

# 動的経路選択行動の室内実験による検証と分析

池田 愛<sup>1</sup>・宮城 俊彦<sup>2</sup>

<sup>1</sup>非会員 東北大学大学院修士課程前期 情報科学研究科 (宮城県仙台市青葉区荒巻字青葉6-6-6)

E-mail: ai@plan.civil.tohoku.ac.jp

<sup>2</sup>正会員 東北大学 情報科学研究科教授 (宮城県仙台市青葉区荒巻字青葉6-6-6)

E-mail: toshi\_miyagi @plan.civil.tohoku.ac.jp

近年、事故や渋滞などの交通問題に対する解決策としてITSが普及し始め、経路情報とドライバーの経路選択行動の関係を把握する事、より有効な経路情報をいかに提供するかが益々重要な課題になりつつある。

本研究では、経路情報において与える情報の種類や質が経路選択行動に与える影響を分析する事を目的とし、ゲーム論的室内実験の設計と結果の統計的分析を行った。この実験では、リグレットマッチング学習理論に基づき経路情報を提供し、選択行動結果とWardrop均衡との比較やプレイヤーの得る報酬の分析等を通じて、経路選択行動への影響の違いを検証している。その結果、動的経路選択行動はその期待値がWardrop均衡を満足する事、リグレットマッチング理論に基づいた情報提供は、強化学習理論に基づく情報提供よりも人々をより合理的選択へ導くという結果を得た。

**Key Words :** *game theory, learning, regret matching, route choice, economic experiment*

## 1. はじめに

人々の経路選択行動を室内実験によって確認しようとする試みは、Iida, Akiyam and Uchida<sup>1)</sup>によって始められた。彼らは2つの異なる実験を2経路ネットワークを対象に行っている。実験1では被験者が前日選択した経路の実際の走行時間が被験者に与えられるが、実験2では実走行時間と予測走行時間が与えられる。実験では理論的均衡値の付近で大きな変動を示しており、定常状態に至る様子はほとんど観測できなかつた。しかし、彼らの実験では反復回数は20回程度と非常に少なく、被験者が道路環境を学習するには極端に短いと考えられる。

Selten<sup>2)</sup>も2経路ネットワークを対象に同様の室内実験を行っている。実験1では、利用した経路の走行時間のみが与えられるが、実験2では利用していない実験の経路の所要時間も被験者に与えられる。Selten<sup>3)</sup>ら実験では累積所要時間が情報として与えられる。この背景には、Elev and Ross<sup>3)</sup>のような実験経済学の分野では繰り返し実験の際の報酬に累積報酬を与えていることが考えられる。Selten<sup>4)</sup>らはElev and Rossの強化学習モデルを用いてシミュレーションし、実験結果を良く再現できることを報告している。Selten<sup>5)</sup>らは200回の反復実験を行っているが、Iida<sup>6)</sup>らの実験結果と同様、理論的均衡値の周辺で変動し、定常状態に収束することはなかつた。Helbing<sup>7)</sup>は、Selten<sup>8)</sup>らの変動の原因が反復回数にあるのではない

かと判断し、更に反復回数を増やしているが定常状態に収束することはなかつた。しかし、分散は減少する傾向が観測されている。その後、Amnon Rapoport<sup>9)</sup>、John L.Hartman<sup>10)</sup>らが室内実験を用いた交通問題の検証を行っている。

繰り返し実験において、どのような情報を被験者に提供するかが重要な課題である。これまでの実験では累積報酬、前日の経験値が利用されている。無論、平均値を与えるということも考えられる。

本研究は、室内実験を通してドライバーの日々の経路選択行動を調査することを目的としているが、リグレット・マッチング・ゲームに基づく経路選択行動を検証しようとする意味ではSelten<sup>11)</sup>らのゲーム論的実験に近い立場をとっている。また、Selten<sup>12)</sup>らの実験1と実験2は、Miyagi<sup>13)</sup>が提案したnaïve driver problemとinformed driver problemに対応している意味で、交通情報の与える経路選択行動が検証できる。

Miyagi<sup>14)</sup>は強化学習に基づく経路選択行動モデル<sup>8,9)</sup>及びリグレット・マッチング理論に基づく経路選択行動モデル<sup>7,10,11)</sup>を提案している。強化学習モデルはnaïve driver problemに対応したモデルであるのに対し、リグレット・マッチング・モデルはinformed driver problemに対応している。すなわち、前者はドライバーが利用した経路の利得情報しか知らない場合で、後者は利用しなかつた経路の利得も知り得る場合である。リグレット・マッ

ング理論は Hart&Mas-colell<sup>12)</sup>によって提案された。この理論の重要な点は、プレーヤーは利得情報しか知らず、他のドライバーの行動は観測できないことを前提にしている点である。Miyagi<sup>7), 10), 11)</sup>は色々なタイプのコスト関数を用いてテストしており、学習行動を内包したこの理論が経路選択モデルに有用であることを示しているが、実証的研究にまでは至っていない。

