日本オペレーションズ・リサーチ学会
2005年春季研究発表会

目標側のエネルギー補充戦略を考慮した捜索資源配分ゲーム

02005680　防衛大学校　*池田圭子　IKEDA Keiko
01504810　防衛大学校　宝崎隆祐　HOHZAKI Ryusuke
01110110　防衛大学校　小宮　享　KOMIYA Toru

1　はじめに

捜索者と目標の２人のプレーヤーが参加する捜索ゲームの中に，捜索者は目標を探知すべく手持ちの捜索資源を投入し，目標は捜索者から逃避すべく移動戦略をとる捜索割当ゲームがある．Hohzaki ら [1] は，従来の研究に対し目標の移動エネルギーを考慮したモデルを提案した．本研究では，目標にエネルギー補充の意思決定が可能であるとした場合の捜索割当ゲームを考える．

2　モデルの前提

捜索者と目標が参加する２人ゼロ和ゲームを考える．

[Ⅰ] 地理空間は n 個のセルから成る離散空間 $\mathbb{K}=\{1,\cdots,n\}$ であり，時間空間は時点 0 から m までの離散時点の集合 $\mathbb{T}=\{0,\cdots,m\}$ とする．

[Ⅱ] 目標はこの空間を移動することにより捜索者からの逃避を図る．目標は初期時点 $t=0$ にセル群 $\mathbb{K}_0\subseteq\mathbb{K}$ のいずれかのセルから出発し，時点 t でのセル i からは次の時点でセル群 $N(t,i)$ のみへ移動できる．セル i から j への移動にはエネルギー $\mu(i,j)$ を消費する．初期の所有エネルギーは e_0 である．目標は各時点において潜伏態勢（0）と補充態勢（1）のいずれかを選択できるとし，態勢の集合を $\mathbb{S}=\{0,1\}$ で表す．ただし，$t=0$ のときの態勢は 0 とする．補充態勢の選択により次の時点でエネルギーは ε だけ増加するが，潜伏態勢では補充はされず増加しない．

[Ⅲ] 捜索者は捜索空間上へ捜索資源を投入することにより目標を探知しようと図る．捜索は時点 τ に開始され，その捜索可能な時間帯を $\hat{\mathbb{T}}:=\{\tau,\cdots,m\}\subseteq\mathbb{T}$ で表す．また，捜索者の使用可能な資源総量は各時点 $t\in\hat{\mathbb{T}}$ において $\Phi(t)$ である．

[Ⅳ] 捜索者の捜索資源配分戦略と目標の移動戦略によるゲームの支払は，目標の移動経路上に投入された捜索資源の重み付き総量の指数関数で表されるとするが，セル i での単位捜索資源量に対する重み（探知しやすさ）は，目標の態勢が 0 ならば α_i，態勢が 1 ならば β_i である．

3　定式化

問題は，捜索者をマキシマイザー，目標をミニマイザーとする１段階の２人ゼロ和ゲームである．時点 t，セル i の捜索空間上の点 (t,i) に投入する捜索資源配分 $\varphi(t,i)$ の集合が捜索者の純粋戦略であり，これを $\varphi=\{\varphi(t,i),t\in\hat{\mathbb{T}},i\in\mathbb{K}\}$ で表すと，前提 [Ⅲ] から，その実行可能領域は $\Psi=\{\varphi(t,i)|\varphi(t,i)\geq 0,\sum_i\varphi(t,i)\leq\Phi(t)\}$ である．

