CONTINUOUS AND DISCRETE SEARCH FOR ONE OF MANY OBJECTS

David ASSAF and Shmuel ZAMIR
Hebrew University of Jerusalem, Mount Scopus, Jerusalem, Israel

Received May 1987
Revised September 1987

By an easy observation we show that the basic result of Blackwell [2], according to which the most inviting strategy is optimal in a discrete search for one object, is also true when the number of objects is random provided the search is made in continuous time. This result does not hold in the discrete search model even when only two boxes are present (contrary to a conjecture of Smith and Kimeldorf [7]). For the case of two boxes, a convenient sufficient condition on the distribution of the number of objects is provided which ensures optimality of the most inviting strategy. As a result, this strategy is shown to be optimal for several important distributions.

optimal search * most inviting strategy * uniformly optimal strategy

1. Introduction and summary

Suppose N objects are hidden in m boxes. The number of objects, N, is unknown, but has a known (prior) distribution $\{P_n\}_{n=1}^{\infty}$ ($P_n = P(N = n)$, $\sum_{n=1}^{\infty} P_n = 1$). The number of boxes, $m \geqslant 2$, is assumed known and we denote the set of boxes by $M = \{1, 2, \ldots, m\}$. Each of the objects is hidden in box i with known probability π_i ($\pi_i > 0$, $\sum_{i=1}^{m} \pi_i = 1$). The objects are hidden indepently of each other, so that if X_i denotes the number of objects in box i, then, given $N = n$, $X = \{X_1, \ldots, X_m\}$ has a multinomial distribution with parameters n and $\pi = (\pi_1, \ldots, \pi_m)$.

Associated with each box i are two quantities which are first described for the continuous search version:

(i) A cost of search for objects in box i for t units of time, which is assumed to be of the form $c_i t$. The vector $c = (c_1, \ldots, c_m)$ is known with $c_i > 0$ for all $1 \leqslant i \leqslant m$.

(ii) A (conditional) probability of finding an object which is hidden in box i when that box is searched for t time units.

This probability is assumed to be of the form $a(i, t) = 1 - e^{-\beta_i t}$ with $\beta_i > 0$, $1 \leqslant i \leqslant m$ known. Searches for different objects hidden in the same box are assumed independent in the sense that if there are k objects in box i, then the probability of finding at least one of them when searching that box for t time units is $1 - (1 - a(i, t))^k$.

In the discrete search version considered in Section 2, only integer multiples of $t = 1$ are allowed and the resulting parameters are denoted by c_i and α_i ($\alpha_i = a(i, 1)$).

The primary objective considered is to search the boxes until an object is found and to do so with a minimal total expected cost. Other objectives are also mentioned in Section 3.

For the special case, $p_1 = 1$, the results are well known (Blackwell [2], Stone [9]); the optimal strategy is to start the search in a 'most inviting' box, i.e., a box i for which the 'marginal probability per unit cost' of finding the object, $\alpha_i \pi_i / c_i$ is highest ($\beta_i \pi_i / c_i$ in the continuous version). This remains the strategy until the object is found but rather than the initial π, we substitute a modified π which reflects the conditional probabilities of the boxes given the current history of unsuccessful searches. A strategy of this type is called a *most inviting strategy*.

When N has an arbitrary distribution, the problem is apparently much harder. The discrete search version was studied by Smith and Kimeldorf [7] (S–K henceforth), and their basic result is that for $m \geqslant 3$, the most inviting strategy is optimal when N has a PP(λ) distribution (Poisson conditioned on $N \geqslant 1$), but if N is not PP(λ) then there exist α, π and c for which this strategy is not optimal. (For an arbitrary N, the most inviting box is one whose h_i/c_i is highest where h_i is the 'current' probability of finding at least one object when searching box i.) The case $m = 2$ is different and S–K conjecture that the most inviting strategy is optimal for any distribution $\{P_n\}_{n=1}^{\infty}$.

The case $m = 2$ is closely studied in Section 2. We first show that the S–K conjecture is false (Example 1). A convenient sufficient condition on the distribution of N which ensures optimality of the most inviting strategy is then provided (Theorem 1). Properties of this condition are studied and as a result, it is shown to hold in several standard cases such as the binomial, and Poisson distributions.

