 = 1$ becomes possible, while the first search is fixed as above. As long as the second search is in a region R_2 that includes R_1, F_2 is given by the random search formula $1-\exp(-A_2/a_2)$. The exclusion probability is $Q_2 = \exp(-a_2/(2\pi\sigma^2))$. Using (9), the optimal value of a_2 is 9.0, and the resulting detection probability after both searches is $P_2 = 0.4376$. However, this is only a local minimum, because the formula for F_2 is incorrect when R_2 is contained in R_1, rather than vice versa. Casual use would take the fact that the locally optimal a_2 is larger than a_1 as verifying that the globally optimal a_2 has the same property, but in fact the possibility that a_2 should be smaller than a_1 has never been accurately investigated. The correct value of a_2 is actually 2.5 in this example, and the best detection probability is 0.4395. This can be verified using (9), but with the chronologically second search being logically first. If the total amount of search effort could have been optimized between the two searches, the detection probability would be 0.4445, again illustrating that myopic search is not globally optimal.

The error produced by casual use in this example is small, as it usually is, but nonetheless real. The possibility that optimal search areas should be "inverted" in the sense that a late search should be made over a smaller region than an earlier one must be reckoned with when effort magnitudes are determined arbitrarily, especially if the chronologically late magnitudes are relatively small.

The computational cost for dealing with potential inversions is significant. If the planned regions are $R_1, \ldots, R_n$ in increasing (not chronological) order, then, because the next region might lie anywhere in the sequence, they must all be remembered, along with associated data, if (9) is to be used to compute the maximized P_{n+1}. A possible maximizing procedure is outlined below.

Let Q_i and F_i be as defined earlier, and suppose that it is desired to add a new search of a new region. If the areas of the current regions are $x_1, \ldots, x_n$, and if the new region to be searched has area x, introduce $x_0 \equiv 0$ and $x_{n+1} \equiv \infty$, and say that the new region is inserted in the ith interval if $x_i \leqslant x < x_{i+1}$ for $0 \leqslant i \leqslant n$. Let $Q(x)$ and $F(x)$ be the exclusion probability and conditional detection probability for the new region. $F(x)$ will also depend on the amount of effort available for the new search, although this is not reflected in the notation. Then, using (9) and induction, it can be shown that the detection probability after all $n+1$ searches are complete is

$$Pd(x) = A_i F(x)(1-Q(x)) + B_i(1-F(x)) + C_i \tag{12}$$

for x in interval i, where

$$A_i = \prod_{j=i+1}^{n}(1-F_j), \qquad B_i = \sum_{j=1}^{i} Q_j F_j (1-F_j),$$
$$C_i = \sum_{j=i+1}^{n} Q_j F_j (1-F_j), \tag{13}$$

with the usual conventions for empty sums and products. In terms of these quantities, the detection probability P_i as defined by (9) is B_i/A_i. Equation (12) defines $Pd(x)$ for all $x \geqslant 0$ because the intervals partition the nonnegative real line.

The maximum of $Pd(x)$ could occur anywhere, but analytic methods may somewhat reduce the labor of finding it if $F(x)$ and $Q(x)$ are each differentiable functions of x. In that case, $Pd(x)$ is continuous everywhere, and differentiable except possibly at interval endpoints. Letting the symbol $'$ (prime) denote derivative with respect to x, from (12) we have, in interval i,

$$Pd'(x) = -A_i Q'(x) F(x) + F'(x)(A_i(1-Q(x)) - B_i). \tag{14}$$

Assume that $F'(x) < 0$ for all $x > 0$, and let $P^*(x) \equiv 1 - Q(x) - F(x)(Q'(x)/F'(x))$. Then, $Pd'(x) > 0$ if and only if $P^*(x) < B_i/A_i$, or equivalently, $P^*(x) < P_i$. If $P^*(x)$ is an increasing function of x, there can be a local maximum in the interior of interval i if and only if both $P^*(x_i) < P_i$ and $P^*(x_{i+1}) > P_{i+1}$. In that case, say that the interval passes the possibility test. Intervals that pass the possibility test must be more closely investigated, but most intervals will fail the test, and are therefore easily disposed of as alternatives for containing the optimal value of x.

Page "Myopic" of *PiledSlab.xls* is set up to calculate the correct insertion interval in the case of the circular random search of a normal prior, in which case $F'(x) < 0$ and $P^*(x)$ is indeed an increasing function of x for all $x > 0$. If an interval passes the possibility test, then the position of the local maximum is calculated using Newton's method to find the place where $Pd'(x) = 0$.

There seems to be little else to exploit in the process of finding a global maximum. There may be multiple local maxima, and any local maximum might be global.

The inversion considerations of this section are of no concern when search effort and search area are both decision variables, as they are when the only constraint is on total search effort. Because the chronological order in which searches are made is irrelevant when the target is stationary, no additional generality ensues when one permits inversions.

5. Summary

We have considered the use of uniform searches over an expanding sequence of regions in a variety of circumstances. We have found an analytic optimal solution in the case of random search of a normal prior using expanding circles, thus generalizing a solution found in World War II. This solution can serve as a good approximation to random search using expanding squares, or a slightly better solution can be achieved by using it as the starting point for a local optimization.

Although circles turn out to be a good approximation to squares, it is not true that an optimal random search is a good approximation to an optimal inverse-cube law search. In the latter, there is a strong tendency for optimal searches to consist of searching a single, carefully chosen slab.

We have also considered problems where the optimization must be done myopically. Reliable use of Equation (9) must account for the possibility of "inversions" where chronologically late searches should be over regions lying within those already searched. Although the improvements are generally small, myopic search that permits inversions is generally better than myopic search that does not.

References

Discenza, J. 1980. A solution for the optimal multiple rectangle problem. K. Haley, L. Stone, eds. *Search Theory and Applications*. Plenum, 261–272.

Engel, D. 1981. Finding optimal search rectangles for stationary targets. Daniel H. Wagner Associates report to the U.S. Coast Guard. Paoli, PA. Revised May 27, 1981.

Frost, J., L. Stone. 2001. Review of search theory: Advances and applications to search and rescue support. U.S. Coast Guard Report CG-D-15-01, §4.5. Retrieved October 14, 2004. www.rdc.uscg.gov/Reports/2001/CGD1501Report.pdf.

IAMSAR. 2003. *IAMSAR Manual*, Vol. 2. International Civil Aviation Organization,

Koopman, B. 1946. Search and screening. Operations Evaluation Group, Report 56, 111–114. U.S. Department of the Navy, Washington, D.C.

Koopman, B. 1956. Theory of search II. *Oper. Res.* **4**(5) 503–531.

Koopman, B. 1980. *Search and Screening*. Pergamon Press, New York.

Navy. 1990. Software design document for the Navy search and rescue (NAVSAR) probability functions. Naval Oceanographic Office, Washington, D.C., (OAML-SDD-35).

Reber, R. 1957. Theoretical evaluation of various search salvage procedures for use with narrow-path locators (part II). Technical report 118, Bureau of Ships, U.S. Navy, Washington, D.C., 75–78.

Stone, L. 1975. *Theory of Optimal Search*. Academic Press, New York.

Washburn, A. 2002. *Search and Detection*, 4th ed. Topics in Operations Research. INFORMS, 7240 Parkway Drive, Hanover, MD.