数理解析研究所講究録
第2078巻 2018年 194-200

目標及び捜索者の双方が個人情報をもつ捜索ゲーム

防衛大学校・理工学研究科　宮田　鉄矢
Tetsuya Miyata
Graduate School of Science and Engineering,
National Defence Academy
防衛大学校・情報工学科　宝崎　隆祐
Ryusuke Hohzaki
Department of Computer Science,
National Defense Academy

1 はじめに

警戒監視，対潜捜索，海難救助のような捜索活動を効率的に行うためには，目標を見つけようとする捜索者の意図だけでなく，捜索者に見つからないように，または見つけてもらえるように行動する目標の意図も考慮して捜索計画を立案することが重要である．そのため，複数のプレイヤーの戦略を分析するゲーム理論を用いた捜索ゲームの研究が盛んである．

本研究では，目標を見つけようとする捜索者に対し見つからないように行動する目標のように，参加者が敵対的に行動する捜索ゲームを取り扱う．このような捜索では，捜索資源の効果的な配分のため，捜索者は目標の時々刻々の存在分布をできるだけ正確に見積もることが重要となる．そのためには目標の初期位置と目標の機動性を左右する初期移動エネルギーを正しく推定することが重要であるが，実際にはこれらの情報は目標の秘匿情報であり，捜索者が正確に知ることはできない．一方，逃避しようとする目標にとって捜索者のもつ捜索資源の探知能力を正しく推定することが肝要であるが，実際にはこの情報は捜索者の秘匿情報であり，目標が知ることは困難である．なお，ゲーム理論ではこれらの秘匿情報を個人情報といい，本稿でも以後個人情報という用語を用いる．

従来研究では，目標のみが個人情報をもつ場合 [1]，あるいは捜索者のみが個人情報をもつ場合のモデル [2] しか考えられてこなかったが，上述したように参加者双方が不確実な相手情報しか獲得できない状況が現実的でかつ一般的であり，本研究ではそのような捜索状況を取り扱う．すなわち，目標の初期位置，初期エネルギー及び捜索者の捜索資源の探知能力に関する情報を本人のみが知る個人情報とし，相手プレイヤーはこれを確率分布でしか予測できない情報不完備捜索ゲームを議論し，これら3つの個人情報がゲームに及ぼす影響を定量的に分析する．

2 モデルの前提と定式化

捜索者と目標のそれぞれが個人情報をもつ，以下の7つの前提からなる捜索ゲームを考える．

A1. 捜索空間は離散セル空間 $\boldsymbol{K}=\{1,\dots,K\}$ と，離散時間空間 $\boldsymbol{T}=\{1,\dots,T\}$ から成るものとする．

A2. 目標は，初期位置 $k_0\in\boldsymbol{K}_0\subseteq\boldsymbol{K}$ と初期エネルギー $e_0\in\boldsymbol{E}_0$ を初期状態としてもつ可能性があり，この初期状態の目標を (k_0,e_0) タイプ目標という．捜索者は，この初期状態 (k_0,e_0) を互いに独立な確率分布 $\{f(k_0),k_0\in\boldsymbol{K}_0\},\{g(e_0),e_0\in\boldsymbol{E}_0\}$ で推定する．この分布は両者にとって既知である．

A3. 目標は捜索空間内を時間と共に移動する．時刻 t にセル i から次時刻 $t+1$ で移動可能なセル群を $\boldsymbol{N}(i,t)$，そのとき消費するエネルギーを $\mu(i,j)$ とし，初期エネルギーが尽きた時は他のセルに移動できない．以上の条件を満たす，(k_0,e_0) タイプ目標が選択可能なパスの集合を $\boldsymbol{\Omega}_{k_0e_0}$ とし，このタイプの目標はその一つを選択して移動する．パス $\omega\in\boldsymbol{\Omega}_{k_0e_0}$ をとる目標の時点 t での位置を $\omega(t)\in\boldsymbol{K}$ で表す．

