Journal of the Operations Research
Society of Japan
Vol. 41, No. 4, December 1998

A SEARCH GAME WITH REWARD CRITERION

Ryusuke Hohzaki   Koji Iida
National Defense Academy

(Received October 28, 1997; Revised June 25, 1998)

Abstract This paper investigates a search game of a searcher and a target. At the beginning of the search, the target selects his path from some options and the searcher determines the distribution of his available search resources into a search space which consists of discrete cells and discrete time points. The searcher gains a value on detection of the target while he expends the search cost depending on the allocation of the search resource. The payoff of the search is the expected reward which is defined as the expected value minus the expected search cost. The searcher wants to maximize the expected reward and the target wants to minimize it. We formulate the problem as a two-person zero-sum game and reduce it to a concave maximization problem. We propose a computational method to obtain an optimal solution of the game. Our method proceeds in such a way that one-sided problems generated from the original game are repeatedly solved and their solutions converge asymptotically to an optimal solution of the game. By some examples, we examine the effect of parameters included in the problem upon an optimal solution to elucidate some characteristics of the solution and the computational time of the proposed method.

1. Introduction

This paper deals with a two-person zero-sum game of a searcher and a mobile target, in which the target is free to select a path among some options and the searcher distributes search resources over a search space in order to detect the target. A number of papers have been published so far on the subject of the search for a moving target. As an early study in 1970, Pollock[14] investigated a search for a moving target on among two regions in a Markovian fashion. In the studies on the optimal search for a moving target, some varieties of moving patterns of the target are considered, for example, a Markovian process[14, 15, 1, 18], a diffusion process[4], a type of selecting a path from some options[7, 3, 8, 6], a so-called conditionally deterministic motion[16] and so on[17].

Most of papers cited above deal with a one-sided optimizing problem for the searcher. There are some papers in which not only the searcher strategy but also the target strategy is taken as the decision elements. Danskin[2] gives a saddle point between the concrete search pattern and the evasion pattern of the target. Washburn[19] investigates a game in which a searcher and a target select their cells in a region by turns and a pay-off function is given by the expected travelling cost of the searcher until the detection of the target. For the same criterion, Kikuta[10, 11] obtains some rigid results of the optimal solution in a special case where a target hides in a cell and a searcher examines cells in sequence without overlooking possibility. In the above studies, the strategy of the searcher is represented by a sequence of searching points.

There is another type of search games where the searcher distributes a limited amount of search efforts on a search space and his strategy is represented by the distribution plan. Iida, Hohzaki and Furui[9] investigate such a game with the pay-off function of the detection probability, where the target selects his path from some options. They solve the game by making use of the exponential form of the pay-off function and reducing the problem to a linear programming problem. As a special case, Hohzaki and Iida[5] study a game with the expected reward taking account of the target value and the search cost, where a searcher is interested in randomizing his look strategy on a given search path and a target is to select his path. The presented paper is the extension of the above two models and its outline is as follows.

A target selects a path from some options at the beginning of the search. A searcher determines a distribution of limited amount of search resources over search space and time. The search is executed at a finite number of time points by the searcher and terminates on detection of the target or at the expiration of the whole search time. By the detection, the searcher gets a value but expends search cost in proportion to the used search resources. We want to solve the two-person zero-sum game of the target and the searcher with the expected reward criterion which is defined as the expected value minus the expected search cost.

In the next section, we describe assumptions of the search model and obtain a pay-off function of the two-person zero-sum game and in Section 3, the game is formulated as a concave maximization problem. In Section 4, we propose a numerical method to solve the game and give optimal strategies of the searcher and the target. In Section 5, some characteristics of the optimal solution and the computational time of the proposed method are examined by numerical examples.

2. Description of Search Game and Formulation

In this section, we describe some assumptions about a search model in detail and formulate the pay-off function of the game.

(1) A search space consists of a finite number of time points $t \in \boldsymbol{T} = \{1, \cdots, T\}$ and discrete space $j \in \boldsymbol{K} = \{1, \cdots, K\}$ called a set of cells.

(2) A target selects a path, say ω, from a finite number of options Ω to move along. A path ω is defined by a set of cells $\omega = \{\omega(t),\ t \in \boldsymbol{T}\}$, where $\omega(t) \in \boldsymbol{K}$ is the target position at time t. He can not change his path during the search, which means that he can not get any information about his opponent's strategy on the way of the game.

(3) A searcher has available resources $u(t)$ at time t, which can be continuously divided and distributed on arbitrary cells as he likes. If he allocates $\varphi(i,t)$ resources on cell i and the target is there, he can detect the target with probability $1 - \exp(-\alpha_i \varphi(i,t))$ where α_i is positive.

(4) On detection of the target on cell i and at time t, the searcher gains value $V(t) > 0$ but expends search cost $c_0(i,t) > 0$ per unit resource. We assume that $V(t)$ is non-increasing for t.

(5) The search terminates on detection of the target or at the end of time points T whichever earlier. On the termination, the searcher gains the expected reward occurred during the search operation and the target loses the same value.

Now we formulate the search game as a single-stage two-person zero-sum game. First we obtain the pay-off function of the game. A pure strategy of the searcher is a distribution of the search resources $\varphi = \{\varphi(i,t)\ , i \in \boldsymbol{K}\ \ t \in \boldsymbol{T}\}$ and that of the target is a selection of his path $\omega \in \Omega$. For these pure strategies, the cumulative search cost $C(t,\varphi)$ and the detection probability $P_1^t(\varphi,\omega)$ during time period $[1,t]$ are given by the following expressions.

$$C(t,\varphi) = \sum_{\tau=1}^{t}\sum_{i=1}^{K} c_0(i,\tau)\varphi(i,\tau)\ , \tag{2.1}$$

$$P_1^t(\varphi,\omega) = 1 - \exp\left(-\sum_{\tau=1}^{t} \alpha_{\omega(\tau)}\varphi(\omega(\tau),\tau)\right) . \tag{2.2}$$

The probability that the target is detected at time t is $P_1^t(\varphi,\omega) - P_1^{t-1}(\varphi,\omega)$. When the detection does not occur by time T, the searcher must pay $C(T,\varphi)$ without any gain. Therefore, the expected reward $R(\varphi,\omega)$ during the whole search time period $[1,T]$ is calculated as follows.