そこで本研究では、Hart&Mas-colell のリグレットマッチング理論による学習をベースとした経路選択モデルに基づき実験を行い、実験結果と理論的研究結果の整合性を検証することで経路選択を行う際の学習理論の有用性を確認する。また、情報の与え方の異なる実験を行い比較することで、提供する情報がドライバーの経路選択行動にどのような影響を与えているのかを検証していく。

## 2. 室内経路選択実験の設計

### (1) 実験の構成

#### a) 概要

一回の実験に 18 名のプレーヤーが参加する。プレーヤー（被験者）は主に東北大学の大学生と大学院生である。この 18 名のプレーヤーは、コンピュータ上で仮想的な経路のもとで繰り返し経路選択を行う。プレーヤーは、経路図を見ながらそれぞれのコンピュータ画面上に表示される情報をもとに、どの経路がより早く行けるかを予測して毎回経路選択を行う。そして、事後的に得られる経路選択結果の情報から、より早く行くことのできる経路を予測してもらう。より短い所要時間でたどり着くことができれば、プレーヤーはより多くの報酬を得ることができる。

プレーヤーには図 1 に示した交通ネットワークが提示される。経路のコスト関数は図 2 に示す線形関数を用いているが、実験参加者はコスト関数を知ることができない。コスト関数は、Selten らの実験と比較するため同じ関数を用いている。また、図 1 も Selten らの実験で用いたものと同じネットワークである。なお、3 経路の場合のネットワーク図は付録に載せる。3 経路では 2 経路の場合に比べやや複雑なネットワークを想定し、不確実性が大きくなる場合の経路選択行動を分析する。

プレーヤーは、報酬に換算された所要時間を経路情報として与えられるが、コスト関数、他のプレーヤーの選択行動は知ることができない。また、ゲームへの参加者数は知らせてはいないが、ある程度は推測できていると判断できる。つまり、プレーヤーはゲームの構造を完全には把握していない不完全情報ゲームを想定しており、プレーヤーは常に合理的行動をとることができない。このような状況を想定して経路選択行動を繰り返して行う中での、

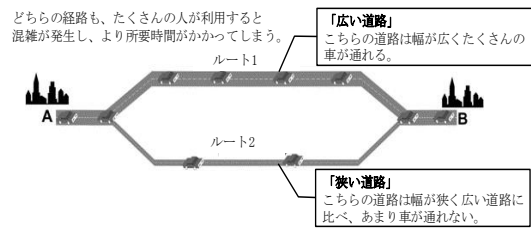


図1 実験中にプレーヤーに示す経路図 (2経路)

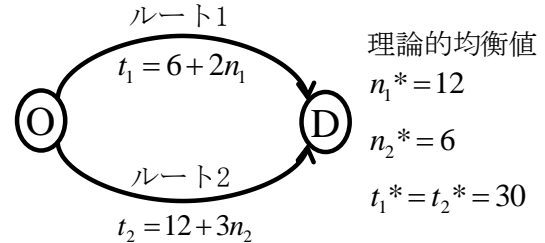


図2 経路のコスト関数と理論的均衡値 (2経路)

プレーヤーの行動の変化や様々な情報からの学習について分析を行う。

なお、今回行った全ての実験は、チューリッヒ大学で開発された経済実験用ソフトウェア z-Tree によって構築したプログラムを用いて行った。

#### b) 情報提供方法

プレーヤーは経路選択を行った後、事後的に経路コストに関する情報(以下、Feedback)を得ることができるが、それは情報提供方法に依存する。

**Treatment1:**プレーヤーは自分の選択した経路コストに関する情報のみを得ることができる場合。

**Treatment2:**プレーヤーは自分の選択した経路コストに関する情報のみならず、選択しなかった他の経路の情報も得ることができる場合

この2つの実験を比較することで、情報提供の有無が経路選択行動に与える影響を分析する。

繰り返し実験では、過去に得られた経路情報をどのような形で提供するかが問題になる。強化学習は Treatment1 を対象にした理論であり、累積の報酬を与えるのが一般的である。Ross & Erev らの実験経済学者あるいは Selten らの実験でも累積報酬が利用されている。一方、Hart&Mas-Colell の提案するリグレットマッチング理論は主に Treatment2 の状況を想定した理論である。この場合、“リグレット (より正確には Hannan リグレットあるいは外部リグレット)” という概念を被験者に説明する手間を省くため、これを“推奨度”と翻訳して情報提供する。一方、リグレット理論をそのままの形で Treatment1 にも適用することはできない。可能ではあるが、理論に忠実に従うリグレット情報を提供するには被験者が現実には行わないような計算プロセスを加える必

