

デイトム検索ゲームと実験搜索学

01504810

防衛大学校

宝崎隆祐

HOHZAKI Ryusuke

防衛大学校情報工学科

井田好彦

IDA Yoshihiko

1. はじめに

目標の探知を意図する海難救助活動や軍事行動等において、暴露された目標の位置や時刻情報（デイトム情報と言う）により動機付けられる搜索活動はデイトム搜索と呼ばれ、古典的な搜索理論においても取り上げられている [1]. その後デイトム搜索は目標と搜索者の間でプレイされるゲームとして拡張され、目標のエネルギー制約を考慮した現実的なモデルも登場している [2]. 近年、実験経済学に見られるように、ゲーム理論の分野ではプレイヤーの嗜好や利得構造をシミュレーションその他を用いて実験的に分析し、理論モデルと比較する作業が盛んである。一般の搜索活動においてもその担い手の多くが人間であることを考えれば、現実的なゲーム的狀況下でどのようにプレイヤーの戦略が採られるのかは興味のあるところである。ここでは、デイトム搜索ゲームに関するシミュレータを作成し、既存の理論 [2] と比較した結果を報告する。

2. 理論モデルとゲームの近似解

海難救助等現実的な搜索ゲームの環境として2次元連続空間上で定義される次のデイトム搜索ゲームを考える。

- (1) 搜索空間を2次元平面 R^2 とし、時刻 $t = 0$ に目標は原点（デイトム点）に存在する。
- (2) 搜索者はこのデイトム情報を得て、タイムレイト $t = \tau$ から時刻 T まで搜索オペレーションを実施する。搜索にあたっては単位時間あたり ρ の搜索資源量が利用可能であり、これを任意の地点に分割・投入して目標探知に努める。
- (3) 最初原点に位置していた目標は、時刻 $t = 0$ 以降搜索空間上を連続的に移動するが、速度 v を使用するにあたっては、単位時間あたりエネルギー量 $\mu(v)$ (v の増加関数) を消費する。また、使用速度は最大速度 S_0 を越えてはならず、初期時点におけるエネルギーの総量は E であるとする。
- (4) 搜索者の投入した搜索資源が目標のとった経路をどれだけカバーしたかを利得尺度とするため、目標の存在確率密度で重み付けた搜索資源量の空間上時間上での累積量により期待利得を定義する。問題は、搜索資源の投入戦略をもつ搜索者をこの利得の最大化プレイヤーとし、移動戦略を採る目標を最小化プレイヤーとする2人ゼロ和ゲームである。

問題は原点を中心として点対称であるから位置座標を原点からの距離 $x \in [0, \infty)$ により表す。搜索者の戦略を時刻 t で地点 x に投入する搜索資源の密度 $h(x, t)$ とし、目標の戦略を移動による存在確率密度 $f(x, t)$ とすると、

期待利得は次式で与えられる。

$$G = \int_{\tau}^T \int_{X_t} h(x, t) f(x, t) 2\pi x dx dt \quad (1)$$

ただし X_t は時刻 t における目標存在領域を表す。モデルの前提から、 $h(x, t) \geq 0$ には $\int_0^{\infty} h(x, t) 2\pi x dx \leq \rho$, $\tau \leq t \leq T$ の制約条件がある。また、 $f(x, t) \geq 0$ に対しては目標の連続的移動を可能とすること、及び $\int_{X_t} f(x, t) 2\pi x dx = 1$, $0 \leq t \leq T$ が必要である。目標の純粋戦略としての移動経路を時刻 t における位置 $x(t)$ で表すと、速度 $v(t) = dx(t)/dt$ を使って2つの制約条件、最大速度制約とエネルギー制約を $v(t) \leq S_0$ 及び $\int_0^T \mu(v(t)) dt \leq E$ で表すことができる。さて、(1) 式の積分核である $h(x, t) f(x, t) 2\pi x$ の連続性及び積分範囲の有界性から、期待利得のミニマックス値とマックスミニ値は一致しゲームの値をもつことが知られている。

目標にとって望ましい移動戦略は、 $f(x, t)$ をできるだけ一様にするにより、搜索者に搜索資源の効果的集中的な投入を許さないことである。その第2は、各時点 t における存在領域 X_t をできるだけ大きくし、搜索者の資源投入を広く薄くさせることである。しかしながら、エネルギー制約があるために、目標はこの2つの要件を同時には十分満足させられない。以下ではこの2つの要件を考慮に入れたゲームの値の下界、上界の2つの理論的近似値を述べる。

下界評価 最大速度制約、エネルギー制約という2つの移動制約を満たしつつ時刻 t を終了段階として実現できる目標の最大到達距離 $z(t)$ は変分法によって求めることができる。各時点 t で最大距離 $z(t)$ に到達できてもそこでエネルギーが尽きれば以後停止せざるを得ないが、常にこの最大距離が実現でき、かつその半径円内で一様な存在確率分布が可能であると仮定することにより目標の機動性を過大評価して得られる期待利得が次の下界値 G_L を与える。 $G_L = \int_{\tau}^T \rho / \pi z(t)^2 dt$ 。

上界評価 ある実現可能な目標経路 $y(t)$ があれば、それより近距離を経由する経路も実現可能であるから、目標は常に半径 $y(t)$ 円内での一様な存在分布を作ることができる。このときの期待利得は $\int_{\tau}^T \rho / \pi y(t)^2 dt$ で与えられ、目標はこれを最小にする移動方法を目指すであろう。最大速度制約とエネルギー制約を満たし、かつこの期待利得を最小にする $y(t)$ は、速度 $v(t) = dy(t)/dt$ を制御ベクトルとする最適制御問題を解くことにより求められる。もちろん、この解以上に目標側にとって都合のよい戦略があり得ることを考えれば、解 $y(t)$ を用いた次式はゲームの値の上界値を与える。 $G_U = \int_{\tau}^T \rho / \pi y(t)^2 dt$ 。

