A WEAKEST LINK MARKED STOPPING PROBLEM

STEVEN A. LIPPMAN,* *University of California, Los Angeles*

SHELDON M. ROSS,** *University of Southern California*

SRIDHAR SESHADRI,*** *New York University*

Abstract

We consider a model in which we have k items to be sold. Potential buyers make offers in a sequential fashion. Once made, the offer is either rejected or marked for acceptance. Once k items have been marked, the items are then sold to the buyers whose offers were marked, but at a price equal to the minimum of the k marked offers. Assuming that the successive offers are independent and identically distributed according to a specified distribution and that there is a fixed cost incurred whenever an offer is rejected, we determine structural results about the optimal policy, present computational approaches for finding the optimal policy, and give some heuristic policies.

Keywords: Marked stopping problem; weakest link; structure of optimal policy; online heuristic

2000 Mathematics Subject Classification: Primary 90C39
Secondary 90C40

1. Introduction

Suppose that independent random variables $X_1, X_2, \ldots$, having a common distribution function F, are observed in sequence. Instantly upon observation, each random variable is either rejected or marked by the decision maker. The problem ends when k observations have been marked, and a return equal to the minimum of the k marked values is then received. In addition, a cost c is assumed to be incurred each time an offer is rejected. The problem is to find the strategy that maximizes the expected net return.

One application for this problem arises when a team of k units has to be assembled. Units are observed sequentially, with each unit having a value. Marking a unit means that it is accepted to the team. The problem ends when k units have been marked and the payoff is the minimum of the marked values. Other applications come from problems in which you have multiple items to sell. For instance, suppose that you have k items to sell, and that offers to buy a single item arrive sequentially. Marking a buyer's offer means that you will sell one of your items to that buyer. However, you must charge the same amount to each buyer, namely the minimum of the marked offers. There is also a positive cost c incurred each time an offer is rejected. Thus, this fits the framework of our problem, with X_i equal to offer i multiplied by k. Another possibility is that you have $k - 1$ items to sell but you only sell when there have been k marked offers and the price charged for all $k - 1$ items is the smallest of the k marked offers. Here X_i would

equal offer i multiplied by $k - 1$. This model has similarities to Vickrey auctions; see [6] and also [4, Chapter 2]. Other examples of multiple marking models appear in [1], [2], and [3].

A variation of this problem that will also be considered, referred to as model 2 in the sequel, is one is which the values come in as before but there is no marking. Rather, at some point, after you have already observed at least k values, you stop. The payoff when you stop is the kth highest value observed minus nc, where now n denotes the number of values that were observed before you stopped. The main difference between this and the first model is that in the first model you must tell each observed unit whether you will accept it or not, whereas in the latter model you only select the units when you stop.

2. Structural results for model 1

To begin let us note that because the successive offers are independent and identically distributed it follows, from standard dynamic programming results, that the optimal decision as to whether to mark any new offer depends only on the value of the current offer, the minimal of the so-far marked offers, and the number of offers that still remain to be marked.

Let $V_j(x)$ denote the maximal expected additional return given that x is the minimal marked offer and j offers still remain to be marked. Also, let $V_j(x, y)$ denote the maximal expected additional return given that x is the minimal marked offer, j offers still remain to be marked, and the value of the next offer is y. Both x and y are allowed to range between $-\infty$ and ∞.

The optimality equations for this model are as follows.

$$V_j(x, y) = \max\{V_{j-1}(\min(x, y)), V_j(x) - c\}, \qquad j > 0,$$

$$V_j(x) = \int_{-\infty}^{\infty} V_j(x, y) \, dF(y), \qquad j > 0,$$

$$V_0(x) = x.$$

The following lemma is rather apparent.

Lemma 1. $V_j(x)$ and $V_j(x, y)$ are both nondecreasing in x and nonincreasing in j.

Proof. Let $w < x$, and suppose that the current minimal marked offer is x. Marking future offers as if this value were w rather than x, the net return for any policy will be at least as large as it would have been if w were the minimal marked offer, showing that $V_j(x)$ and $V_j(x, y)$ are both increasing in x. Similarly, if there are j additional items to be marked then, marking as if there were $j + 1$, the return from any policy will be at least as large as what would be obtained if $j + 1$ values still had to be marked.