表1 実験方法詳細

実験	実験1-1	実験2-1	実験1-2	実験2-2	実験A-1	実験A-2
ベースの学習モデル	リグレットマッチング理論(RMT)			強化学習モデル(RT)		
実験方法	Treatment1		Treatment2		Treatment1	Treatment2
被験者数	18名					
繰り返し回数	150ピリオド				100ピリオド	
経路数	2経路	3経路	2経路	3経路	2経路	
経路選択 ステージ	現段階のピリオド数					
	自分の選択した経路に関する情報					
	前ピリオドでの報酬額			前ピリオドでの所要時間		
	そのピリオドまでの平均報酬額			そのピリオドまでの累積報酬額		
他の経路に に関する情報	そのピリオドまでの各ルートの平均報酬額		各経路の推奨度		前ピリオドでの各ルートの所要時間	
	自分の選択した経路					
経路選択 結果表示 ステージ	被験者自身の選択結果					
	自分の選択した経路の所要時間					
	そのピリオドで獲得した報酬額					

要がある. この問題を回避するため, 本研究では宮城・石黒が提案した近視眼的リグレットを Treatment1 の場合には採用した.

ここで N 人ゲーム  $\Gamma = (N, (S^i)_{i \in N}, (u^i)_{i \in N})$  として,  $S^i$  をプレイヤー i の取りうる戦略集合,  $k \in S^i$  をプレイヤー i の戦略,  $u^i$  をプレイヤー i の効用とし, 繰り返し回数  $t=1, 2, \dots$  とする時, Treatment2 で用いた推奨度 R は次式で定義される.

$$\begin{aligned}
 DH_t^i(k) &= \frac{1}{t} \sum_{\tau \leq t} \{u^i(k, s_\tau^{-i}) - u^i(s)\} \\
 &= \frac{1}{t} \sum_{\tau \leq t} u^i(k, s_\tau^{-i}) - \bar{u}^i(s) \quad (1) \\
 &= (t \text{ 期までの経路 } k \text{ の平均報酬額}) \\
 &\quad - (t \text{ 期までの自分の平均報酬額}) \\
 &= \text{推奨度 } R
 \end{aligned}$$

$$q_{t+1}^i(k) = \frac{\{DH_t^i(k)\}_+}{\sum_{k' \in S^i} \{DH_t^i(k')\}_+} \quad (2)$$

$DH_t^i(k)$  をプレイヤー i の Hannan リグレットと呼ぶ.

$\bar{u}^i(s_\tau)$  はプレイヤー i の平均利得である. 上式右辺第 1 項は過去の選択において常に戦略  $k$  を選択していた場合に成立する平均利得とみなすことができ, 第 2 項は実際の平均利得である. Hannan リグレットに基づいて行動するという事は今までの過去の選択の履歴と 1 つの選択  $k$  を比較することである.

リグレットマッチング理論では, Hannan リグレット DH の大きい経路のほど, より大きな選択確率が割り振られる. したがって,  $R_H$  は経路選択において推奨する度合いの大きい経路になる. ただし, リグレットマッチング理論では, プレイヤーがこの推奨度の大きい経路を必ず選択することを仮定してはいない. プレイヤーには, 「推奨度とは, 値が大きければ大きいほどより早く目的地まで行ける確率が高い経路であることを示す」と説明する. また, この推奨度の値はあくまでも過去の結果からの予測であり, 必ず結果がその推奨度の通りになるとは限らないことも知らせる. また, 理論的には DH の大きさに基づいて次期の戦略  $k$  の選択確率  $q$  を算出し, この確率に基づきプレイヤーは行動すると仮定している.

一方, Treatment1 で使用する近視眼的リグレット  $MH_t^i(k)$  は次式で定義される.

$$MH_t^i(k) = u^i(k, s_t^{-i}) - \bar{u}^i(s) \quad (3)$$

$$q_{t+1}^i(k) = \left(1 - \frac{\delta}{t^\gamma}\right) \frac{\{MH_t^i(k)\}_+}{\sum_{k' \in S^i} \{MH_t^i(k')\}_+} + \frac{\delta}{t^\gamma} \frac{1}{|S^i|} \quad (4)$$

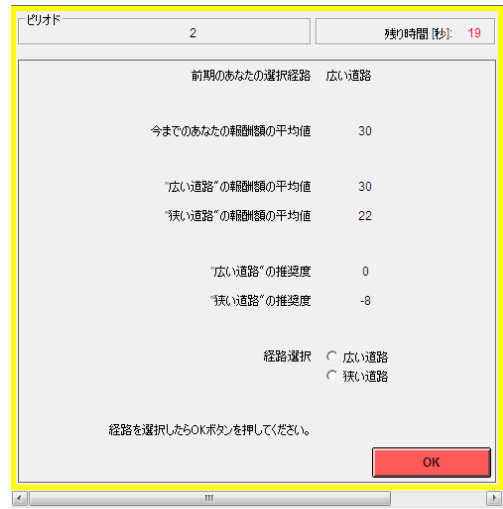


図3 経路選択ステージ画面 (Treatment2)

Hannan リグレットは今までの過去の選択の履歴と 1 つの過去に選択  $k$  をとり続けた場合の比較であったが, 近視眼的 Hannan リグレットは現在の選択  $k$  と現在の平均値の比較である. ここで,  $|S^i|$  はプレイヤー i の選択可能な戦略数である  $\gamma^i \in (0, \frac{1}{2})$  は学習パラメーター,

$\delta^i \in (0, 1)$  はプレイヤー i 固有の探索パラメーターである.