$$R(\varphi,\omega) = \sum_{t=1}^{T}(V(t) - C(t,\varphi))(P_1^t(\varphi,\omega) - P_1^{t-1}(\varphi,\omega)) - C(T,\varphi)(1 - P_1^T(\varphi,\omega)) \tag{2.3}$$

$$= V(T)P_1^T(\varphi,\omega) + \sum_{t=1}^{T-1}(\Delta C(t,\varphi) - \Delta V(t))P_1^t(\varphi,\omega) - C(T,\varphi) \tag{2.4}$$

where

$$\Delta C(t,\varphi) = C(t+1,\varphi) - C(t,\varphi) = \sum_{i=1}^{K} c_0(i,t+1)\varphi(i,t+1), \tag{2.5}$$

$$\Delta V(t) = V(t+1) - V(t). \tag{2.6}$$

Because the pay-off function $R(\cdot)$ is strictly concave for $\{\varphi(i,t),\ i \in \boldsymbol{K}\ t \in \boldsymbol{T}\}$, we do not need to consider the mixed strategy for the searcher as a solution of the game, which is known by the game theory concerning with the concave-convex game[13]. The mixed strategy of the target is presented by probability $\pi(\omega) \geq 0$ with which the target selects path $\omega \in \Omega$. The relation $\sum_{\omega\in\Omega}\pi(\omega) = 1$ must be satisfied, of course. Hereafter, we use notation ω and π as a specific target path and a specific path selection probability, respectively. When the searcher takes a pure strategy $\{\varphi(i,t)\}$ and the target takes a mixed strategy π, the expected reward of the search $\widetilde{R}(\varphi,\pi)$ is formulated as follows, which is a pay-off of this game.

$$\widetilde{R}(\varphi,\pi) = \sum_{\omega\in\Omega}\pi(\omega)R(\varphi,\omega) . \tag{2.7}$$

Strategies of the target and the searcher have to satisfy the next conditions.

$$\pi(\omega) \geq 0\ ,\quad \omega \in \Omega\ , \tag{2.8}$$

$$\sum_{\omega\in\Omega}\pi(\omega) = 1\ , \tag{2.9}$$

$$\varphi(i,t) \geq 0\ ,\quad i \in \boldsymbol{K}\quad t \in \boldsymbol{T}\ , \tag{2.10}$$

$$\sum_{i=1}^{K}\varphi(i,t) \leq u(t)\ ,\quad t \in \boldsymbol{T}\ . \tag{2.11}$$

We denote the set of feasible solutions of the target satisfying conditions (2.8) and (2.9) by Π and that of the searcher satisfying (2.10) and (2.11) by Φ.

3. Solution of the Game

Since the pay-off is given by Eq.(2.7), the game is solved by finding φ^*, π^* satisfying $\widetilde{R}(\varphi^*,\pi^*) = \min_\pi\max_\varphi\widetilde{R}(\varphi,\pi) = \max_\varphi\min_\pi\widetilde{R}(\varphi,\pi)$ or $\widetilde{R}(\varphi^*,\pi) \geq \widetilde{R}(\varphi^*,\pi^*) \geq \widetilde{R}(\varphi,\pi^*)$ for arbitrary $\pi \in \Pi$ and $\varphi \in \Phi$. Considering that conditions (2.8)~(2.11) give a closed domain, it is evident that the game has a finite real number as a value of the game. The problem $\max_\varphi\min_\pi\widetilde{R}(\varphi,\pi)$ is transformed as follows.

$$\max_{\varphi\in\Phi}\min_{\pi\in\Pi}\widetilde{R}(\varphi,\pi)$$

$$= \max_{\varphi\in\Phi}\min_{\pi\in\Pi}\sum_{\omega\in\Omega}\pi(\omega)R(\varphi,\omega) = \max_{\varphi\in\Phi}\min_{\omega\in\Omega}R(\varphi,\omega) = \max_{\varphi\in\Phi}\{\nu|\ R(\varphi,\omega) \geq \nu,\ \omega \in \Omega\}\ .$$

In the above transformation, an optimal π^* is given by $\pi^*(\omega) = 0$ for $\{\omega|R(\omega,\varphi) > \nu = \min_{\omega'\in\Omega}R(\omega',\varphi)\}$. In the result, we can formulate the game as a concave maximization problem (P0).

$$(P0) \quad \max_{\varphi} \ \nu \tag{3.1}$$

$s.t.$

$$R(\varphi,\omega) \geq \nu, \ \omega \in \Omega \ , \tag{3.2}$$

$$\varphi(i,t) \geq 0 \ , \quad i \in \boldsymbol{K} \ \ t \in \boldsymbol{T} \ , \tag{3.3}$$

$$\sum_{i=1}^{K} \varphi(i,t) \leq u(t) \ , \quad t \in \boldsymbol{T} \ . \tag{3.4}$$

We should note that the problem does not contain $\pi(\omega)$. We take $\pi(\omega)$, $\mu(i,t)$ and $\lambda(t)$ as Lagrangean multipliers corresponding to conditions (3.2), (3.3) and (3.4), respectively. Afterward, we will state the valid reason by that the multiplier $\pi(\omega)$ can be regarded as the target's mixed strategy. Using these multipliers, we define a Lagrangean function $L(\nu,\varphi;\pi,\lambda,\mu)$ as follows.

$$\begin{aligned} L(\nu,\varphi;\pi,\lambda,\mu) &= \nu + \sum_{\omega\in\Omega} \pi(\omega)(R(\varphi,\omega)-\nu) + \sum_{t} \lambda(t)(u(t) - \sum_{i} \varphi(i,t)) \\ &\quad + \sum_{t}\sum_{i} \mu(i,t)\varphi(i,t) \ . \end{aligned} \tag{3.5}$$

Using this function, an optimal solution of problem (P0) has the next necessary and sufficient conditions, so-called Kuhn-Tucker conditions.