The continuous search version for $N \geqslant 1$ was apparently never studied. In Section 3 we make a simple observation the consequence of which is that the most inviting strategy is optimal for any m and any distribution $\{P_n\}_{n=1}^{\infty}$. Moreover, this strategy is shown to be uniformly optimal (optimal for any given budget) and in a sense is also the strategy that provides the best information on the distribution of N. This actually means that the non-optimality of the most inviting strategy in the discrete case is due to '*non-divisibility*' of the search effort.

Some recent papers which deal with other aspects of the problem when $N \geqslant 1$ are [1], [3], [4], [5], [8] and [6]

We would like to thank Isaac Meilijson for many helpful discussions.

2. Discrete search

When searching for at least one object in discrete time, a search strategy is a sequence $s = (s_1, s_2, \ldots)$, where $1 \leqslant s_t \leqslant m$ is the box to be searched at stage t if no object was found in the first $t-1$ stages. Let h_i be the probability of success (i.e., finding at least one object) when searching box i. That is,

$$h_i = 1 - \sum_{n=1}^{\infty} (1 - \alpha_i \pi_i)^n P_n = 1 - \mathrm{E}(1 - \alpha_i \pi_i)^N,$$

$1 \leqslant i \leqslant m$.

Following an unsuccessful search of box k the number of objects $N^{(k)}$ has the modified distribution $\{P_n^{(k)}\}_{n=1}^{\infty}$ given by

$$P_n^{(k)} = \frac{(1 - \alpha_k \pi_k)^n P_n}{\mathrm{E}(1 - \alpha_k \pi_k)^N},$$

and the revised (conditional) probability of success when searching box i is

$$h_i^{(k)} = \frac{\mathrm{E}(1 - \alpha_k \pi_k)^N - \mathrm{E}(1 - \alpha_i \pi_i - \alpha_k \pi_k)^N}{\mathrm{E}(1 - \alpha_k \pi_k)^N} \quad \text{if} \quad i \neq k, \tag{2}$$

and

$$h_k^{(k)} = \frac{\mathrm{E}(1 - \alpha_k \pi_k)^N - \mathrm{E}(1 - 2\alpha_k \pi_k + \alpha_k^2 \pi_k)^N}{\mathrm{E}(1 - \alpha_k \pi_k)^N}. \tag{3}$$

The most inviting strategy is one which searches at each stage box i, where i is a maximizer of the 'current' h_i/c_i. This strategy was suggested by Blackwell [2] who proved its optimality for $N = 1$ (in the sense of minimizing the total expected cost until an object is found). What are the distributions of N for which the most inviting strategy is optimal? S–K [7] basically answer this question for $m \geqslant 3$ and conjecture that for $m = 2$ the most inviting strategy is optimal for any distribution of N.

Our first observation is that this conjecture is false, as can be seen from the following example.

Example 1. Let $\alpha_1 = 0.999$, $\alpha_2 = 0.9$, $c_1 = 0.53$, $c_2 = 1$, $\pi_1 = 0.01$, $\pi_2 = 0.99$, and take N as a two point random variable such that $p_1 = \frac{1}{2}$ and $p_n = \frac{1}{2}$ for some large n ($n = 500$ will do). The most inviting strategy is checked to have the form $S = (1, 2, 2, s_4, s_5, \ldots)$. The expected cost of using this strategy can be written as

$$C(S) = C(1, 2, 2) + q(1, 2, 2) C(R \mid 1, 2, 2),$$

where $C(1, 2, 2)$ is the expected cost of the first three stages (stopping at each stage if an object is found), $q(1, 2, 2)$ the probability of not finding any objects in the first three stages and $C(R \mid 1, 2, 2)$ the expected cost of the search sequence $R = (s_4, s_5, \ldots)$ given that no object was found in the first three stages.

Consider the search strategy $S' = (2, 2, 1, s_4, s_5, \ldots)$. Then,

$$C(S') = C(2, 2, 1) + q(2, 2, 1) C(R \mid 2, 2, 1).$$

Clearly $q(1, 2, 2) = q(2, 2, 1)$ and $C(R \mid 1, 2, 2) = C(R \mid 2, 2, 1)$. Careful computation reveals that for $n = 500$ these quantities are $C(1, 2, 2) = 1.075$; and $C(2, 2, 1) = 1.060$. Thus, S' is better than S and the most inviting strategy is not optimal.