### c) 実験の流れ

各実験での提供する情報, 実験方法の詳細な内容をまとめたものを表1に示す. 表1では2経路を対象にした場合を実験1, 3経路の場合を実験2と分類している. Seltenらの実験の追試を行っているが, これを実験Aとしている. また, 繰り返し反復される実験における1回の試行をピリオドと呼ぶ.

1ピリオドは2つのステージで構成されている.

- 1) 経路選択ステージ (20秒) (図3)
- 2) 経路選択結果表示ステージ (10秒)

実際の経路選択実験の際に表示されるステージの画面を図3に示す。経路選択ステージでは、それぞれの実験で決められたfeedbackの内容を表示する。ここで表示される情報は、前ピリオドまでの結果を反映している。プレイヤーは、その情報を参考に、次にどの経路を選択したらより早く目的地にたどり着けるかを予測し、繰り返し経路選択を行ってもらふ。経路選択結果表示ステージでは、そのピリオドの選択結果のみを表示している。本実験では、繰り返し回数を150ピリオドとした。

#### d) 予備実験

Selten らの実験と比較するため、彼らと同じ実験を予備実験として行った。予備実験では、強化学習理論をベースとした情報提供を行うため、本実験で提供する情報とは別の情報を与える。主に以下の二点が異なる。

Feedback の内容の相違点

- ・前ピリオドまでの報酬額の平均値  
→ 前ピリオドまでの累積報酬
- ・ピリオドまでの全経路の「推奨度」  
→ 前ピリオドでの全経路の報酬

(ただし Treatment2 の実験のみ)

前述したように、Selten らの実験は Erev and Roth の強化学習モデルを前提に実験が構成されているので、強化学習モデルとリグレットマッチングモデルの比較検証することができる。

予備実験で得られた経路利用者数のみの結果を表 2 に示す。

### 3. 室内経路選択実験分析結果

リグレットマッチング理論に基づく 2 経路での実験結果について述べる。

#### (1) 経路選択行動におけるリグレットマッチング理論の整合性の検証

##### a) 利得情報とWardrop均衡

ピリオドごとの経路利用者数に着目する。図 4 より理論的均衡値の 6 へは完全に収束することなく振動し続けていることが分かる。しかし、前半の振動に比べて、後半の振動は小さくなっていることが読み取れる。経路利用者数の二乗誤差を示したのが図 5 である。振動が徐々に小さくなっていることが読み取れる。しかし、平均値は表 3 に示す理論的均衡値にとっても近い値を示している。経路利用者数の度数分布を示す図 6 より、どの実験においても均衡値付近での頻度が高く、また頻度分布は正規分布している。これより、全体の平均値が均衡値に近い値を示しているだけでなく、その頻度も高いことがわかる。この現象から推論できることは、動的、確率的に変

表 2 統計的結果 (実験 A-1, A-2)

実験	理論的均衡値	ルート2利用者数平均	標準偏差
実験A-1	6人	5.96	2.02
実験A-2		6.03	1.88

表 3 統計的結果 (実験 1-1, 1-2)

実験	理論的均衡値	ルート2利用者数平均	標準偏差
実験1-1	6人	6.13	1.88
実験1-2		6.03	1.41

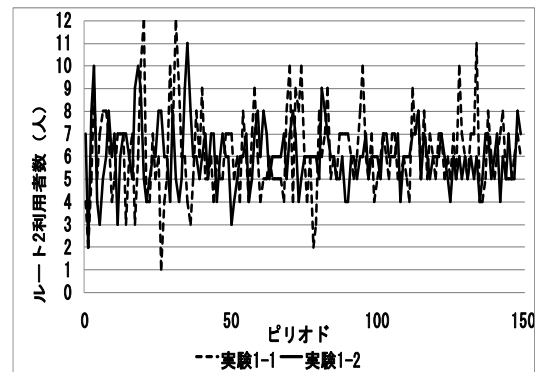


図 4 ルート 2 利用者数の推移 (実験 1-1, 実験 1-2)