It follows, from the preceding lemma and the optimality equations, that it is optimal to mark offer y when in state j, x, y if $y \geq x$. Indeed, the optimal policy is to mark offer y if

$$V_{j-1}(y) \geq V_j(x) - c.$$

Letting

$$y_j(x) = \min\{y \colon V_{j-1}(y) \geq V_j(x) - c\},$$

it follows, from the monotonicity of $V_{j-1}(y)$, that the optimal policy marks offer y when in state j, x, y if $y > y_j(x)$, does not mark y if $y < y_j(x)$, and is indifferent about marking y if $y = y_j(x)$. (The critical value $y_j(x)$ can be nonpositive even when all offers are positive.)

Proposition 1. $y_j(x)$ *is nondecreasing in* x.

Proof. Suppose that $x_1 > x_2$. To show that $y_j(x_1) \geq y_j(x_2)$, we must show that if $V_{j-1}(y) \geq V_j(x_1) - c$ then $V_{j-1}(y) \geq V_j(x_2) - c$. But, the preceding follows because $V_j(x)$ is a nondecreasing function of x.

Proposition 2. $y_j(x)$ *is nonincreasing in* j.

Proof. To begin note that

$$V_j(x) = \int V_j(x, y) \, dF(y)$$

$$= (V_j(x) - c) F(y_j(x)) + \int_{y_j(x)}^x V_{j-1}(y) \, dF(y) + V_{j-1}(x) \bar{F}(x),$$

where $\bar{F}(x) = 1 - F(x)$. Integrating the right-hand side of the preceding by parts ($u = -V_{j-1}(y)$ and $dv = d\bar{F}(y)$) yields

$$V_j(x) = -V_{j-1}(x) \bar{F}(x) + V_{j-1}(y_j(x)) \bar{F}(y_j(x)) + \int_{y_j(x)}^x \bar{F}(y) V_{j-1}'(y) \, dy$$

$$+ V_{j-1}(x) \bar{F}(x) + (V_j(x) - c) F(y_j(x))$$

$$= V_{j-1}(y_j(x)) \bar{F}(y_j(x)) + \int_{y_j(x)}^x \bar{F}(y) V_{j-1}'(y) \, dy + (V_j(x) - c) F(y_j(x)).$$

However, because the optimal policy is indifferent between marking or rejecting when the offer is identically $y_j(x)$,

$$V_{j-1}(y_j(x)) = V_j(x) - c. \tag{1}$$

Therefore,

$$c = \int_{y_j(x)}^x \bar{F}(y) V_{j-1}'(y) \, dy$$

$$= \int_0^x \bar{F}(y) V_{j-1}'(y) \, dy - \int_0^{y_j(x)} \bar{F}(y) V_{j-1}'(y) \, dy. \tag{2}$$

Differentiation now yields

$$0 = \bar{F}(x) V_{j-1}'(x) - \bar{F}(y_j(x)) V_{j-1}'(y_j(x)) y_j'(x). \tag{3}$$

However, differentiating (1) shows that

$$V_{j-1}'(y_j(x)) y_j'(x) = V_j'(x),$$

which, in conjunction with (3), yields

$$\bar{F}(x) V_{j-1}'(x) = \bar{F}(y_j(x)) V_j'(x). \tag{4}$$

Because $\bar{F}$ is nonincreasing and $y_j(x) \leq x$, it follows, from (4), that

$$V_{j-1}'(x) \geq V_j'(x). \tag{5}$$

However, from (2) we have

$$c = \int_{y_j(x)}^x \bar{F}(y) V_{j-1}'(y) \, dy = \int_{y_{j+1}(x)}^x \bar{F}(y) V_j'(y) \, dy,$$

which, using (5), yields the result

$$y_j(x) \geq y_{j+1}(x).$$

3. Computations for model 1

Consider the problem starting at the point where there are still j offers that need to be marked and x is the smallest of the offers that have been marked. Now regard this as a stopping problem that stops when the next offer is marked, with a final return of $V_{j-1}(\min(x, y))$ obtained if the marked value is y. Because the optimal policy need never recall a previously unmarked offer (received in state (j, x)), we can let the state of this stopping problem be the maximum of all the received offers. Hence, with $M_x = \min(x, X)$, the one stage look-ahead policy (see [5, pp. 54–55]) would mark an offer y if and only if

$$V_{j-1}(y) \geq F(y)V_{j-1}(y) + \mathrm{E}[V_{j-1}(M_x) \mid X > y]\bar{F}(y) - c.$$