$$\pi(\omega)(R(\varphi,\omega)-\nu) = 0 \ , \quad \omega \in \Omega \ , \tag{3.6}$$

$$R(\varphi,\omega) \geq \nu, \ \omega \in \Omega \ , \tag{3.7}$$

$$\pi(\omega) \geq 0 \ , \ \omega \in \Omega \ , \tag{3.8}$$

$$\frac{\partial L}{\partial \nu} = 1 - \sum_{\omega\in\Omega} \pi(\omega) = 0 \ , \tag{3.9}$$

$$\sum_{i=1}^{K} \varphi(i,t) \leq u(t) \ , \quad t \in \boldsymbol{T} \ , \tag{3.10}$$

$$\varphi(i,t) \geq 0 \ , \quad i \in \boldsymbol{K} \ \ t \in \boldsymbol{T} \ , \tag{3.11}$$

$$\lambda(t) \geq 0, \ \mu(i,t) \geq 0 \ , \quad i \in \boldsymbol{K} \ \ t \in \boldsymbol{T} \ , \tag{3.12}$$

$$\frac{\partial L}{\partial \varphi(i,t)} = \sum_{\omega\in\Omega} \pi(\omega) \frac{\partial R(\varphi,\omega)}{\partial \varphi(i,t)} - \lambda(t) + \mu(i,t) = 0 \ , \quad i \in \boldsymbol{K} \ \ t \in \boldsymbol{T} \ , \tag{3.13}$$

$$\lambda(t)\left(u(t) - \sum_{i=1}^{K} \varphi(i,t)\right) = 0 \ , \quad t \in \boldsymbol{T} \ , \tag{3.14}$$

$$\mu(i,t)\varphi(i,t) = 0 \ , \quad i \in \boldsymbol{K} \ \ t \in \boldsymbol{T} \ . \tag{3.15}$$

As known from conditions (3.6), (3.8) and (3.9), an optimal multiplier $\pi(\omega)$ satisfies the same conditions as an optimal mixed strategy of the target must do, that is, $\pi(\omega) \geq 0$, $\sum_{\omega\in\Omega} \pi(\omega) = 1$, $\pi(\omega) = 0$ for $\{\omega | R(\omega,\varphi) > \nu\}$. Furthermore, the summation of the first and second terms of the Lagrangean function $L(\cdot)$ becomes $\sum_{\omega} \pi(\omega) R(\varphi,\omega) = \widetilde{R}(\varphi,\pi)$ for an optimal multiplier π. By this reason, we use variable $\{\pi(\omega)\}$ originally denoting the target mixed strategy as the Lagrangean multiplier. Considering these discussions, we notice that expressions (3.10)$\sim$(3.15) give the necessary and sufficient conditions of an optimal solution φ^*_π for a problem $\max_\varphi \widetilde{R}(\varphi,\pi)$, that is, a problem of maximizing the expected reward given the probability of the target path selection π. This kind of one-sided optimizing problem has been already studied and an algorithm to give an optimal solution is proposed by Iida and Hohzaki[8]. The above discussion brings us an idea to solve our game by such a way that we repeatedly solve the problem $\max_\varphi \widetilde{R}(\varphi,\pi)$ while varying π until Eqs. (3.6) and (3.7) hold. We will state a computational algorithm using the idea in the next section.

4. Algorithm to Solve the Game

We state some useful lemmas before the proposition of an algorithm of solving the game.

Lemma 1 *A problem* $\max_{\varphi} \widetilde{R}(\varphi, \pi)$ *has the same optimal solution if ratios of* $\pi(\omega)$ *remain unchanged among* $\omega \in \Omega$.

The validity of this lemma is self-evident because $\widetilde{R}(\varphi, \pi)$ is linear for $\{\pi(\omega)\}$. For a positive real number β, we have $\max_{\varphi} \widetilde{R}(\varphi, \beta\pi) = \beta \max_{\varphi} \widetilde{R}(\varphi, \pi)$ and two maximizing problems have the same solution. This means that flexible methods are possible for solving $\max_{\varphi} \widetilde{R}(\varphi, \pi)$. That is, an optimal solution φ^* for $\max_{\varphi} \widetilde{R}(\varphi, \pi)$, where non-negative $\{\pi(\omega)\}$ does not necessarily satisfy condition (2.9), stays optimal for $\max_{\varphi} \widetilde{R}(\varphi, \hat{\pi})$ with the following $\{\hat{\pi}(\omega)\}$,

$$\hat{\pi}(\omega) = \pi(\omega) / \sum_{\omega' \in \Omega} \pi(\omega') , \quad \omega \in \Omega \tag{4.1}$$

which satisfies condition (2.9). Thus condition (2.9) is not so tight for our algorithm in a sense that we can change π regardless of their summation. Whenever necessary, we can normalize $\{\pi(\omega)\}$ by (4.1).

We denote an optimal solution for π by φ^*_{π}, that is, $\widetilde{R}(\varphi^*_{\pi}, \pi) = \max_{\varphi} \widetilde{R}(\varphi, \pi)$. Similarly, we use φ^*_{ω} as an optimal searcher's strategy of maximizing the expected reward for a specific target path ω, that is, $R(\varphi^*_{\omega}, \omega) = \max_{\varphi} R(\varphi, \omega)$. This solution can be obtained from $\max_{\varphi} \widetilde{R}(\varphi, \pi)$ in the case of $\pi(\omega) = 1,\ \pi(\omega') = 0,\ \forall \omega' \neq \omega$.

The next lemma deals with the change of $\pi(k)$ for a specific target path k.

Lemma 2 *Modify* $\{\pi(\omega)\}$ *to* $\{\pi'(\omega)\}$ *as follows,*

$$\pi'(\omega) = \begin{cases} \pi(k) + \Delta\pi(k), & \omega = k \\ \pi(\omega), & \forall \omega \neq k \end{cases} .$$

Then the expected reward for the path k *increases if* $\Delta\pi(k) > 0$ *and decreases if* $\Delta\pi(k) < 0$, *that is,*

$$if \ \ \Delta\pi(k) > 0, \quad R(\varphi^*_{\pi'}, k) \geq R(\varphi^*_{\pi}, k) , \tag{4.2}$$

$$if \ \ \Delta\pi(k) < 0, \quad R(\varphi^*_{\pi'}, k) \leq R(\varphi^*_{\pi}, k) . \tag{4.3}$$

Proof: The next relation holds.

$$\begin{aligned} \widetilde{R}(\varphi^*_{\pi}, \pi') &= \sum_{\omega} \pi(\omega) R(\varphi^*_{\pi}, \omega) + \Delta\pi(k) R(\varphi^*_{\pi}, k) = \widetilde{R}(\varphi^*_{\pi}, \pi) + \Delta\pi(k) R(\varphi^*_{\pi}, k) \\ &\leq \widetilde{R}(\varphi^*_{\pi'}, \pi') = \widetilde{R}(\varphi^*_{\pi'}, \pi) + \Delta\pi(k) R(\varphi^*_{\pi'}, k) . \end{aligned}$$

Therefore,

$$0 \leq \widetilde{R}(\varphi^*_{\pi}, \pi) - \widetilde{R}(\varphi^*_{\pi'}, \pi) \leq \Delta\pi(k)(R(\varphi^*_{\pi'}, k) - R(\varphi^*_{\pi}, k)) .$$

Here, we complete the proof. **Q.E.D.**

From this lemma, we know that the expected reward of a specific target path could be under our control by varying of $\{\pi(\omega)\}$ to a certain extent. As a stopping rule of our computational algorithm, we can make use of the following lemma.