A4. 捜索者は捜索資源のタイプの集合 $\boldsymbol{S}$ のうち，いずれか1つのタイプ $s\in\boldsymbol{S}$ を選択して使用する．これを s タイプの捜索者と呼ぶ．目標は，資源のタイプを確率分布 $\{h(s),s\in\boldsymbol{S}\}$ で推定する．この分布は両者にとって既知である．

A5. 捜索者は，時刻 τ 以降の時間帯 $\widehat{\boldsymbol{T}}=\{\tau,\tau+1,\ldots,T\}$ に，捜索資源を捜索空間内に投入して目標を捜索する．s タイプ捜索者の時刻 t での使用可能資源量は $\Phi_s(t)$ であり，任意に分割可能である．

A6. セル i に存在する目標に対し，s タイプの捜索資源を x だけ投入したとき，捜索者は確率

$$1-\exp(-\alpha_i^s x) \tag{1}$$

で目標を探知する．ここで，α_i^s は s タイプの単位捜索資源をセル i に投入したときの探知効率を表す非負のパラメータである．

A7. 目標を探知した時点，あるいは最終時点 T でゲームは終了する．目標を探知すると捜索者は利得 1 を得て，目標は同量を失う．

前提 A7 から，問題は目標探知確率を支払関数とする 2 人ゼロ和ゲームである．

ここで，プレイヤーの戦略を明確にしておく．(k_0,e_0) タイプ目標は混合戦略をとるものとし，パス $\omega\in\boldsymbol{\Omega}_{k_0e_0}$ を選択する確率を $\pi_{k_0e_0}(\omega)$，その確率分布を $\boldsymbol{\pi}_{k_0e_0}=\{\pi_{k_0e_0}(\omega)\geq 0,\omega\in\boldsymbol{\Omega}_{k_0e_0}\mid\sum_{\omega\in\boldsymbol{\Omega}_{k_0e_0}}\pi_{k_0e_0}(\omega)=1\}$ で表す．一方，s タイプの捜索者の捜索資源配分計画を $\boldsymbol{\varphi}_s=\{\varphi_s(i,t)\geq 0,i\in\boldsymbol{K},t\in\widehat{\boldsymbol{T}}\mid\sum_{i\in\boldsymbol{K}}\varphi_s(i,t)=\Phi_s(t)\}$ で表す．$\varphi_s(i,t)$ は，時刻 t にセル i へ投入する s タイプの捜索資源配分量である．

以上の戦略表現を用いて，支払関数である目標探知確率を求めよう．目標が一つのパス ω をとり，s タイプ捜索者が計画 $\boldsymbol{\varphi}_s$ をとれば，式 (1) から目標探知確率は次式となる．

$$R_{s,k_0e_0}(\boldsymbol{\varphi}_s,\omega)=1-\exp\left(-\sum_{t\in\widehat{\boldsymbol{T}}}\alpha_{\omega(t)}^s\varphi_s(\omega(t),t)\right)$$

以降では定式化を容易にするため，探知確率ではなく，次式で表される非探知確率を支払関数として議論を進めていく．

$$Q_{s,k_0e_0}(\boldsymbol{\varphi}_s,\omega)=\exp\left(-\sum_{t\in\widehat{\boldsymbol{T}}}\alpha_{\omega(t)}^s\varphi_s(\omega(t),t)\right)$$

この式から，(k_0,e_0) タイプ目標が混合戦略 $\boldsymbol{\pi}_{k_0e_0}$ をとるときの期待支払は，

$$Q_{s,k_0e_0}(\boldsymbol{\varphi}_s,\boldsymbol{\pi}_{k_0e_0})=\sum_{\omega\in\boldsymbol{\Omega}_{k_0e_0}}\pi_{k_0e_0}(\omega)Q_{s,k_0e_0}(\boldsymbol{\varphi}_s,\omega) \tag{2}$$