目標の時点 t における移動セルを $\omega(t)$，態勢を $\sigma(t)$，残存エネルギーを $e(t)$ とすると，前提 [Ⅱ] から，目標の経路制約は $\omega(t+1)\in N(t,\omega(t))$，エネルギー制約は $e(0)=e_0$，$e(t)=e(t-1)-\mu(\omega(t-1),\omega(t))+\varepsilon\sigma(t-1)\geq 0$ $(t=1,\cdots,m)$ であり，また $\mu(\omega(t),\omega(t+1))\leq e(t)$ $(t=0,\cdots,m-1)$ である．これを満たす経路・態勢を実行可能パスと呼び，その実行可能領域を $\hat{\mathbb{P}}$ とする．この時，前提 [Ⅳ] による捜索資源配分 φ と実行可能パス $\rho=(\omega,\sigma)$ の支払関数は $R(\varphi,\rho)=1-\exp(-\sum_{t\in\hat{\mathbb{T}}}\varphi(t,\omega(t))\{\alpha_{\omega(t)}(1-\sigma(t))+\beta_{\omega(t)}\sigma(t)\})$ と表される．この支払関数が φ について凹であることに注意すれば，凸ゲームに関するこれまでの研究から均衡解は純粋戦略で与えられることが分かる．一方，パス ρ を選択する確率を $\pi(\rho)$ とし，パスの選択に関する混合戦略を $\pi=\{\pi(\rho),\rho\in\hat{\mathbb{P}}\}$ で定義すると，捜索者の純粋戦略 φ，目標の混合戦略 π による期待支払は $R(\varphi,\pi)=\sum_\rho\pi(\rho)R(\varphi,\rho)$ となる．以上から，ゲームの値は次式で与えられる．ただし，Π は，$\pi(\rho)\geq 0$，$\rho\in\hat{\mathbb{P}}$，$\sum_{\rho\in\hat{\mathbb{P}}}\pi(\rho)=1$ の条件を満たすパス選択確率 π の実行可能領域である．

$$(P^p)\quad \max_\varphi\min_\pi\ R(\varphi,\pi)\quad \text{s.t. } \varphi\in\Psi,\pi\in\Pi.$$

この問題の最適解は目的関数を $\sum_t\varphi(t,\omega(t))\{\alpha_{\omega(t)}(1-\sigma(t))+\beta_{\omega(t)}\sigma(t)\}$ と変えても同じであるため，以下ではこの線形式を支払関数として議論する．

4　解法

目標の実行可能パス全体の数は一般にセル数×態勢数に対する時間のベキ乗で増加し，大きなサイズの問題に対しては取り扱いが困難であるため，ここでは目標の戦略として各時点における存在確率を用いる．また，エネルギー消費関数 $\mu(\cdot)$，初期エネルギー e_0，補充エネルギー ε は非負整数値とする．取り得るエネルギーの最

大値$e_{MAX} = e_0 + (m-1)\varepsilon$により，エネルギー全体は$\boldsymbol{E} = \{0, \cdots, e_{MAX}\}$で表せる．目標が時点$t$にセル$i$, 態勢$s$, 所有エネルギー$e$の状態$(t,i,s,e)$にある目標の確率を$q(t,i,s,e)$で表す．また，目標が状態$(t,i,s,e)$から，次の時点$t+1$でセル$j$, 態勢$r$に推移する確率を$v(t,i,j,s,r,e)$とする．目標が状態$(t,i,s,e)$から次の時点$t+1$で移動可能なセル群は$C(t,i,e) = \{j \in N(t,i) | \mu(i,j) \le e\}$, 状態$(t,i,s,e)$へ移動可能な前の時点$t-1$でのセル群は，時点$t-1$での態勢$r$を用いて$C^*(t,i,r,e) = \{j \in \boldsymbol{K} | i \in C(t-1, j, e+\mu(j,i) - r\varepsilon)\}$で表現できる．

ここで，変数$h(t)$を時点t以降の最適捜索資源配分による期待支払の最大値と定義する．時点tでの期待支払および$t+1$以降の最大期待支払を考えれば，$h(t)$は以下の動的計画法により定式化できる．

$$\begin{aligned} h(t) &= \max_{\varphi}\{\textstyle\sum_i \alpha_i \varphi(t,i) \sum_e q(t,i,0,e) \\ &\qquad + \textstyle\sum_i \beta_i \varphi(t,i) \sum_e q(t,i,1,e) + h(t+1)\} \\ &= \Phi(t) \max_i \{\alpha_i \textstyle\sum_e q(t,i,0,e) + \beta_i \sum_e q(t,i,1,e)\} \\ &\quad + h(t+1). \end{aligned} \tag{1}$$

ここで，$\alpha_i \sum_e q(t,i,0,e) + \beta_i \sum_e q(t,i,1,e)$を時点$t$, セル$i$での被探知効力と呼ぶ．時点$\tau$以降の捜索による最大期待支払$h(\tau)$を最小にするように目標は行動するから，(1)式及び目標存在確率の保存則から目標の存在と移動に関する最適解q^*, v^*を与える次の線形計画問題が作成できる．ただし，$h(m+1) = 0$とする．

(P^e)　$\min\ h(\tau)$

s.t.