Lemma 3 *Assume that, for some target paths* $\omega_1, \omega_2, \cdots, \omega_s \in \Omega$ *and a real number* ν, *the next conditions are satisfied.*

$$R(\varphi^*_{\pi^*}, \omega_1), \cdots, \ R(\varphi^*_{\pi^*}, \omega_s) > \nu , \tag{4.4}$$

$$R(\varphi^*_{\pi^*}, \omega) = \nu, \ \forall \omega \in \Omega / \{\omega_1, \cdots, \omega_s\} , \tag{4.5}$$

$$\pi^*(\omega) = 0, \ \omega \in \{\omega_1, \cdots, \omega_s\} . \tag{4.6}$$

Then $\pi^* \in \Pi$, $\varphi^*_{\pi^*} \in \Phi$ *is an optimal solution of the game and the value of the game is* ν.

Proof: Using $\Pi' = \{\pi \in \Pi |\ \pi(\omega) = 0,\ \omega \in \{\omega_1, \cdots, \omega_s\}\}$, we have $\widetilde{R}(\varphi^*_{\pi^*}, \pi) = \nu$ for $\pi \in \Pi'$ and $\widetilde{R}(\varphi^*_{\pi^*}, \pi) > \nu$ for $\pi \in \Pi / \Pi'$. Therefore, $\widetilde{R}(\varphi^*_{\pi^*}, \pi) \geq \nu,\ \forall\ \pi \in \Pi$. Furthermore, from the optimality of $\varphi^*_{\pi^*}$ for π^*, $\tilde{R}(\varphi^*_{\pi^*},\pi^*) \geq \tilde{R}(\varphi,\pi^*)$, $\forall\ \varphi \in \Phi$. Consequently, we have the next inequality for arbitrary $\varphi \in \Phi$ and $\pi \in \Pi$.

$\tilde{R}(\varphi^*_{\pi^*},\pi) \geq \nu = \tilde{R}(\varphi^*_{\pi^*},\pi^*) \geq \tilde{R}(\varphi,\pi^*)$.

That indicates that $\{\varphi^*_{\pi^*},\pi^*\}$ is an optimal solution of the game. **Q.E.D.**

Now, we are ready to explain an algorithm for computing the optimal strategies of the game. Its essence is to find an optimal searcher's strategy φ^*_π while varying π and make them converge to a solution of the game, which is accomplished when Lemma 3 is satisfied. Every time the optimization of the searcher's strategy φ for an π is terminated, we sort $\{R(\varphi,\omega),\ \omega \in \Omega\}$ in the order of value. Then, we increase $\pi(\omega)$ for path ω with small $R(\varphi,\omega)$ in order to enlarge the expected reward of the path ω and decrease positive $\pi(\omega)$ for path ω with large $R(\varphi,\omega)$ within non-negative value in order to make $R(\varphi,\omega)$ smaller, which is expected by Lemma 2. While we increase or decrease $\pi(\omega)$, it is not necessary to consider condition $\sum_\omega \pi(\omega) = 1$ as seen from Lemma 1. If we can make $\pi(\omega) = 0$ for ω with larger $R(\varphi,\omega)$ and make $R(\varphi,\omega)$ have the same value for ω with $\pi(\omega) > 0$, then we will have obtained an optimal solution of the game from Lemma 3. Our algorithm to give a solution of the game is as follows.

(Step1) Initialize $\{\pi(\omega),\ \omega \in \Omega\}$, e.g. uniform distribution $\{\pi(\omega)\} = \{1/|\Omega|\}$. Set $l = 0$.

(Step2) For π, solve the one-sided problem $\max_\varphi \tilde{R}(\varphi,\pi)$ and find its optimal solution φ^*_π by Iida and Hohzaki's method[8].

(Step3) Sort values of $\{R(\varphi^*_\pi,\omega),\ \omega \in \Omega\}$ in the order $W_1 < W_2 < \cdots < W_M$, where W_k denotes the k-th smallest value, and classify $\omega \in \Omega$ according to the value. Assume that a subset $\Omega_1 \subseteq \Omega$ has the smallest expected reward W_1, Ω_2 the second smallest one W_2 and so on, and Ω_M has the largest one W_M, namely, $\Omega_j = \{\omega |\ R(\varphi^*_\pi,\omega) = W_j\}$, $j = 1,\cdots,M$.
If Lemma 3 holds, the algorithm terminates. Current φ^*_π and π normalized by formula (4.1) are the optimal searcher's strategy and target's strategy of the game, respectively.
Otherwise, go to (Step4).

(Step4) If l is even, increase $\pi(\omega)$ for a certain $\omega \in \Omega_1$ by $\Delta\pi(\omega)$ which is determined by a method discussed later.
If l is odd, decrease positive $\pi(\omega)$ with the largest $R(\varphi^*_\pi,\omega)$ by $\Delta\pi(\omega)$ which is determined by a method stated later, too.
Increase l by one, $l = l + 1$, and go back to (Step2).