となり，s タイプの捜索者の期待支払は，目標タイプの分布を考慮した

$$Q_s(\boldsymbol{\varphi}_s,\boldsymbol{\pi})=\sum_{k_0\in\boldsymbol{K}_0}\sum_{e_0\in\boldsymbol{E}_0}f(k_0)g(e_0)Q_{s,k_0e_0}(\boldsymbol{\varphi}_s,\boldsymbol{\pi}_{k_0e_0}) \tag{3}$$

となる．一方，(k_0,e_0) タイプ目標の期待支払は，式 (2) に捜索者のタイプの分布を考慮し，

$$Q_{k_0e_0}(\boldsymbol{\varphi},\boldsymbol{\pi}_{k_0e_0})=\sum_{s\in\boldsymbol{S}}h(s)Q_{s,k_0e_0}(\boldsymbol{\varphi}_s,\boldsymbol{\pi}_{k_0e_0}) \tag{4}$$

となる．ここで，各タイプ s の捜索者は式 (3) を最小化しようとし，各タイプ (k_0,e_0) の目標は式 (4) を最大化しようとするが，これは次式を共通の期待支払として最小化，または最大化することと同値である．

$$Q(\boldsymbol{\varphi},\boldsymbol{\pi})=\sum_{k_0\in\boldsymbol{K}_0}\sum_{e_0\in\boldsymbol{E}_0}f(k_0)g(e_0)\sum_{s\in\boldsymbol{S}}h(s)Q_{s,k_0e_0}(\boldsymbol{\varphi}_s,\boldsymbol{\pi}_{k_0e_0})$$

捜索者は $Q(\boldsymbol{\varphi},\boldsymbol{\pi})$ を最小化するように，目標はそれを最大化するように行動する．

3 最適戦略の導出

捜索者の最適戦略は $Q(\boldsymbol{\varphi},\boldsymbol{\pi})$ のミニマックス最適化により得られ，最終的に次の凸計画問題に定式化される．これを解けば捜索者の最適戦略 $\boldsymbol{\varphi}^*$ を求めることができる．

$$(\mathrm{P_S})\quad \min_{\boldsymbol{\varphi},\boldsymbol{\xi}} \sum_{k_0\in\boldsymbol{K}_0}\sum_{e_0\in\boldsymbol{E}_0} f(k_0)g(e_0)\xi_{k_0e_0}$$

$$\text{s.t.}\quad \xi_{k_0e_0} \geq \sum_{s\in\boldsymbol{S}} h(s)\exp\left(-\sum_{t\in\widehat{\boldsymbol{T}}}\alpha^s_{\omega(t)}\varphi_s(\omega(t),t)\right), \omega\in\boldsymbol{\Omega}_{k_0e_0}, k_0\in\boldsymbol{K}_0, e_0\in\boldsymbol{E}_0,$$

$$\sum_{i\in\boldsymbol{K}}\varphi_s(i,t)=\Phi_s(t), t\in\widehat{\boldsymbol{T}}, s\in\boldsymbol{S},$$

$$\varphi_s(i,t)\geq 0, i\in\boldsymbol{K}, t\in\widehat{\boldsymbol{T}}, s\in\boldsymbol{S}.$$

目標の最適戦略は，上で得られた $\boldsymbol{\varphi}^*$ を用いて求める．まず，$\boldsymbol{\varphi}^*$ に対して最適反応となる目標の戦略 $\boldsymbol{\pi}$ は，問題 $\max_{\pi} Q(\boldsymbol{\varphi}^*,\boldsymbol{\pi})$ により得られる．また，$\boldsymbol{\varphi}^*$ が目標の戦略 $\boldsymbol{\pi}$ に対し最適反応であるためには，$\boldsymbol{\varphi}^*_s$ は凸計画問題 $\min_{\boldsymbol{\varphi}_s} Q_s(\boldsymbol{\varphi}_s,\boldsymbol{\pi})$ の最適解となり，KKT 条件を満たしているはずである．以上の 2 つの条件を組み合わせることにより，目標の最適戦略 $\boldsymbol{\pi}^*$ を導出する線形計画問題が以下のように得られる．