That is, it would stop if $y \geq x$ or if $y < x$ and

$$c \geq \mathrm{E}[V_{j-1}(M_x) - V_{j-1}(y) \mid X > y]\bar{F}(y).$$

Because of the monotonicity of $V_{j-1}(y)$, the preceding implies that the one stage look-ahead policy would mark an offer y if and only if

$$c \geq \mathrm{E}[(V_{j-1}(M_x) - V_{j-1}(y))^+].$$

Because the state variable y cannot decrease (as recall of old offers is allowed) and $V_{j-1}(y)$ is increasing in y, it follows that the set of stopping states of the one stage look-ahead policy is a closed set, implying that it is the optimal policy.

Now suppose that X is a finite, integer valued random variable with $P_i = \mathrm{P}\{X = i\}$, $\sum_{i=0}^{m} P_i = 1$. Letting

$$H_{j-1}(i : x) = \mathrm{E}[(V_{j-1}(M_x) - V_{j-1}(i))^+]$$

$$= \sum_{k=i}^{x} P_k(V_{j-1}(k) - V_{j-1}(i)) + (V_{j-1}(x) - V_{j-1}(i)) \sum_{k=x+1}^{m} P_k,$$

it follows that

$$y_j(x) = \min\{i : c \geq H_{j-1}(i : x)\}.$$

The search for the minimal value of i in the preceding is facilitated by using the fact that $H_{j-1}(i)$ is nonincreasing in i.

Once $y_j(x)$ is determined, $V_j(x)$ can be obtained from

$$V_j(x) = \mathrm{E}[V_{j-1}(M_x) \mid X \geq y_j(x)] - c\frac{\mathrm{P}\{X < y_j(x)\}}{\mathrm{P}\{X \geq y_j(x)\}}.$$

That is, starting with

$$y_1(x) = \min\left\{i : c \geq \sum_{k=i}^{x}(k - i)\,P_k + (x - i)\sum_{k=x+1}^{m} P_k\right\},$$

we can determine

$$V_1(x) = \mathrm{E}[M_x \mid X \geq y_1(x)] - c\frac{\mathrm{P}\{X < y_1(x)\}}{\mathrm{P}\{X \geq y_1(x)\}}$$

for each value of x, then use this to determine the function $y_2(x)$, and so on. This continues until we have the values of $V_{n-1}(x)$ which we then use to find $V_n(\infty)$. Computations can be reduced by using the fact that $y_j(x)$ is nondecreasing in x and nonincreasing in j.

4. A heuristic online algorithm for model 1

We can obtain a lower bound on $V_i(a)$ by considering the policy which states that when i offers still remain to be marked, with a the minimal of the already marked offers, all future offers will be accepted provided that they are at least b, for some specified $b \leq a$. Using the fact that the expected value of a nonnegative random variable is the integral of its tail-distribution function, it is easy to see that $g_i(a, b)$, the expected additional net return of this policy, is

$$g_i(a, b) = b + \int_b^a \left(\frac{\bar{F}(t)}{\bar{F}(b)} \right)^i dt - ic \frac{P\{X < b\}}{P\{X \geq b\}}.$$

When $i = k$, set $a = \infty$. With

$$g_i(a) = \sup_{b \leq a} g_i(a, b),$$

it follows that $g_i(a) \leq V_i(a)$. Now, whereas it is optimal to reject offer $y < x$ when in state j, x, y if and only if

$$V_{j-1}(y) \leq V_j(x) - c,$$

we propose the online heuristic policy that rejects such an offer if and only if $y < x$ and $g_{j-1}(y) \leq g_j(x) - c$. We refer to it as an online policy because it can be implemented by computing only those values of g_i that are needed.