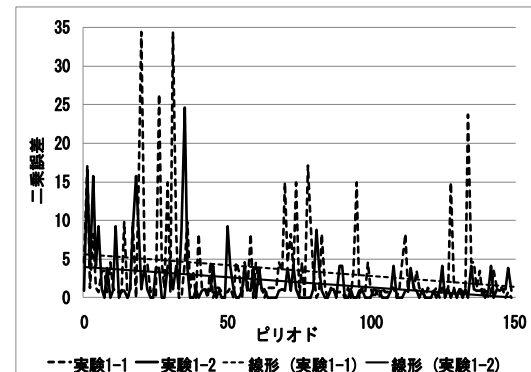


図 5 経路利用者数の二乗誤差 (実験 1-1, 実験 1-2)

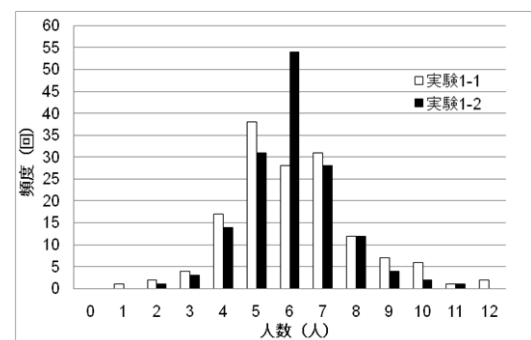


図 6 ルート 2 利用者数の頻度分布 (実験 1-1, 実験 1-2)

動する経路選択行動は、期待均衡として Wardrop 均衡が実現するというのである。

## b) リグレットの収束性

図 7 に実験 1 の Treatment2 の場合のリグレット平均値の挙動を示す。そもそもリグレットマッチング理論では、人々が繰り返し行う中で学習することで式 (1) から算出される全プレイヤーのリグレット  $R$  の平均値は 0 に漸近することがわかっている。図 7 より、最初は値が大きかったリグレットの値も、実験の繰り返し回数の増加に伴い 0 に漸近していくという理論同様の傾向があると推測される。つまり、事後的にリグレットの情報を得ることで人々は適応行動をとり、より最適な選択をしていくという傾向があることが実験的に示された。

## (2) 情報提供の有効性の検証

### a) 利用しなかった経路の情報提供の効果

表 3 や図 4 からわかるように、他経路の情報を与えた場合の方が与えなかった場合よりも利用者数の平均値は均衡値に近く、また標準偏差の値も小さい。Wilcoxon-Mann-Whitney 検定 (以下 WMW 検定) において有意水準 5% の両側検定を行った結果、実験 1-2 の方が標準偏差は有意に小さいという結果が得られたことから同様のことが言える。実験 1, 実験 2 そして実験 A の全ての実験に対して同様の結果を得た。つまり、自分の選択していない経路に関する情報を得られるということは、次の経路を選択していく際に非常に大きな影響を持ち、有用な情報だと言える。

次に、他経路情報の有無によるプレイヤーの経路変更回数に注目する。図 8 より、他経路の情報を得られることで、繰り返し回数の増加とともに経路変更者数は大きく減少していることがわかる。経路変更回数の平均値において WMW 検定を行った結果、有意水準 5% の両側検定より実験 1-2 の平均値は実験 1-1 比べて有意に小さいという結果を得た。図 9 はプレイヤー毎の経路変更回数と最終報酬額の関係を示した図である。経路変更回数が少ないプレイヤーほど獲得報酬額は増加する傾向が確認できる。これはスピアマンの順位相関係数における有意水準 10% 以下の両側検定により、相関関係がある事が確認できた。システム全体での報酬額の変化に着目すると図 10 の様になった。これより、どちらの実験でもシステム全体での報酬額平均値は繰り返しに伴い均衡値へと収束していく傾向が確認できた。さらに実験 1-2 ではよりシステム最適状態へ近づく結果となった。つまり、プレイヤーは情報を得られることで学習を行い、限られた情報の中で適応的行動をとり、できる限り最適な選択を行っているかと推測できる。このプレイヤーの学習には情報の有無が影響しており、他経路の情報を得られることでより早くシステムが均衡へ収束する傾向があ

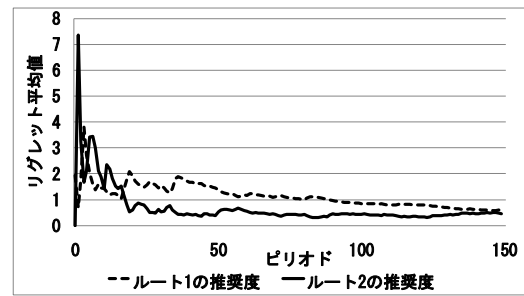


図 7 リグレット平均値の推移 (実験 1-2)

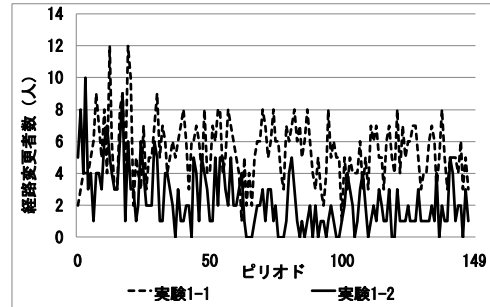


図 8 ピリオド毎の経路変更者数 (実験 1-1, 実験 1-2)