$$(\mathrm{P_T})\quad \max_{\boldsymbol{\pi},\boldsymbol{\lambda}} \sum_{k_0\in\boldsymbol{K}_0}\sum_{e_0\in\boldsymbol{E}_0} f(k_0)g(e_0)\sum_{\omega\in\boldsymbol{\Omega}_{k_0e_0}}\pi_{k_0e_0}(\omega)\sum_{s\in\boldsymbol{S}}h(s)\exp\left(-\sum_{t\in\widehat{\boldsymbol{T}}}\alpha^s_{\omega(t)}\varphi^*_s(\omega(t),t)\right)$$

s.t. $\varphi^*_s(i,t)>0$ なる $(i,t,s)\in\boldsymbol{K}\times\widehat{\boldsymbol{T}}\times\boldsymbol{S}$ に対して，

$$\lambda_s(t)=\alpha^s_i\sum_{k_0\in\boldsymbol{K}_0}\sum_{e_0\in\boldsymbol{E}_0}f(k_0)g(e_0)\sum_{\omega\in\boldsymbol{\Omega}^{k_0e_0}_{it}}\pi_{k_0e_0}(\omega)\exp\left(-\sum_{t'\in\widehat{\boldsymbol{T}}}\alpha^s_{\omega(t')}\varphi^*_s(\omega(t'),t')\right),$$

$\varphi^*_s(i,t)=0$ なる $(i,t,s)\in\boldsymbol{K}\times\widehat{\boldsymbol{T}}\times\boldsymbol{S}$ に対して，

$$\lambda_s(t)\geq\alpha^s_i\sum_{k_0\in\boldsymbol{K}_0}\sum_{e_0\in\boldsymbol{E}_0}f(k_0)g(e_0)\sum_{\omega\in\boldsymbol{\Omega}^{k_0e_0}_{it}}\pi_{k_0e_0}(\omega)\exp\left(-\sum_{t'\in\widehat{\boldsymbol{T}}}\alpha^s_{\omega(t')}\varphi^*_s(\omega(t'),t')\right),$$

$$\sum_{\omega\in\boldsymbol{\Omega}_{k_0e_0}}\pi_{k_0e_0}(\omega)=1, k_0\in\boldsymbol{K}_0, e_0\in\boldsymbol{E}_0,$$

$$\pi_{k_0e_0}(\omega)\geq 0, \omega\in\boldsymbol{\Omega}_{k_0e_0}, k_0\in\boldsymbol{K}_0, e_0\in\boldsymbol{E}_0.$$

ただし $\lambda_s(t)$ はラグランジュ乗数であり，$\boldsymbol{\Omega}^{k_0e_0}_{it}=\{\omega\in\boldsymbol{\Omega}_{k_0e_0}|\omega(t)=i\}$ である．

4 数値例

本節では，3 節で提案した捜索者と目標の最適戦略を求める手法を具体的な数値例に適用し，プレイヤーの最適戦略を分析する．また，情報のもつ価値について，特に目標の初期エネルギーに関する情報に焦点を当て，その価値を定量的に評価する．

4.1 パラメータ設定

図 1 に示す 18 個のセルに分割した捜索空間を，目標が左から右へ移動する場合を考える．各セルの記号 ♯ の後ろにセル番号を記しており，その下の数字は捜索資源ごとの探知効率 α^s_i を示しているが，この詳細については後述する．時間空間は $\boldsymbol{T}=\{1,\ldots,6\}$ とする．