5. Model 2

In model 2, after observing at least k offers, it is allowable to stop and receive a final reward equal to the kth largest value so far observed. In addition, a cost c is incurred each time a new offer is observed, and the objective is to maximize the total expected net return. The state of this dynamic programming problem is the vector $\boldsymbol{x} = (x_k, x_{k-1}, \ldots, x_1)$, where $x_k \leq x_{k-1} \leq \cdots \leq x_1$ are the k largest offers yet received in increasing order. With $V(\boldsymbol{x})$ defined to be the maximal expected net return from then on, given that the current state is $\boldsymbol{x}$, the optimality equation is

$$V(\boldsymbol{x}) = \max \left\{ x, F(x_k) V(\boldsymbol{x}) + \int_{x_k}^{\infty} V(\mathbf{o}(\{y, x_{k-1}, \ldots, x_1\})) \, dF(y) - c \right\},$$

where $\mathbf{o}(\{y, x_{k-1}, \ldots, x_1\})$ is the order statistics of the set of values $\{y, x_{k-1}, \ldots, x_1\}$.

We can prove the following result exactly as we proved Lemma 1.

Lemma 2. *If $\boldsymbol{x} = (x_k, \ldots, x_1)$ and $\boldsymbol{y} = (y_k, \ldots, y_1)$ are states such that $y_i < x_i$ for each i then $V(\boldsymbol{y}) \leq V(\boldsymbol{x})$.*

Proposition 3. *If it is optimal to stop when in state $\boldsymbol{x} = (x_k, x_{k-1}, \ldots, x_1)$ then it is optimal to stop when in state $\boldsymbol{x}' = (x, x_{k-1}, \ldots, x_1)$ when $x_k \leq x \leq x_{k-1}$.*

Proof. Suppose that the initial state is $\boldsymbol{x}$, and consider any policy π. Throughout act as if the initial state were $\boldsymbol{x}'$ rather than $\boldsymbol{x}$, and follow policy π. The net return when you stop will then be within $x - x_k$ of what it would have been if the actual initial state were $\boldsymbol{x}'$, showing that $V(\boldsymbol{x}) \geq V_{\pi}(\boldsymbol{x}') - (x - x_k)$, where V_{π} signifies the expected return when policy π is employed. Because this is true for all policies π, it follows that

$$V(\boldsymbol{x}) \geq V(\boldsymbol{x}') - (x - x_k).$$

Because it is optimal to stop in state x, $V(x) = x_k$, showing that

$$V(x') \leq x.$$

Thus, by the optimality equation, $V(x') = x$ and the result is proven.

Corollary 1. *If it is optimal to stop when in state $x = (x_k, x_{k-1}, \ldots, x_1)$ then it is optimal to stop when in state $y = (y_k, y_{k-1}, \ldots, y_1)$ when $y_k \geq x_k$ and $y_j \leq x_j$, $j \neq k$.*

Proof. Suppose that it is optimal to stop when in state $x = (x_k, x_{k-1}, \ldots, x_1)$. Then, by Proposition 3, it is also optimal to stop when in state $(y_k, x_{k-1}, \ldots, x_1)$, showing that

$$V(y_k, x_{k-1}, \ldots, x_1) = y_k.$$

The result now follows from Lemma 2 as

$$V(y) \leq V(y_k, x_{k-1}, \ldots, x_1) = y_k.$$

Corollary 2. *Assuming that $\mathrm{E}[X] > c$, let $\underline{v}$ be such that*

$$c = \mathrm{E}[(X - \underline{v})^+].$$

Then it is optimal to stop in any state x having $x_k \geq \underline{v}$.

Proof. If the $k - 1$ highest offers are all equal to ∞ then the problem reduces to the classical one item problem. Thus, using the result from this problem (see [5, pp. 55–56]), it is optimal to stop if $x_k \geq \underline{v}$. It now follows, from Corollary 1, that it is optimal to stop in any state x having $x_k \geq \underline{v}$.

Corollary 3. *With*

$$\bar{v} = \sup_y \left\{ y + \int_y^\infty \left(\frac{\bar{F}(t)}{\bar{F}(y)} \right)^k \mathrm{d}t - \frac{kc}{\bar{F}(y)} \right\},$$

it is optimal to continue in any state x having $x_k \leq \bar{v}$.

Proof. Let V denote the maximal expected net return before any offers have been received. Then it is clear that it is optimal to continue in any state x having $x_k \leq V$. (If the state is $(x, x, \ldots, x)$ then it is optimal to continue if and only if $x \leq V$.) With $R(y)$ equal to the expected net return from the policy that stops when there are k offers that all exceed y,

$$R(y) = y + \int_y^\infty \left(\frac{\bar{F}(t)}{\bar{F}(y)} \right)^k \mathrm{d}t - \frac{kc}{\bar{F}(y)}.$$

Because $V \geq \sup_y R(y)$, the result follows.