Facing this negative result we proceed now to look for useful sufficient conditions for the optimality of the most inviting strategy.

Let $g_N(s) = \sum_{n=0}^{\infty} P_n s^n$ be the (probability) generating function of the random variable N.

Theorem 1. *For $m = 2$, if $\log g_N(s)$ is concave on $0 < s \leqslant 1$ then the most inviting strategy is optimal.*

Proof. Observe first that if $f(s)$ is a concave function on $(0, 1)$ then for any x, y, z in $(0, 1)$ such that $x + y < 1$, the following inequality holds:

$$f((1-x)(1-z)) + f((1-y)(1-z)) \geqslant f((1-x-y)(1-z)) + f(1-z). \qquad (4)$$

This follows readily from the definition of concavity noting that

$$(1-x)(1-z) = \frac{x}{x+y}(1-x-y)(1-z) + \frac{y}{x+y}(1-z)$$

and

$$(1-y)(1-z) = \frac{y}{x+y}(1-x-y)(1-z) + \frac{x}{x+y}(1-z).$$

Writing (4) for $f(s) = \log g_N(s)$ and taking exponential of both sides of the inequality we obtain

$$\mathrm{E}[(1-x)(1-z)]^N \mathrm{E}[(1-y)(1-z)]^N \geqslant \mathrm{E}[(1-x-y)(1-z)]^N \mathrm{E}(1-z)^N. \qquad (5)$$

Next note that if the distribution of N satisfies (5) then the conditional distribution given an unsuccessful search in box k also satisfies (5). In fact by (1) the modified probabilities for N are of the form $P_n^{(k)} = (1-u)^n P_n / \mathrm{E}(1-u)^N$. Substituting this in (5) yields the same inequality with the original distribution P_n and $(1-z)$ replaced by $(1-\bar{z}) = (1-z)(1-u)$. As a result we conclude that concavity of $\log g_N(s)$ implies that (5) holds for the updated distribution of N following any search history.

Next we show that if box 2 is more inviting than box 1 then (5) implies that it remains more inviting following an unsuccessful search of box 1. In fact, we show that (5) implies

$$\frac{h_2^{(1)}}{h_1^{(1)}} \geqslant \frac{h_2}{h_1} \qquad (6)$$

To prove (6) substitute $x = \alpha_1 \pi_1$ and $y = \alpha_2 \pi_2$ in (2) to obtain using (5) with $z = 0$ that $h_2^{(1)} \geqslant h_2$. Also $(1 - 2\alpha_1\pi_1 + \alpha_1^2\pi_1)^n \geqslant (1 - 2\alpha_1\pi_1 + \alpha_1^2\pi_1^2)^n = (1 - \alpha_1\pi_1)^{2n}$. Using this and $\mathrm{E}W^2 \geqslant (\mathrm{E}W)^2$, we have $h_1^{(1)} \leqslant h_1$ and inequality (6) now follows.

Inequality (6) implies that if box 2 is more inviting than box 1 it remains more inviting following an unsuccessful search of box 1.

The proof of our theorem is now completed by using the basic result which states: If a most inviting box remains most inviting as long as it is not searched then the most inviting strategy is optimal (see [1, Lemma 2 and Theorem 2). □

We say that the random variable N is GLC if it has a log concave generating function. The following simple propositions are worth noting for future applications:

Proposition 1 (closure under convolution). *If N_1 and N_2 are independent and both GLC then $N_1 + N_2$ is GLC.*

Proof. Straightforward. □

Proposition 2 (closure under conditioning on positiveness). *If N is an integer valued non-negative GLC then N^+ defined by $P(N^+ = k) = P(N = k \mid N \geqslant 1)$ is also GLC.*