$$\begin{aligned} & h(t) \ge \Phi(t)\{\alpha_i \textstyle\sum_e q(t,i,0,e) + \beta_i \sum_e q(t,i,1,e)\} \\ &\qquad + h(t+1),\ t = \tau, \cdots, m,\ i \in \boldsymbol{K}, \\ & q(t,i,s,e) = \textstyle\sum_r \sum_{j \in C(t,i,e)} v(t,i,j,s,r,e), \\ &\quad t = 0, \cdots, m-1,\ i \in \boldsymbol{K},\ s \in \boldsymbol{S},\ e \in \boldsymbol{E}, \\ & q(t,i,s,e) \\ & = \textstyle\sum_r \sum_{j \in C^*(t,i,r,e)} v(t-1,j,i,r,s,e+\mu(j,i) - r\varepsilon), \\ &\quad t = 1, \cdots, m,\ i \in \boldsymbol{K},\ s \in \boldsymbol{S},\ e \in \boldsymbol{E}, \\ & \textstyle\sum_{i \in \boldsymbol{K}_0} q(0,i,0,e_0) = 1, \\ & \textstyle\sum_s \sum_i \sum_e q(t,i,s,e) = 1,\ t \in \boldsymbol{T}, \\ & v(t,i,j,s,r,e) \ge 0, \\ &\quad t = 0, \cdots, m-1,\ i,j \in \boldsymbol{K},\ s,r \in \boldsymbol{S},\ e \in \boldsymbol{E}. \end{aligned}$$

また，捜索者の最適戦略φ^*はこの問題の双対問題から得られる．

5　数値例

パラメータ設定$n = 10$, $m = 10$, $\boldsymbol{K}_0 = \{1\}$, $N(t,i) = \boldsymbol{K}$, $e_0 = 7$, $\varepsilon = 2$, $\mu(i,j) = (i-j)^2$, $\tau = 2$, $\alpha_i = 0.2$, $\beta_i = 0.8$, $\Phi(t) = 1$の例題に対し，時点$t = 5$までの最適解を抜粋したのが図1である．縦軸にセルを，横軸に時点をとり，目標の被探知効力及び最適捜索資源配分の大きさを四角により相対表示している．また，態勢0,1の最適な選択確率を影付及び斜線付四角により示しているが，各時点におけるその和は1である．

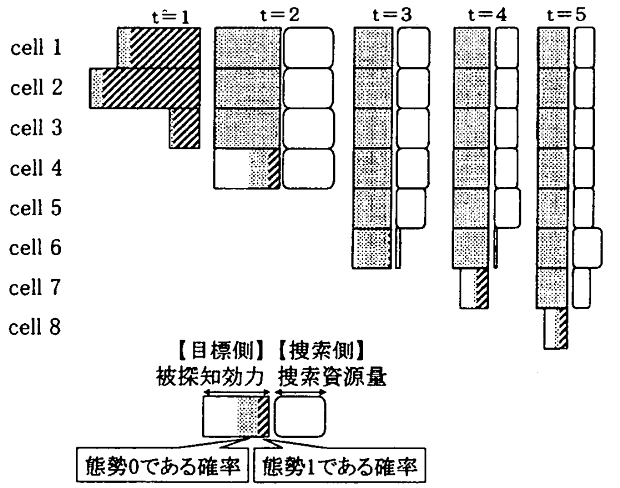

図1　最適解における被探知効力と捜索資源配分

プレーヤーの最適戦略に関し次の性質が見られる．目標が態勢1をとる確率は，出発したセルから遠いセルで高い．これは残存エネルギー量と存在セル領域の拡張に対する効率性が影響を与えているものと考えられる．捜索者の各時点での捜索資源投入量はセルによって多少のばらつきがある．これは，目標の存在確率と逃避セルの拡大防止を考慮した結果であると考えられる．

6　まとめ

本研究では目標にエネルギー補充戦略を付加することにより，従来の移動戦略のみを考慮したモデルより柔軟な逃避戦略が可能となった．補充態勢をとる意思決定により，間接的に移動戦略に正の影響を及ぼし，直接的にゲームの支払（被探知率）に負の影響を受ける効果を得た．このように状態を変化させ，2つの要素間でのトレードオフを可能にする手法は，様々な問題に応用可能であると思われる．

参考文献

[1] R. Hohzaki, K. Iida, T. Komiya, "*Discrete Search Allocation Game with Energy Constraints*", Journal of the Operations Research Society of Japan, 45(1), pp. 93-107, 2002.