3. デイタム検索ゲーム用シミュレータ

前節で提案した理論式を検証するため、二人のオペレータが目標役、探索者役となって対戦するシミュレータを製作し、シミュレーション実験を行う。シミュレータのハードウェアは、RS232Cケーブルで繋がれた2台のパーソナル・コンピュータ(PC)から構成される。この2台のPCを目標用コンソールと探索者用コンソールと見立て、それぞれを一人のオペレータが操作することにより各プレイヤーは探索空間内を移動しながら探索ゲームをプレイする。ただし、オペレータの操作状況はお互いに見ることはできない。

両プレイヤーに対しこのような移動操作を提供する本シミュレータの機能は、必ずしも前節の理論モデル(1)-(4)の仮定には合致しない。模擬できる機能は、2次元平面の探索海域、タイムレイト τ 及びオペレーション終了時刻 T 、さらには目標の最大速度 S_0 や初期エネルギー量 E 、エネルギー消費率 $\mu(v)$ によるエネルギー制約である。これに反し、探索者用コンソールは探索者の移動操作を行わせるためのものであり、理論モデルにおける探索資源の投入戦略を機能としては提供しない。しかし、探索資源は目標探知のために使用するものであることを考え、探知事象をシミュレートする機能を仲介させることにより、理論モデルとシミュレーション結果を比較させることができる。

理論モデルにおける利得は、目標経路上にうまく投入できた有効な探索資源量を示しているが、ランダム検索と呼ばれる一般的な探索オペレーションにおける探知確率は、この有効探索資源量の指数関数で与えられることが知られている。すなわち、ゲームの値 G に対し、探知確率 $P(G)$ は次式で評価できる。

$$P(G) = 1 - \exp(-\beta G) \quad (2)$$

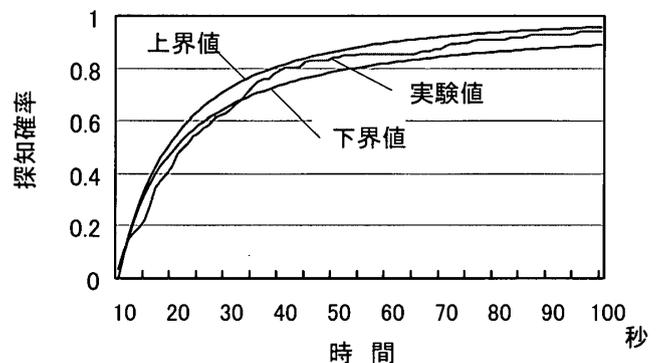
ここで、 β は有効探索資源 G の探知効率性を表す環境パラメータである。一方、シミュレーションにおける探知事象は容易に設定できる。探索理論では、探知センサーの特性を表す値として有効探索幅がある。探索環境が変わらない場合には、センサーによる目標の探知確率は主としてセンサーと目標との距離に依存する。この依存性はそれぞれのセンサーに固有のものであるが、このセンサーを、目標とセンサーが近接しあう場合にある距離 W 以内では確実に探知し、それ以上離れてすれ違っても決して探知できないという理想的なセンサーと見なしたとき、この距離 W を有効探索幅と呼ぶ。したがって、このシミュレータでは探索者の持つ探知センサーの有効探索幅を設定できるようにし、シミュレーション中に目標が探索者の有効探索幅以内に接近すれば探知が起こったとする。本シミュレーションにおいて、両プレイヤーはコンソール操作により平面上を移動し、探知が生じた時点または探索時間が満了した時点で1回のゲームは終了する。

4. シミュレーション実験

デイタム点を中心とする半径1000の円形の探索海域を探索空間とし、 $\tau = 10$ (秒)、 $T = 100$ (秒)、 $S_0 = 50$ /秒、 $E = 10000$ と設定した。また、探索資源の使用率を $\rho = 1$ で、シミュレーションにおける探索者の速度及び有効探索幅をそれぞれ $u = 250$ /秒、 $W = 75$ とした。ゲームの値の下界 G_L と上界 G_U は、エネルギー消費率を $\mu(v) = v^2$ として求めた。

実施したシミュレーション実験では計9名のオペレータが総数166回の探索オペレーションを行い、その平均値としてデータを得ている。このように少人数被験者による多数回のシミュレーションを実施したため、オペレータにはコンソール操作の習熟みならず効果的な戦略実施に関する学習効果が見られたが、これはゲームにおいて暗黙裏に前提とされているプレイヤーの合理的な判断・行動の規範に合致する。

目標側戦略に関する特徴のひとつは、探索開始後の移動速度をどのように変化させるかに見て取れるが、紙数の関係上、探索時間に対する探知確率の変化に関してのみシミュレーション実験と理論式との比較結果を下図に記した。ただし、(2)式の形状パラメータとして $\beta = 30000$ を設定した。前提としたモデルや設定環境の違いにも拘わらず、理論値とシミュレーションの実験結果とはそのおおよその形状が合致している。



4. まとめ

この報告では、デイタム検索ゲームに関する理論研究を人間の介入するシミュレーション実験により検証し、良く合致することを見た。多くの探索活動が人間を意思決定の主体としている現状からは、実データが少ない他の探索ゲームに対してもこのようなシミュレーション検証のような、いわば実験探索学のススメが必要とされる。

参考文献

- [1] B.O. Koopman, *Search and Screening*, Pergamon, pp.221-227, 1980.
- [2] R. Hohzaki and A.R. Washburn, *JORSJ*, **46**, pp.306-318, 2003.