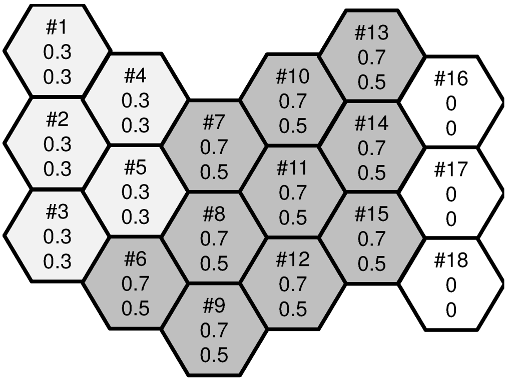

図 1: 数値例のセル空間と捜索資源別探知効率

目標は時点 1 にセル 1 または 3 を出発し，時点 6 までにセル 16,17,18 のいずれかに到達するように移動する．捜索者は目標がどちらのセルから出発するかを確率分布 $f(1) = f(3) = 0.5$ で推定する．目標の初期エネルギーは 5 または 20 であり，捜索者はこれを確率分布 $g(5) = g(20) = 0.5$ で推定する．なお，初期エネルギー 5 の目標は各時点で所在するセルに留まるか隣接セルに移動する通常型，初期エネルギー 20 の目標は各時点で所在するセルに留まるか 1 つ隔てたセルに移動する高機動型である．通常型目標は現在のセルから隣接セルに移動するとエネルギーを 1 だけ消費し，高機動型目標は 1 つ隔てたセルに移動するとエネルギーを 4 だけ消費する．通常型目標は時間内に目的セルに到達するためには常に前進しなければならず，時点 6 で残存エネルギーは 0 となる．高機動型目標は時点 2 から時点 6 まで常に移動すると 20 のエネルギーを消費する．このタイプの目標は目的セルの方向とは逆方向に進んでも時間内に目的セルに到達できる．このように移動したときに両タイプの目標が取り得るパスの数を，表 1 に示す．表 1 は目標のタイプを (初期位置，初期エネルギー) で表し，タイプ毎に取り得るパスの数を示している．パスの数を比べると，通常型目標のタイプ (1,5) 及び (3,5) のパスの数より，高機動型目標のタイプ (1,20) 及び (3,20) のパスの数はかなり多い．

表 1: 数値例における目標のタイプ別パス数

目標のタイプ	パスの数
(1,5)	7
(3,5)	21
(1,20)	679
(3,20)	1028

捜索は時点 2 から開始できるものとし，捜索時間は $\hat{\boldsymbol{T}} = \{2, ..., 5\}$，捜索資源のタイプは 1 または 2 で，目標はこれを確率分布 $h(1) = h(2) = 0.5$ で推定する．探知効率 α_i^s は図 1 の各セル内に示しており，タイプ 1，タイプ 2 の α_i^s を上下に併記している．セル 1 から 5 ではどちらのタイプも探知効率が 0.3 と同じ能力だが，セル 6 以降はタイプ 1 が 0.7，タイプ 2 が 0.5 とタイプ 1 の方が高い能力をもつ．なお，目的セル 16,17,18 では捜索を行わないため探知効率を 0 としている．時点ごとの使用可能資源量はタイプ $s = 1, 2$ に対し，それぞれ常時 $\Phi_1(t) = 1, \Phi_2(t) = 5$ である．

4.2 最適戦略の特徴

この数値例に対し，3 節で提案した解法を用いてゲームの値や捜索者と目標の最適戦略を求めた．その結果，ゲームの値である非探知確率は 0.53 となり，捜索者は目標を 0.47 の確率で探知できる．捜索者の資源タイプごとに目標の非探知確率を表 2 にまとめた．資源量の少ないタイプ 1 はタイプ 2 より全体的に非探知確率が高いが，どの資源タイプでも通常型目標の非探知確率は低く，高機動型目標の非探知確率は高いことが分かる．このようになった理由と，捜索者と目標の最適戦略の特徴を以降で詳細に分析していく．

表 2: 捜索資源タイプ別及び目標タイプ別の非探知確率

	捜索資源のタイプ	
目標のタイプ	タイプ 1	タイプ 2
(1,5)	0.33	0.12
(3,5)	0.51	0.17
(1,20)	0.96	0.58
(3,20)	0.95	0.59