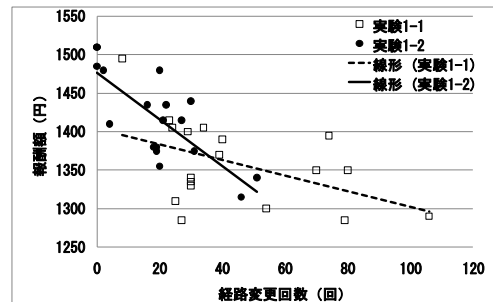


図 9 経路変更回数と最終報酬額の相関関係

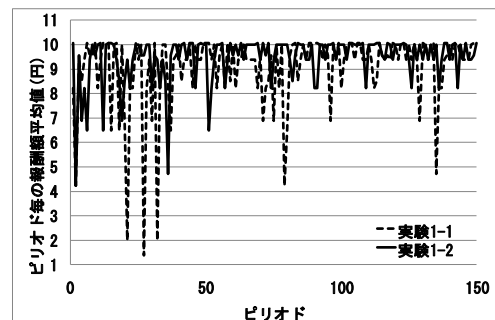


図 10 ピリオド毎の平均報酬額の推移

ると考えられる。

### b) 経路選択行動の合理性

交通情報提供そして情報提供の質の違いによってプレイヤーの学習行動がどう変化するかを見るためにユール係数を利用する。まず、以下のような 2 つのプレイヤー

の反応モードを仮定する。

- モード1: 前ピリオドの報酬に応じて行動する  
「合理的行動」 (C-,S+)
- モード2: 前ピリオドの報酬とは無関係に行動  
する「非合理的行動」 (C+,S-)

すなわち、推奨度が大きいあるいは推奨度が小さいという情報を得たとき、推奨度の大きい経路へ変更する行動をとったとき、あるいは小さい経路から他経路に変更した行動をとったとき、合理的行動をとったと判断する。逆のケースは非合理的と判断するのである。言い換えれば、経路2の方が早く行けるという情報を得たのに、経路1を選択した場合には、このプレイヤーは非合理的行動を行ったと判断する。無論、プレイヤーは常に合理的行動をとり続ける訳ではない。逆もそうである。したがって、頻度で判断する必要がある。ここでは、均衡値での報酬額10と前ピリオドでの報酬額を比較して分類を行う。この分類結果より、それぞれのプレイヤーのユール係数を計算し、実験全体を通したプレイヤーの行動が合理的か非合理的かを決定する。ユール係数は以下の式で定義できる。

$$Q = \frac{c_- \times s_+ - c_+ \times s_-}{c_- \times s_+ + c_+ \times s_-} \quad -1 < Q < 1$$

なお、ユール係数は50ピリオド毎の頻度ベースに計算している。また、今回は以下の場合分けによって行動の分類を行っている。

- (i)  $Q > 0.5$  ⇒ 合理的行動
- (ii)  $-0.5 < Q < 0.5$  ⇒ 分類なし(曖昧な行動)
- (iii)  $Q < -0.5$  ⇒ 非合理的行動

計算結果を図11に示す。実験1-1に比べ実験1-2では合理的選択をしたプレイヤーの割合が多いことがわかる。また、実験1-1では、分類なしに属しているプレイヤーの割合が多い。この結果より、他経路の情報を得られることで、プレイヤーはより合理的行動ができると推測できる。また、自分の選択した経路情報のみの場合には、プレイヤーの選択における不確実性が大きいと推測できる。

これらの結果より、人々の経路選択において、他経路の情報を提供することは効果的であり、これによりシステムのより早い収束が見込めるとともにシステム全体での混雑の減少が期待できると考えられる。

### (3) リグレットマッチングモデルの有効性

ユール係数を用いて経路情報の提供の仕方の違いによる学習効果をリグレットマッチングモデルと強化学習モデルで比較する。頻度の取り方は前と同様50ピリオド

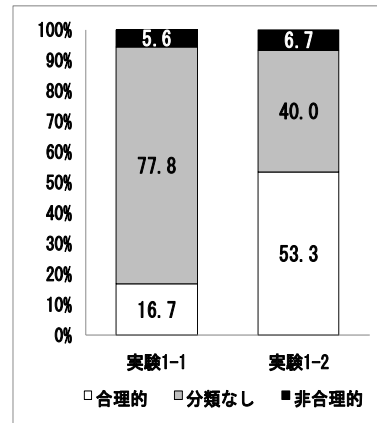


図11 反応モードの分類結果

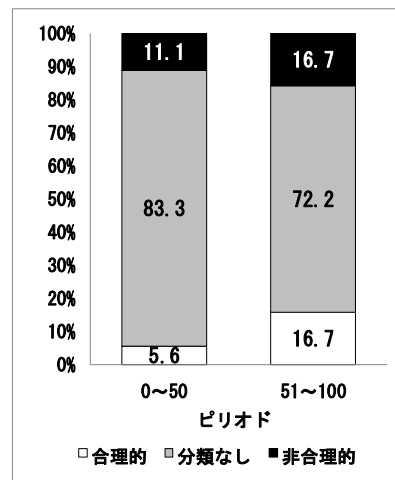


図12 「所要時間」情報に対する行動方法の推移

(実験 A-2)