We propose a method to calculate $\Delta\pi(\omega)$ in (Step4) as follows. The largest expected reward for path k is obtained from $R(\varphi^*_k,k) = \max_{\varphi\in\Phi} R(\varphi,k)$ which is equivalent to the problem $\max_{\varphi\in\Phi} \tilde{R}(\varphi,\pi)$ in the case of $\pi(k) = 1$, $\pi(\omega) = 0$, $\forall\ \omega \neq k$. We denote the maximum reward for path k by $\overline{R}_k$. On the other hand, we estimate the minimum reward for path k by $\min_{\omega\neq k} R(\varphi^*_\omega,k)$ though it may not be a correct estimation. We denote the estimated minimum reward by $\underline{R}_k$. The maximum or minimum expected reward for path k corresponds to the case of $\pi(k) = 1$ or 0, respectively. We expect that the expected reward for path k changes linearly from $\underline{R}_k$ to $\overline{R}_k$ as $\pi(k)$ varies from 0 up to 1 and control the decrease/increase of the expected reward by changing $\pi(k)$. Our control rule is as follows. If we want to enlarge the expected reward of path k by γ, we increase $\pi(k)$ by

$$\Delta\pi(k) = \frac{\gamma}{\overline{R}_k - \underline{R}_k}\ , \tag{4.7}$$

and if we want to decrease it by γ, we decrease $\pi(k)$ not beyond current value $\pi(k)$ by $\Delta\pi(k)$.

In (Step3), we classify a set of paths Ω according to the value of their expected reward. Assume that path ω for which the expected reward $R(\varphi_\pi^*, \omega)$ is the largest and $\pi(\omega) > 0$ is classified into a subset $\Omega_{M'}$. We determine γ of Eq. (4.7) as follows.

(i) In the case of $M' = 1$, Lemma 3 is satisfied and the algorithm terminates.

(ii) In the case of $M' = 2$, the expected reward W_1 of $\omega \in \Omega_1$ ought to lift up to and the expected reward $W_{M'}$ of $\omega' \in \Omega_{M'}$ ought to put down to the intermediate value of W_1 and $W_{M'}$. That is,

$$\gamma = \frac{W_{M'} - W_1}{2} . \tag{4.8}$$

Using γ, we increase value $\pi(k)$ for a certain $k \in \Omega_1$ by $\Delta\pi(k)$ of Eq. (4.7) if l is even, and decrease $\pi(k)$ for a certain $k \in \Omega_{M'}$ by $\Delta\pi(k)$ if l is odd.

(iii) In the case of $M' > 2$, if l is even, we set

$$\gamma = W_2 - W_1 \tag{4.9}$$

for a $k \in \Omega_1$ in order to enlarge the expected reward W_1 up to W_2 and increase $\pi(k)$ by $\Delta\pi(k)$ determined by Eq. (4.7).

If l is odd, we set

$$\gamma = W_{M'} - W_{M'-1} \tag{4.10}$$

in order to shorten $W_{M'}$ up to $W_{M'-1}$ and decrease $\pi(k)$ by $\Delta\pi(k)$ determined by (4.7) for a $k \in \Omega_{M'}$.

Now, we finish explaining our computational algorithm for optimal searcher's and target's strategies. The validity of our algorithm can be proved as follows. Our original problem is a maximizing problem (P0) which is equivalent to $\max_{\varphi\in\Phi} \min_{\pi\in\Pi} \tilde{R}(\varphi, \pi)$. We prove that $\tilde{R}(\varphi_\pi^*, \pi)$ becomes smaller step by step varying π in our algorithm and converges to $\min_\pi \tilde{R}(\varphi_\pi^*, \pi) = \min_\pi \max_\varphi \tilde{R}(\varphi, \pi)$ at last. If so, we might say that our algorithm moves in such a way that it solves a Lagrangean dual problem of (P0) which is just a minimizing problem $\min_\pi \tilde{R}(\varphi_\pi^*, \pi)$.

Assume that we get a multiplier π and a one-sided optimal solution φ_π^* at a certain stage in our algorithm. From (4.1), we can normalize π whenever we like and by this reason, we assume $\sum_\omega \pi(\omega) = 1$. In (Step4), we change current π to a new π' by increasing $\pi(k)$ by $\Delta\pi(k)$. Its resulting difference ΔR of the total expected reward is estimated as follows.

$$\begin{aligned}\Delta R &= \sum_\omega \frac{\pi'(\omega)}{1+\Delta\pi(k)} R(\varphi_{\pi'}^*, \omega) - \tilde{R}(\varphi_\pi^*, \pi) \\ &= \frac{1}{1+\Delta\pi(k)} \left[\tilde{R}(\varphi_{\pi'}^*, \pi) + \Delta\pi(k) R(\varphi_{\pi'}^*, k) - (1+\Delta\pi(k))\tilde{R}(\varphi_\pi^*, \pi)\right] \\ &= \frac{1}{1+\Delta\pi(k)} \left[\Delta\pi(k)(R(\varphi_{\pi'}^*, k) - \tilde{R}(\varphi_\pi^*, \pi)) - (\tilde{R}(\varphi_\pi^*, \pi) - \tilde{R}(\varphi_{\pi'}^*, \pi))\right] .\end{aligned}$$

Because of $k \in \Omega_1$, we have $R(\varphi_\pi^*, k) < \tilde{R}(\varphi_\pi^*, \pi)$. If $\Delta\pi(k)$ is enough small, the ineqautlity $R(\varphi_{\pi'}^*, k) < \tilde{R}(\varphi_\pi^*, \pi)$ still holds for π' from the continuity of $R(\cdot, k)$. Taking account of $\tilde{R}(\varphi_\pi^*, \pi) - \tilde{R}(\varphi_{\pi'}^*, \pi) \geq 0$, it results in $\Delta R < 0$. When $R(\varphi_{\pi'}^*, k) < \tilde{R}(\varphi_\pi^*, \pi)$ is not satisfied for $\Delta\pi(k)$ estimated by (4.7), we may set $\Delta\pi(k)$ smaller, e.g. $\Delta\pi(k) = \Delta\pi(k)/2$.

Similarly, by the decrease of $\Delta\pi(k)$ for $k \in \Omega_{M'}$, the change of the expected reward is estimated as follows.

$$\begin{aligned}\Delta R &= \sum_\omega \frac{\pi'(\omega)}{1-\Delta\pi(k)} R(\varphi_{\pi'}^*, \omega) - \tilde{R}(\varphi_\pi^*, \pi) \\ &= \frac{1}{1-\Delta\pi(k)} \left[\Delta\pi(k)(\tilde{R}(\varphi_\pi^*, \pi) - R(\varphi_{\pi'}^*, k)) - (\tilde{R}(\varphi_\pi^*, \pi) - \tilde{R}(\varphi_{\pi'}^*, \pi))\right] .\end{aligned}$$

The fact $k \in \Omega_{M'}$ guarantees $\tilde{R}(\varphi_\pi^*, \pi) < R(\varphi_\pi^*, k)$. Then we have $\tilde{R}(\varphi_\pi^*, \pi) < R(\varphi_{\pi'}^*, k)$ for enough small $\Delta\pi(k)$. If the estimation (4.7) does not satisfy this relation, we may decrease $\Delta\pi(k)$, e.g. $\Delta\pi(k) = \Delta\pi(k)/2$, until the inequality holds. We have the inequality $\widetilde{R}(\varphi^*_\pi, \pi) - \widetilde{R}(\varphi^*_{\pi'}, \pi) \geq 0$ of course. In the result, we obtain $\Delta R < 0$.