捜索時間 $\hat{T}$ の時点ごとの最適戦略を図 2 に示す．図は，目標の最適パス選択確率をもとに計算した目標存在確率をタイプ別に示し，捜索者の最適資源配分量も資源タイプ別に示している．$e_0 = 5$ の通常型目標と $e_0 = 20$ の高機動型目標の存在確率をそれぞれ白と黒の四角柱で，タイプ $s = 1$ と $s = 2$ の捜索資源配分量を白と黒の円柱で示している．四角柱の高さが高いほど存在確率が高く，円柱の高さが高いほど資源配分量が多い．

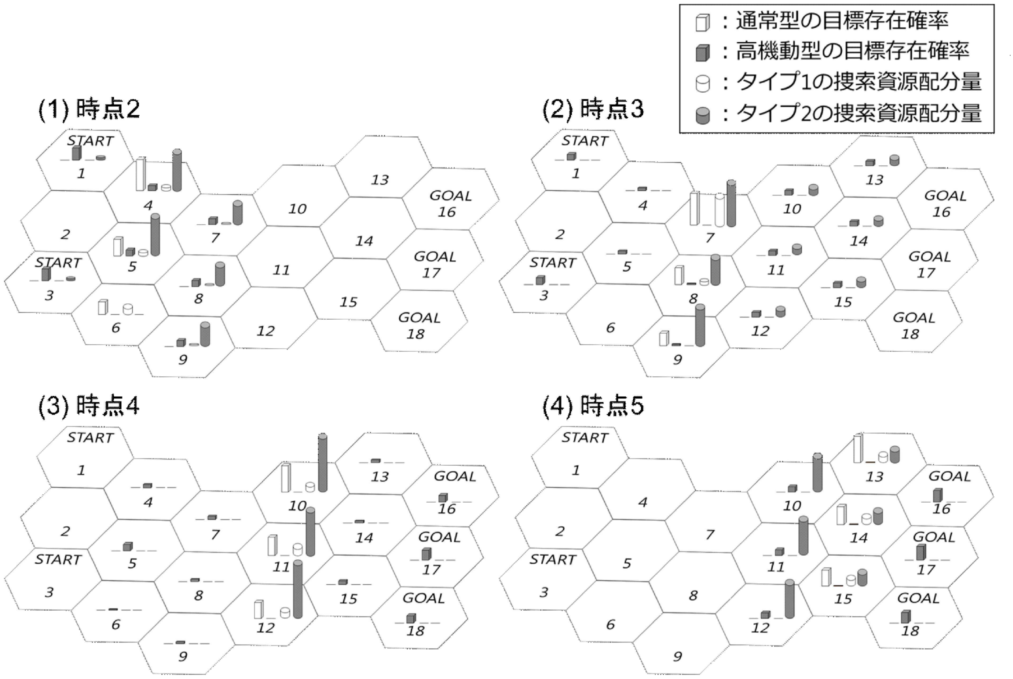

図 2: 目標存在確率と捜索者の最適資源配分

まず，目標の最適戦略の特徴について述べる．各時点の目標存在確率をタイプ別に見ていくと，通常型目標はどの時点においても3つのセルに固まって分布しているのに対し，高機動型目標は広く分布している．この理由は，通常型目標は時点6までに目的セルに到達するには常に前進する必要があり，そのためこのような偏った存在分布となる．しかし，このような分布では集中的な捜索資源配分を受けるため，機動性が高い高機動型目標が広範囲に分布することでそれを防いでいる．