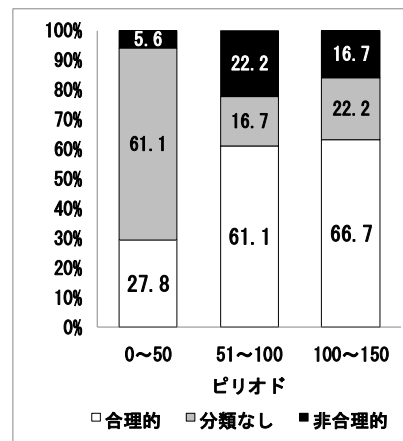


図13 「推奨度」の情報に対する行動方法の推移

(実験 1-2)

毎とした。結果をそれぞれ図12,13に示す。実験A-2では、所要時間の情報に対して明確に合理的な行動をとっている人は前半、後半ともに20%以下である。図12より「分類なし」の割合がほとんどで、曖昧な行動をし

ている人が多いと考えられる。ただし、前半の 50 ピリオド後半の 50 ピリオドを比較すると、後半の方がやや合理的な行動をしている人が増加している。これに対して実験 1-2 では、情報として与えられている推奨度の値に対し合理的に行動している人の割合が全体的に高いことがわかる。また、時間の経過とともに、情報に対して合理的に行動する人の割合が大きく増加している。また、実験 1-2 と A-2 の実験結果を比較すると、振動の大きさは実験 1-2 の方が小さくより早い収束傾向を示し、プレーヤー全体ではより高報酬額を獲得していた。これらの結果より、所要時間という情報に対して、推奨度という一つの指標となる情報を提供することで、人々の行動をより合理的な行動へと導くことができると推測できる。また、推奨度という情報は経路選択の際に有効的で、人々の学習を促す効果が高いと考えられる。

#### (4) 3経路のケース

3経路の場合は、2経路と類似した傾向が得られた。しかし、2経路の場合に比べ不確実性が高いため経路選択のばらつきは大きく、収束スピードも緩やかである。経路利用者数の理論値と実験の平均値を表4に示す。また、経路利用者数の推移結果については付録に載せる。

### 4. おわりに

本研究では、人々の学習に着目し、情報提供と人々の経路選択にはどのような相関関係があるのかについて、経済実験を応用し分析を行った。人々の限定合理性を仮定したリグレットマッチング学習理論に基づいた情報提供を行い、日々の通勤を例とした不完全情報下での繰り返しゲームにおける人々の経路選択動向を観察した。

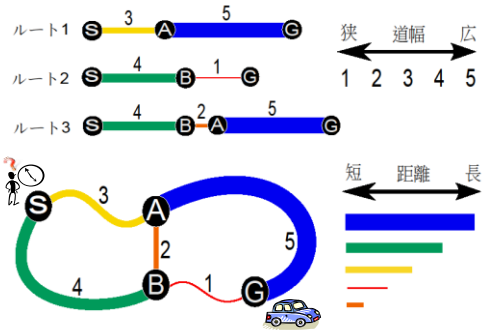
その結果、プレーヤーが経路の利得情報のみを知り得るという条件の下では、Wardrop 均衡は日々達成される均衡ではなく、期待値として達成される期待均衡であると判断できた。無論、実験の反復を非常に大きく取れば定常状態に近付くかも知れないという期待はあるが、これを実験的に確認するのはかなり無理がある。また、自分の選択していない経路の情報を得られることで、システムの収束は早く人々はよりよい選択行動を行えるということが分かった。また、強化学習をベースとした実験結果と比較すると、リグレットマッチング理論をベースとした学習理論に基づく情報提供の方が、人々はより合理的な経路選択行動をとる傾向があると分かった。

**謝辞：**本研究は文部科学省科学研究科補助費の平成22～24年度基盤研究(B)の援助を受けて行われた研究成果の一部である。

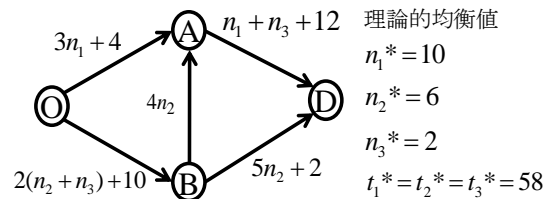
表4 経路利用者数の統計結果 (実験2-1, 2-2)