From the above discussion, we can decrease $\widetilde{R}(\varphi^*_\pi, \pi)$ by adopting adequate $\Delta\pi(k)$ each time (Step2)~(Step4) are repeated. From the concavity and the finiteness of the expected reward, we conclude that our algorithm makes $\widetilde{R}(\varphi^*_\pi, \pi)$ converge to a minimum value $\min_\pi \widetilde{R}(\varphi^*_\pi, \pi)$. Thus we have proved that our algorithm has the so-called global convergence property[12]. However, we can not have clarified the rate of convergence of it because Iida and Hohzaki's method which is used in (Step2) to obtain a one-sided optimal distribution of search efforts after changing $\pi(\omega)$ is difficult to be analyzed in term of the rate of convergence.

5. Numerical Examples

Here, we investigate some characteristics of optimal strategies of the searcher and the target by the sensitivity analysis in some examples and the computational time of the proposed method.

Consider a search in discrete cells $\boldsymbol{K} = \{1, \cdots, 5\}$ and discrete time points $\boldsymbol{T} = \{1, \cdots, 10\}$. The target has four options of paths, that is, $|\Omega| = 4$. There are so many system parameters included in the game that we can not exhaust the sensitivity analysis concerning with all parameters. We keep the next parameters constant through all examples.

$c_0(i,t) = 1\ ,\quad u(t) = 5\ ,\quad i \in \boldsymbol{K} \quad t \in \boldsymbol{T}\ .$

(1) Effect of target paths

(Case 1) The search efficiency on cells and the value of the target are assumed to be constant.

$\alpha_i = 0.2\ ,\quad V(t) = 20\ ,\quad i \in \boldsymbol{K} \quad t \in \boldsymbol{T}\ .$

Routes of four target paths are illustrated in Fig.1 and shown in Table 1. The target on Path 3 or 4 stays always at cell 3 or 2, respectively. Paths 1 and 2 run across the cell space left to the right and right to the left.

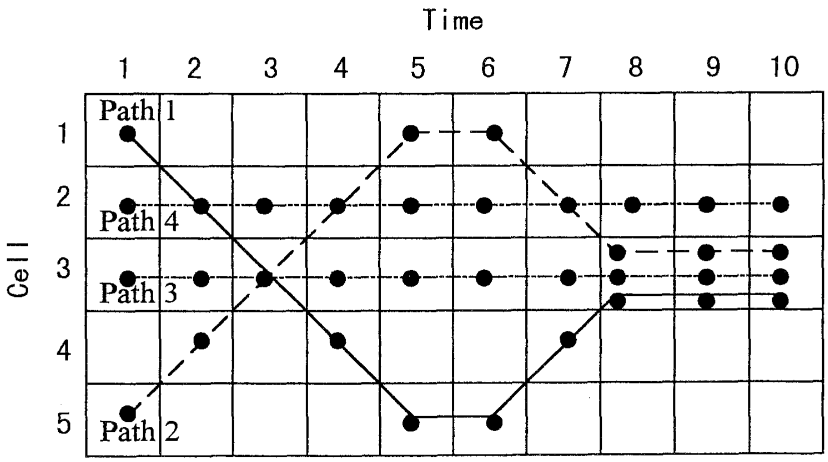

Figure 1. Target Paths

Table 1. Options of target paths

Paths \t	1	2	3	4	5	6	7	8	9	10
Path 1	1	2	3	4	5	5	4	3	3	3
Path 2	5	4	3	2	1	1	2	3	3	3
Path 3	3	3	3	3	3	3	3	3	3	3
Path 4	2	2	2	2	2	2	2	2	2	2

In this case, an optimal strategy of the target is $\pi^* = \{0.076, 0.028, 0.449, 0.447\}$ which means the target should mainly select Paths 3 and 4 with almost equal probability and should not use other paths. An optimal searcher's strategy, say an optimal distribution of search resources, is given in Table 2. Blank entry indicates no distribution. The searcher should allocate his search resources in cells 2 and 3 which Paths 3 and 4 run through, especially in cross points between Paths 3, 4 and the other paths. At many time points, only a part of available resources $u(t) = 5.0$ is used and the searcher exhausts them only at time points 9 and 10 when the probability of the target's existence focuses on cell 2 and 3. The value of the game is $\widetilde{R}(\varphi^*_{\pi^*}, \pi^*) = 8.03$.

Table 2. Optimal distribution of search resources

Cells \t	1	2	3	4	5	6	7	8	9	10
Cell 1										
Cell 2		2.244		2.035	0.704	1.088	2.868		2.342	2.500
Cell 3	0.826		2.975		1.226	1.055		3.647	2.658	2.500
Cell 4										
Cell 5										

(Case 2) We change options of the target's paths as shown in Fig.2 and Table 3 while other parameters remain unchanged as Case 1.