Proof. A variable N is GLC if and only if $g_N'' \cdot g_N - (g_N')^2 \leqslant 0$. Since $g_N'' \geqslant 0$ it follows that for any $\alpha \geqslant 0$, $\beta > 0$, $\log(\beta g_N - \alpha)$ is also concave. In particular: $\log g_{N^+}(s) = \log[(g_N(s) - P_0)/(1 - P_0)]$ is concave. □

Proposition 3 (closure under translations). *If N is GLC and $n \geqslant 1$ is a fixed integer then $N + n$ is also GLC.*

Proof. Since $\log g_n(s) = n \log s$ is concave, the proof follows from Proposition 1. □

An immediate consequence of Theorem 1 and the above propositions is that the most inviting strategy is optimal for the following distributions of N:

(a) N is constant.
(b) N is Bernoulli ($P_0 + P_1 = 1$). (Immediate since $g_N(s) = P_0 + P_1 s$.)
(c) N is Poisson or positive Poisson. (For Poisson $\log g_N(s) = \lambda(s-1)$ and for positive Poisson use Proposition 2.)
(d) N is binomial or positive binomial (combine (b) with Propositions 1 and 2).
(e) $P_n + P_{n+1} + P_{n+2} = 1$ for some $n \geqslant 1$. (Verify directly for $n = 1$ and apply Proposition 3.) It may be of interest to note that this is not true for $n = 0$ (for example, $P_0 = 0.98$, $P_1 = P_2 = 0.01$).
(f) N is a convolution of two Bernoulli r.v. with different values of P. Such distributions on $\{0, 1, 2\}$ will be GLC by (b) and Proposition 1. However, not all distributions on $\{0, 1, 2\}$ are GLC as shown in example (e).

Finally we remark that the geometric distribution is not GLC and hence Theorem 1 is not applicable. In fact it is proved in Sharlin [6] that for geometrically distributed N the most inviting strategy may not be optimal.

3. Continuous search

In this section we make some observations showing that the non-optimality of the most inviting strategy is due to the *non-divisibility* of the search effort. In fact in a continuous time model the most inviting strategy is optimal in a very strong sense for any distribution of N.

A *search strategy* is a function $f\colon R \times M \to R$ (here R denotes $[0, \infty)$), such that

$$\sum_{i=1}^{m} c_i f(b, i) \leqslant b \quad \text{for all} \quad b \geqslant 0, \tag{7}$$

$f(b, i)$ is non-decreasing in b

$$\text{for each} \quad 1 \leqslant i \leqslant m. \tag{8}$$

Here $b \geqslant 0$ represents some budget and $f(b, i)$ is the total time of search allocated to box i when the budget is b. Requirement (8) states that we cannot 'change our minds' and spend less time in box i then the time we have already spent there.

The probability of finding at least one object when using the strategy f and a given budget b is given by

$$P_{N,b}(f) = 1 - \sum_{n=1}^{\infty} P_n \left[\sum_{i=1}^{m} \pi_i \, e^{-\beta_i f(b,i)} \right]^n. \tag{9}$$

For any fixed $b \geqslant 0$, define a *b-allocation* as a function $g\colon M \to R$ satisfying $\sum_{i=1}^{m} c_i g(i) \leqslant b$. The probability of finding at least one object using the b-allocation g is given by $P_N(g) = P_{N,b}(f)$ with f satisfying $f(b, i) = g(i)$, $1 \leqslant i \leqslant m$. Let $G(b)$ be the set of all b-allocations and define $V(b) = \sup_{g \in G(b)} P_N(g)$.

A strategy f^* is called *b-optimal* if $P_{N,b}(f^*) = V(b)$ and *uniformly optimal* if $P_{N,b}(f^*) = V(b)$ for all $b \geqslant 0$.

First observe that (9) can be rewritten in terms of $P_{1,b}(f)$, the probability of success under f when there is one object,

$$P_{N,b}(f) = 1 - \sum_{n=1}^{\infty} P_n (1 - P_{1,b}(f))^n = 1 - \mathrm{E}(1 - P_{1,b}(f))^N. \tag{10}$$

Since the constraints (7) and (8) do not depend on the distribution of N we readily have the following:

Observation. A search strategy f^* is uniformly optimal for any random N if and only if it is uniformly optimal for $N = 1$.