ルートNO.	理論的均衡値	実験方法	平均値	標準偏差
ルート1 利用者数	10人	Treatment1	9.87	2.13
		Treatment2	9.91	1.78
ルート2 利用者数	6人	Treatment1	5.73	1.84
		Treatment2	5.86	1.90
ルート3 利用者数	2人	Treatment1	2.39	1.51
		Treatment2	2.23	1.31

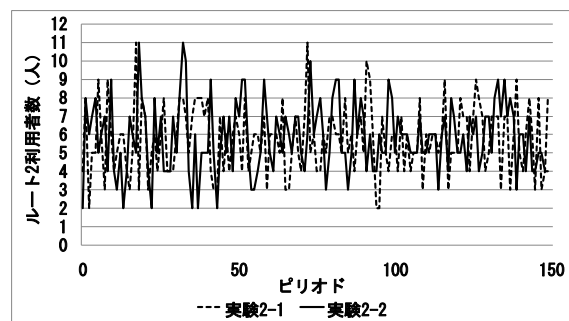
#### 付録



図A-1 実験中にプレーヤーに示す経路図 (3経路)



図A-2 経路のコスト関数と理論的均衡値 (3経路)



図A-3 ルート2利用者数の推移 (実験2-1, 実験2-2)

#### 参考文献

- 1) Iida, Y., T. Akiyama and T. Uchida: Experimental analysis of dynamic route choice behavior, *Transportation Research Part B*, Vol.26B, No.1, pp.17-32, 1992.
- 2) Selten, R., Chmura, T., Pitz, T., Kube, S. & Schreckenberg, M.: Commuters route choice behavior, *Game and Economic Behavior* 58, pp.394-406, 2007
- 3) Erev, I., Roth, A.E., Rewdicting how people play games: Reinforcement learning in experimental games with unique mixed strategy equilibria.

- Amer. Econ. Rev. 88(4), pp848-881, 1988.
- 4) Helbing, D.: Dynamic Decision Behavior and Optimal Guidance through Information Services: Models and Experiments, *Human Behaviour and Traffic Networks*, pp47-95, 2004.
  - 5) Rapoport, A., Kugler, T., Dugar, S., Gisches, E.J.: Choice of route in congested traffic networks: Experimental tests of the Braess Paradox, *Game and Economic Behavior* 65, pp538-571, 2009
  - 6) Hartman, J.L.: A Route Choice Experiment With an Efficient Toll, *working paper*, 2007
  - 7) T. Miyagi: Multiagent learning models for route choices in transportation networks: An integrated approach of regret-based strategy and reinforcement learning, Proceedings of the 11<sup>th</sup> International Conference on Travel Behavior Research, Kyoto, 2006.
  - 8) Miyagi T.: A modelling of route choice behaviour in transportation networks: an approach from reinforcement learning”, WIT Press, Southampton, UK, pp. 235-244, 2004.
  - 9) Miyagi T.: A Reinforcement Learning Model with Endogenously Determined Learning-Efficiency Parameters: Applications to Route Choice Behaviour in Congested Networks”, Proceedings of Joint 2<sup>nd</sup> International Conference on Soft Computing and Intelligent Systems and 5<sup>th</sup> International Symposium on Advanced Intelligent Systems, 2004.
  - 10) T. Miyagi: Modeling of Adaptive Learning Behavior of Drivers in Dynamic Environment, Proc. of 14th International Conference on Recent Advances in Retailing and Consumer Services Science, San Francisco, US, 2007.
  - 11) T. Miyagi and M. Ishiguro: Modelling of Route Choice Behaviours of Car-Drivers under Imperfect Travel Information, (an invited paper), Proc. of 14th International Conference on Urban Transport and Environment, pp. 551-560, WIT Press, 2008.
  - 12) Hart, S. & Mas-colell, A.: A simple adaptive procedure leading to correlated equilibrium, *Econometrica* 68, pp.1127-1150, 2000.

(2011.8.5 受付)

## Experimental Design and Investigation of Day-to-Day Route-Choice Behavior in Network

Ai IKEDA, and Toshihiko MIYAGI

Recently ITS system become widely spread as a solution against traffic issues such as traffic accidents and traffic jam, so it become more important to find relationship between routing information and choice behavior of drivers to the information and to provide more effective information. In this paper, we design gametheoretical laboratory experiments and perform statistic analysis to analyze the effects the types and quality of routing information on route choice behavior. In our experiments, we provide information based on the “Regret-matching” learning theorem, and verify the differences of the effect on route choice behaviors through comparison of observed value with theoretical value “wardrop equilibrium” and observing driver’s payoff. As a result, we show that the expectation values of the dynamic route choice behavior tend to converge to wardrop equilibrium and information supplement based on Regret-matching theorem leads driver’s choices to more reasonable options compared to one based on reinforcement learning theorem.