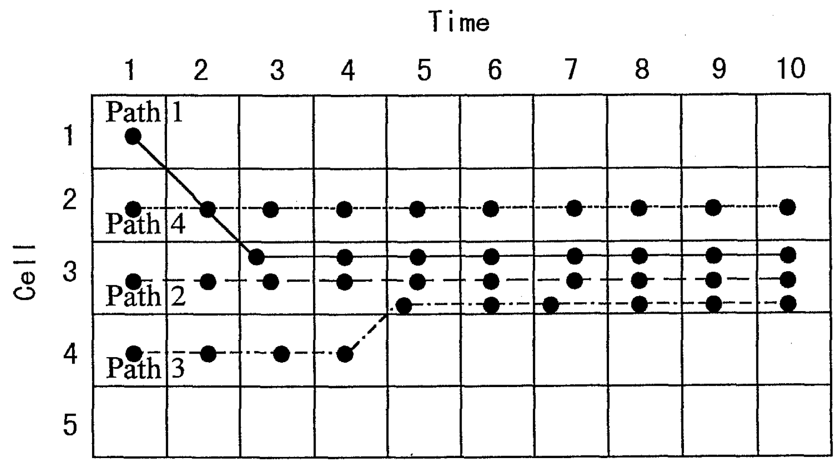

Figure 2. Target Paths

Table 3. Options of target paths

Paths \t	1	2	3	4	5	6	7	8	9	10
Path 1	1	2	3	3	3	3	3	3	3	3
Path 2	3	3	3	3	3	3	3	3	3	3
Path 3	4	4	4	4	3	3	3	3	3	3
Path 4	2	2	2	2	2	2	2	2	2	2

In this case, an optimal mixed strategy of selecting paths is $\pi^* = \{0.0, 0.285, 0.285, 0.430\}$. Path 1 comes across all other paths and it is convenient for the searcher. That is why the target should never select it. The probability of target's taking Paths 2 and 3 is approximately equal to the probability for Path 4, which makes it difficult for the searcher to anticipate where the target is each time. The searcher takes an optimal strategy presented by Table 4. The distribution of search resources in cell 3 begins after time 5 when Path 2 gets together with Path 3. The value of the game is $\tilde{R}(\varphi^*_{\pi^*}, \pi^*) = 7.74$.

Table 4. Optimal distribution of search resources

Cells \t	1	2	3	4	5	6	7	8	9	10
Cell 1										
Cell 2	0.645	0.686	0.734	0.794		0.442	1.792	2.156	2.500	2.500
Cell 3					2.270	2.426	1.792	2.156	2.500	2.500
Cell 4										
Cell 5										

(Case 3) A special set of the target paths is given in Case 3. We set Ω so that each target path stays alone at a cell which is illustrated in Table 5. Other parameters are the same as in Case 1.

Table 5. Options of target paths

Paths \t	1	2	3	4	5	6	7	8	9	10
Path 1	1	1	1	1	1	1	1	1	1	1
Path 2	2	2	2	2	2	2	2	2	2	2
Path 3	3	3	3	3	3	3	3	3	3	3
Path 4	4	4	4	4	4	4	4	4	4	4

In this case, optimal strategies of the target and the searcher are the balanced probabilities $\pi^* = \{0.25, 0.25, 0.25, 0.25\}$ and $\varphi(i,t) = 0$, respectively. The searcher has no way to estimate which of four cells the target is in and can not perform an efficient search at all. The value of the game is 0.

(2) Effect of the target value

(Case 4) Even if we have the same circumstance as Case 3, an optimal solution changes depending on other parameters. If we increase the target value up to $V(t) = 50$, an optimal strategy of the searcher becomes as follows,

$\varphi(i,1) = 1.21\ ,\ \ i \in \boldsymbol{K}\ ,\ \ \varphi(i,t) = 1.25\ ,\ \ i \in \boldsymbol{K}\ \ t \in \{2, \cdots, 10\}\ ,$

while that of the target remains unchanged. The detection probability of the target in this case may be low even though the search efforts are distributed but the reward is expected

to be larger because of the high target value. This fact encourages the searcher to begin the search. The value of the game is $\widetilde{R}(\varphi^*_{\pi^*}, \pi^*) = 25.15$.
(3) Effect of the search efficiency of cells
(Case 5) We change α_i of Case 1 as follows.

$\alpha_1 = 0.1$, $\alpha_2 = 0.2$, $\alpha_3 = 0.3$, $\alpha_4 = 0.4$, $\alpha_5 = 0.5$,

which means that the search efficiency of cells per unit search resource becomes larger in the order of cell numbers $1, 2, \cdots, 5$.

In this case, an optimal strategy of the target is $\pi^* = \{0.075, 0.021, 0.361, 0.543\}$ and that of the searcher is given in Table 6.

Table 6. Optimal distribution of search resources

Cells \ t	1	2	3	4	5	6	7	8	9	10
Cell 1										
Cell 2	0.371	1.946		1.838	1.351	1.348	2.781	0.421	3.000	3.000
Cell 3	0.476		2.596		0.585	0.842		3.052	2.000	2.000
Cell 4										
Cell 5										

The target tends to select Path 4 than Path 3 comparing with Case 1 because Path 4 stays at Cell 2 with less detectability for the target. According to the target's strategy, the searcher focuses a little more allocation of search resources on Cell 2 than on Cell 3. The little unbalance of the target's path selection and $\alpha_i \geq 0.2$ in Cell 2, 3 and 4 make the value of the game lift up a little to 9.93 from 8.03 of Case 1.
(4) Computational time

It is thought that there are three parameters of having much effect on the computational time of the algorithm proposed in Section 4: the number of cells $K = |\boldsymbol{K}|$, the number of target paths $|\Omega|$ and the number of time points $T = |\boldsymbol{T}|$. Noting that the proposed algorithm repeats the adjustment of a balance between $|\Omega|$ expected rewards for all target paths, it is anticipated that the parameter $|\Omega|$ directly influences the number of the repetition of the algorithm. As a sub-procedure, the algorithm uses Iida and Hohzaki's method which is also an iterative method of revising a feasible solution better at each time point and making a sequence of feasible solutions converge to an optimal one. That is why the number of time points T plays an important role in the computational time of their method. A large K makes many variables $\{\varphi(i,t),\ i \in \boldsymbol{K}\ \ t \in \boldsymbol{T}\}$ and might increase the computational time of solving the game to some extent. To examine the extent of these effects on the computational time, we take statistics of CPU-times by the following example.

The whole target paths are randomly generated. That is, a cell is selected randomly from K cells T times and then a path is generated. This procedure is repeated $|\Omega|$ times and a set of target paths is made. The detectability parameter α_i of cell i is selected randomly from interval $[0.1, 1]$. The value parameter of the detection and the cost parameter are kept constant, $V(t) = 20$ and $c_0(i,t) = 1$, respectively. In the search game, the total amount $u(t) = 5$ of search efforts are supposed to be available to the searcher. Thus, a game problem has been generated which is solved by the proposed method to measure CPU-time. IBM personal computer Aptiva B97(Pentium 200MHz) and the programming language BASIC are used on the computation.