The case $N = 1$ is studied in Stone [9] using Lagrange multiplier techniques. It is proven (see Chapter 2, mainly Theorems 2.1.2, 2.4.3 and 2.4.4) that the most inviting strategy is indeed uniformly optimal for the model described in this paper, and in fact for more general models as well. Our affirmative results for random N in the continuous search model are summarized in the following theorem.

Theorem 2. *The most inviting strategy f^* with any random N:*
(a) Is uniformly optimal.
(b) Minimizes the posterior expected value of N. Furthermore it provides the highest rate of decrease for the (conditional) expectation of N.
(c) Minimizes the expected cost of searching until an object is found.

Proof. (a) Immediate from the observation and the results for $N = 1$.

(b) We claim here that if $\mathrm{E}_b^{(f)}N$ is the conditional expectation of N following an unsuccessful search with budget b according to the strategy f then

$$\text{for all} \quad b > 0 \quad \mathrm{E}_b^{(f^*)}N \leqslant \mathrm{E}_b^{(f)}N \quad \text{for all } f. \tag{11}$$

In fact

$$\mathrm{E}_b^{(f)}N = \sum_{i=1}^{\infty} nP_{n,b}^{(f)}, \tag{12}$$

where

$$P_{n,b}^{(f)} = \frac{P_n(1-P_{1,b}(f))^n}{\mathrm{E}(1-P_{1,b}(f))^N}$$
$$= P(N=n \mid \text{unsuccessful } f,\ b \text{ search}). \tag{13}$$

So for all f, $P_{n,b}^{(f^*)}/P_{n,b}^{(f)} = [(1-P_{1,b}(f^*))/(1-P_{1,b}(f))]^n$ is non-increasing in n since by the b-optimality of f^* $P_{1,b}(f^*) \geqslant P_{1,b}(f)$. It follows that $\{P_{n,b}^{(f^*)}\}$ is stochastically smaller than or equal to $\{P_{n,b}^{(f)}\}$ implying (11).

Now by (12), (13) and (9) we have

$$\frac{1}{c_i}\frac{\partial}{\partial f(b,i)}\mathrm{E}^{(f)}N = -\frac{\beta_i \pi_i^{(f)}}{c_i}\mathrm{var}^{(f)}(N).$$

Since f^* searches at each instant at box i which maximizes the current value of $\beta_i \pi_i / c_i$, this is also the box which provides the highest decrease of $\mathrm{E}^{(f)}N$. This completes the proof of (b).

(c) Let $B(f)$ be the (random) cost of searching with strategy f until an object is found, then

$$P(B(f) \leqslant b) = P_{N,b}(f),$$

and hence

$$\mathrm{E}B(f) = \int_0^{\infty} (1 - P_{N,b}(f))\,\mathrm{d}b,$$

and this is minimized for f^* since by (a) $P_{N,b}(f^*) \geqslant P_{N,b}(f)$ for all b, for all f. □

References

[1] David Assaf and Shmuel Zamir, "Optimal sequential search: A Bayesian approach", *Ann. Statist.* 3, 1213–1221 (1985).

[2] David Blackwell, "Notes on dynamic programming", Unpublished Notes, University of California, Berkeley, CA, 1962.

[3] R. Hassin and M. Henig, "Dichotomous search for random objects on an interval", *Math. Oper. Res.* 9, 301–308 (1984).

[4] F.P. Kelley, "An optimal search with unknown detection probabilities", *Y. Math. Anal.* **88**, 422–432 (1982).

[5] F.P. Kelley, "A remark on search and sequencing problem", *Math. Oper. Res.* 7, 154–157 (1982).

[6] Ariela Sharlin, "Optimal search for one of many objects hidden in two boxes", Discussion Paper, Department of Statistics, The Hebrew University of Jerusalem, 1985.

[7] Furman H. Smith and George Kimeldorf, "Discrete sequential search for one of many objects", *Ann. Statist.* **3**, 906–915 (1975).

[8] Furman H. Smith and George Kimeldorf, "Binomial searching for a random number of multinomially hidden objects", *Manag. Sci.* 25, 1115–1126 (1979).

[9] Lawrence Stone, *Theory of Optimal Search*, Academic Press, New York, 1975.