次に，捜索者の最適戦略の特徴について述べる．捜索者は目標の時々刻々の存在範囲を推測して資源を配分するが，正確な目標のタイプを知らないため通常型目標と高機動型目標の双方の動きを考慮して資源を配分する．このとき，表1からも分かるように，通常型目標のパスの本数は少なくその存在圏は限られた範囲になるため，捜索者はどの時点でも存在圏の狭い通常型目標に重点的に資源配分をしている．特に時点3と4ではその傾向が顕著である．そのため，通常型目標の非探知確率は表2のように低くなる．一方，高機動型目標の存在圏は広く，特に時点3と4ではほぼ全てのセルに分布しているため，高機動型目標の探知は困難である．しかし，移動開始直後の時点2と移動終了直前の時点5では，比較的狭い範囲に分布しているため，捜索者はこのタイミングを利用して資源を多く配分している．捜索者と目標の戦略には以上の特徴がある．

4.3 情報の価値

情報の価値は，ある情報について相手プレイヤーも知っている場合，つまりプレイヤー間の共有知識であるときのゲームの値とそれが個人情報であるときのゲームの値を比較して評価する．ここでは，目標の初期エネルギー情報の価値について分析するためエネルギーが5である確率 $g(5)$ を変化させるが，他のパラメータは4.1節での設定を用い，初期エネルギー情報以外の情報についても $f(1) = f(3) = 0.5$，$h(1) = h(2) = 0.5$ の個人情報のままとする．

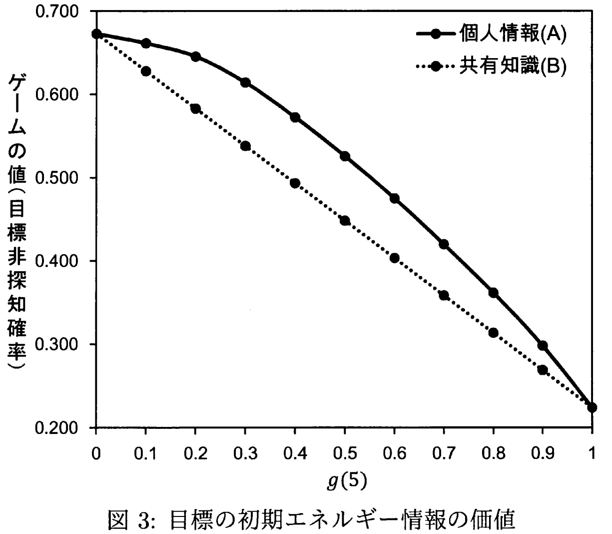

図 3: 目標の初期エネルギー情報の価値

図3に初期エネルギー情報の価値を示す．図では横軸に $g(5)$ を [0,1] 間で 0.1 刻みに変化させたとき，$g(5)$ を個人情報として扱った場合と共有知識として扱った場合のゲームの値を縦軸にプロットして，それぞれ実線と点線で描いた．両曲線の差が初期エネルギー情報の価値を表している．$g(5) = 0$ 及び1では，双方のゲームの値が一致するが，このときは捜索者が正確な目標の初期エネルギーを知っているためである．

情報の価値は，探知が困難な高機動型の確率が高い $g(5)=0.1\sim0.4$ の方が通常型の確率が高い $g(5)=0.6\sim0.9$ より大きく，$g(5)=0.4$ のときに最大となる．これは存在圏が広く捜索者にとって探知困難な高機動型である確率がやや高い情報の方が，捜索者にとって有難いからである．

5 おわりに

本研究では，捜索者と目標の双方が個人情報をもつ捜索ゲームを情報不完備捜索ゲームとしてモデル化し，捜索者及び目標の最適戦略をそれぞれ非線形計画問題及び線形計画問題に定式化して導出する手法を提案した．これにより，確率的に推測するしかない相手の個人情報がゲームの結果に及ぼす影響を定量的に分析する手法を提案できた．数値例による分析から，個人情報がゲームに及ぼす影響を具体的にかつ定量的に評価することができた．

参考文献

[1] T. Matsuo and R. Hohzaki, " A search game with incomplete information about target's energy ", *Scientiae Mathematicae Japonicae*, Vol.80, No.1, pp.57-76, 2017.

[2] R. Hohzaki, "A search game with incomplete information on detective capability of searcher", *Contributions to Game Theory and Management*, Vol.10, pp.129-142, 2017.