Example. Suppose that F is the exponential distribution with mean 5, $c = 0.01$, and $k = 5$. Then

$$0.01 = \mathrm{E}[(X - \underline{v})^+] = 5\mathrm{e}^{-\underline{v}/5}.$$

Thus, $\underline{v} = 5 \ln(500) \approx 31.05$. Also,

$$R(y) = y + 1 - 0.05\mathrm{e}^{y/5},$$

yielding $\bar{v} = 5 \ln(100) - 4 \approx 19.05$.

5.1. A heuristic online algorithm for model 2

Whereas the actual optimal policy is computationally difficult to determine (see Section 5.2 for an exact algorithm for a special case), we suggest the following heuristic policy. Suppose that the current state is x. For a specified j, $j = 1, \ldots, k$, consider the policy that continues until the k largest values are all greater than x_j. Then, with $x_0 = \infty$, the expected net return from this point on, call if $E_j(x)$, is given by

$$E_j(x) = x_j + \int_{x_j}^{x_{j-1}} \left(\frac{\bar{F}(t)}{\bar{F}(x_j)} \right)^{k-j+1} \mathrm{d}t - \frac{c(k-j+1)}{\bar{F}(x_j)}. \tag{6}$$

The heuristic rule we suggest is as follows. Stop in state x if and only if

$$x_k > \max_{j=1,\ldots,k} E_j(x).$$

5.2. The special case, $k = 2$

Now consider the important special case, $k = 2$. For instance, you might have a single item to sell, but when you stop you must sell it for the second highest offer.

Let V denote the maximal expected net return; let $V(y)$ denote the maximal expected net return when a single unit has been observed and its value is y. Say that the state is (z, y), $z \le y$, when the two highest values yet observed are z and y, and let $V(z, y)$ denote the maximal expected additional net return when the state is (z, y).

It follows, from Corollary 1, that if it is optimal to continue when in state (z, z) then it is optimal to continue in state (z, y). Thus, being in state (z, z) is like beginning the problem, from which we can conclude that

$$V(z, z) = \begin{cases} V & \text{if } z \le V, \\ z & \text{if } z > V. \end{cases}$$

By monotonicity, we have

$$V(z, y) = V, \qquad z \le y \le V.$$

If $y > V$ then because it is optimal to stop when in state (y, y) it follows, from Corollary 1, that there exists a critical value $a(y) \le y$ such that it is optimal to continue in state (z, y), $z \le y$, if $z \le a(y)$ and optimal to stop if $z \ge a(y)$. Consequently, for $y > V$,

$$V(z, y) = \begin{cases} V(y) & \text{if } z \le a(y), \\ z & \text{if } z \ge a(y). \end{cases}$$

Because the optimal policy is indifferent between stopping and continuing when in state $(a(y), y)$,

$$a(y) = V(a(y), y) = V(y).$$

Now, recall that $\underline{v}$, satisfying

$$c = \mathrm{E}[(X - \underline{v})^+],$$

is the smallest value v such that it is optimal to stop in state (v, ∞). We consider two cases.

Case 1. The state is (z, y), where $y \ge \underline{v}$.

Let $a = a(y)$ and $\bar{F} = 1 - F$. Using the facts that the optimal policy is indifferent between stopping or continuing when in state (a, y) and that, because $y \geq \underline{v}$, it will stop when in state (y, s), we obtain

$$a = -c + F(a)a + \int_a^y s\,\mathrm{d}F(s) + \bar{F}(y)y$$

or

$$a\bar{F}(a) = -c + \int_a^y s\,\mathrm{d}F(s) + \bar{F}(y)y. \tag{8}$$

Numerically solving the preceding for a yields the value of $a = a(y)$ when $y \geq \underline{v}$.

Case 2. The state is (z, y), where $V < y < \underline{v}$.

For Case 2,

$$a = -c + F(a)a + \int_a^y s\,\mathrm{d}F(s) + \int_y^\infty V(y, s)\,\mathrm{d}F(s). \tag{9}$$

In this case, we need the values of $V(y, s)$ to obtain $a(y)$. The following bootstrapping procedure overcomes this problem and also provides an insight into the optimal strategy.