For a combination of K, $|\Omega|$, T, a hundred problems are generated and a mean value of CPU-times required to solve the problems is calculated. For every combination of $K =$

$5, 10$, $|\Omega| = 5, 10$ and $T = 5, 10$, the average CPU-times are given in Table 7.

Table 7. CPU-time

K	$\|\Omega\|$	T	CPU-time(sec)
5	5	5	5.28
10	5	5	8.34
5	10	5	24.00
5	5	10	21.09
10	5	10	29.64
10	10	5	27.83
5	10	10	81.49
10	10	10	112.20

As we expected, $|\Omega|$ has much effect on the computational time. The CPU-time for the case of $(K, |\Omega|, T) = (5, 10, 5)$ is about 4 times as large as the case of $(K, |\Omega|, T) = (5, 5, 5)$. The same tendency is seen in the cases of $(K, |\Omega|, T) = (5, 5, 10)$ and $(K, |\Omega|, T) = (5, 10, 10)$ or the cases of $(K, |\Omega|, T) = (10, 5, 10)$ and $(K, |\Omega|, T) = (10, 10, 10)$. The varying of T verifies that it has approximately the same effect as $|\Omega|$ on the CPU-time. Comparing with these two elements, the effect of K on the CPU-time is very small.

6. Conclusions

This paper investigates a search game with two players, say, a searcher and a target. At the beginning of the search, the target selects his path from some options and the searcher determines his distribution of his available search resources into a search space which consists of discrete cells and discrete time points. In Introduction, we survey previous researches concerning with search problems for a moving target. The assumptions that the target selects his path and the searcher allocates the search resources into the search space are thought to be basic assumptions used in many studies.

In our model, the searcher gains a value on detection of the target while he expends the search cost depending on the allocation of the search resource. The criterion of the problem is the expected reward which is defined as the expected value minus the expected search cost. The criterion can contains the criterion of the detection probability of the target, which many studies deal with as most important criterion, and then, in that sense, it is a generalized criterion which makes us view the problem from the point of the cost-performance. The searcher acts as a maximizer of the expected reward and the target as a minimizer.

There are not so many papers dealing with the search game where the strategy of the searcher is represented by the distribution plan of search efforts. We formulate the search problem as a two-person zero-sum game with a pay-off function of the expected reward. By the theorems of the convex game, we know that an optimal solution is found in the region of the mixed strategy of the target and the pure strategy of the searcher. We reduce the game to a concave maximization problem and propose a computational algorithm to give an optimal solution. The algorithm proceeds in such a way that one-sided problems generated from the original game are repeatedly solved and a sequence of feasible solutions generated by the algorithm converges to an optimal solution, which is proved. By some examples, we examine the effect of some system parameters on an optimal solution, and analyze some characteristics of the solution and the computational time of the proposed algorithm.

The pay-off of the game has a characteristic form that it is linear for one variable set and concave for another variable set, by which we propose a computational algorithm. We describe this idea in the context of solving a search game. However, this method could be applied to many problems in not only the search theory but also other fields, we think.

The problem is formulated on discrete search space and time. If we want to theoretically discuss the continuous version of the problem and obtain an optimal solution which is represented by a function on the continuous search space and time, we would need some different tools from ours, e.g. the calculus of variations. However, if the approximate estimation of the optimal function is acceptable, our method could be applied by dividing the search space and time into many discrete regions. In this case, we must consider a trade-off between the precision of the solution and the computational time on deciding how many regions we take.

References

[1] S.S. Brown: Optimal search for a moving target in discrete time and space. *Operations Research*, **28**(1980) 1275–1289.

[2] J.M. Danskin: A helicopter versus submarine search game. *Operations Research*, **16**(1968) 509–517.

[3] J.N. Eagle: The optimal search for a moving target when the search path is constrainted. *Operations Research*, **32**(1984) 1107-1115.

[4] O. Hellman: On the optimal search for a randomly moving target. *SIAM J. Appl. Math.*, **22**(1972) 545–552.

[5] R. Hohzaki and K. Iida: A search game with a moving target when a search path is given (in Japanese). *Abstracts of the 1996 Spring National Conference of ORSJ*(1996) 152–153.

[6] R. Hohzaki and K. Iida: Optimal strategy of route and look for the path constrainted search problem with reward criterion. *European J. of Operational Research*, **100**(1997) 236–249.

[7] K. Iida: An optimal distribution of searching effort for a moving target (in Japanese). Keisei Kagaku, **16**(1972) 204–215.

[8] K. Iida and R. Hozaki: The optimal search plan for a moving target minimizing the expected risk. *Journal of the Operations Research Society of Japan*, **31**(1988) 294–320.

[9] K. Iida, R. Hohzaki and S. Furui: A search game for a mobile target with the conditionally deterministic motion defined by paths. *Journal of the Operations Research Society of Japan*, **39**(1996) 501–511.

[10] K. Kikuta: A hide and seek game with traveling cost. *Journal of the Operations Research Society of Japan*, **33**(1990) 168–187.

[11] K. Kikuta: A search game with traveling cost. *Journal of the Operations Research Society of Japan*, **34**(1991) 365–382.

[12] H. Konno and H. Yamasita: *Nonlinear Programming* (in Japanese) (Nikka-giren, Tokyo, 1984).

[13] G. Owen: *Game Theory* (Academic Press, New York, 1995).

[14] S.M. Pollock: A simple model of search for a moving target. *Operations Research*, **18**(1970) 883–903.

[15] L. Saretsaro: On the optimal search for a target whose motion is a Markov process. *Journal of Applied Probability*, **10**(1973) 847–856.

[16] L.D. Stone: *Theory of Optimal Search* (Academic Press, New York, 1975).

[17] W.R. Stromquist and L.D. Stone: Constrained optimization of functionals with search

theory applications. *Mathematics of Operations Research*, **16**(1981) 518–529.
[18] A.R. Washburn: Search for a moving target: The FAB algorithm. *Operations Research*, **31**(1983) 739–750.
[19] A.R. Washburn: Search-evasion game in a fixed region. *Operations Research*, **28**(1980) 1290–1298.

Ryusuke Hohzaki
Department of Applied Physics
National Defense Academy
1-10-20 Hashirimizu, Yokosuka,
239-8686, Japan
E-mail: hozaki@cc.nda.ac.jp