Step 0: Choose small $\varepsilon > 0$. Set $k = 1$, $l^0 = \infty$, and $l^1 = \underline{v}$.

Step 1: Use (8) or (9) to determine $a(y)$ for $y \in [l^k, l^{k-1})$.

Step 2: Determine only the *new* values of $V(y, s)$, $y \in [a(l^k), \infty)$, $s \geq y$.

Step 3: If $l^k - a(l^k) < \varepsilon$ stop, otherwise set $k = k + 1$ and $l^k = a(l^{k-1})$. Proceed to step 1.

We can show that l^k is a monotone decreasing sequence by differentiating (9) with respect to y and rearranging to obtain

$$\bar{F}(a(y))a'(y) = F(a^{-1}(y)) - F(y).$$

Thus, the derivative of $a(y)$ is strictly less than 1. Letting $a^n(y) = a(a(\cdots(y)\cdots))$, we see that a^n converges to a limit that is reached in step 3.

6. Numerical results

In this section we present results for the special case, $k = 2$, for both models. We compare the performance of the online algorithms and also compare the two models.

For both models, an alternate method to determine V and the optimal control is to use successive approximations. To implement this, we need to discretize the state space and also assume that F has a compact support. The state can be conveniently represented as the set of numbers $S = \{1, 2, \ldots, K\}$. We implemented the exact methods proposed above and a successive approximation method in order to carry out a comparison between the results as well as to estimate the rate of convergence of the latter method. We show results for the uniform distribution over $[0, 100]$ discretized in step sizes of 1 and the truncated normal distribution on $[0, 100]$ with mean equal to 50 and standard deviation equal to 16.7. The value of c was set equal to 5, 10, 15, and 20. For both models, the successive approximations method converged, that is, the value function converged to within five decimal places, within 20 iterations for $c = 5$, and within 10 iterations for $c = 20$. The value determined by both methods agreed.

TABLE 1: The values of V, the lower bound, and the value produced by the online algorithm for model 1.

c	Uniform			Normal (truncated)		
	V	Lower bound	Online	V	Lower bound	Online
5	59.29	58.86	58.56	48.46	48.09	47.56
10	48.09	47.47	46.29	44.02	43.70	42.83
15	41.81	41.06	38.92	42.22	41.99	41.23
20	38.09	37.22	35.17	41.39	41.24	40.77

TABLE 2: The values of V, the lower bound, and the value produced by the online algorithm for model 2.

c	Uniform			Normal (truncated)		
	V	Lower bound	Online	V	Lower bound	Online
5	59.79	58.80	59.75	48.85	40.73	48.41
10	48.78	48.58	48.73	44.36	40.71	43.72
15	42.66	38.38	42.59	42.47	40.72	41.61
20	41.00	41.00	41.00	41.55	41.00	41.00

For model 1, we show the values of V in Table 1 alongside the lower bound produced by g_k and the value produced by the online algorithm. It is clear that the distribution affects the value. Surprisingly, not only is the lower bound given by g_k greater than the value given by the online algorithm, but it is also rather close to the optimal value.

We show the values of V for model 2 in Table 2 alongside the lower bound corresponding to the maximum in (6) and the value produced by the online algorithm. The main observations from viewing these results are that the value of V varies with the distribution, the online algorithm performs better compared to model 1, and the lower bound is poor in this case. Also, even though we might expect V to be higher in model 2, the increase is marginal.

References

[1] BRUSS, F. T., AND FERGUSON, T. S. (1997). Multiple buying or selling with vector offers. *J. Appl. Prob.* **34,** 959–973.

[2] LIPPMAN, S. AND ROSS, S. M. (2007). Variability is beneficial in marked stopping problems. To appear in *Econom. Theory.*

[3] MACQUEEN, J. B. AND MILLER, R. G., JR. (1960). Optimal persistence policies. *Operat. Res.* **8,** 362–380.

[4] MILGROM, P. (2004). *Putting Auction Theory to Work.* Cambridge University Press.

[5] ROSS, S. M. (1983). *Introduction to Stochastic Dynamic Programming.* Academic Press, San Diego, CA.

[6] VICKREY, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *J. Finance* **XVI,** 8–37.