

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 5 月 20 日現在

機関番号：11301

研究種目：基盤研究（B）

研究期間：2010～2012

課題番号：22360201

研究課題名（和文）経路選択行動における強化学習理論とアルゴリズムに関する実証的研究

研究課題名（英文）Theory of Reinforcement Learning and Algorithms of Route Choice in Transportation Networks

研究代表者 宮城 俊彦 (MIYAGI TOSHIHIKO)

東北大学・大学院情報科学研究科・教授

研究者番号：20092968

研究成果の概要（和文）：ゲーム的学習理論を用いてエージェントの日々の経路選択行動を定式化し、不完全情報下における選択行動モデルを誘導するとともに、Nash 均衡への収束性を理論的に明らかにしている。また、理論と統合的な適応学習アルゴリズムを開発し、アルゴリズムの妥当性について数値実験を用いて検証している。加えて、経路選択モデルの構造パラメータを推定する方法を提案した。室内実験によって得られた day-to-day 経路選択データへの適用は満足いくものであった。

研究成果の概要（英文）：This research shows that an individual traveler in transportation networks is rigorously modeled as an adaptive learning agent who receives travel information through day-to-day experience and makes his decision so as to reinforce his action depending the realized payoffs. An adaptive learning algorithm consistent with the theory is proposed and proved that it leads the system to a Nash equilibrium with probability one. The proposed algorithms have tested numerically by using example networks with various ill-defined link cost functions and examined a rapid convergence of the algorithms. In addition, we have proposed an estimation method for the structure parameters included in the route choice model. The application to the data of the day-to-day route choice obtained by the indoor experiments was satisfactory.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010 年度	2,800,000	840,000	3,640,000
2011 年度	1,700,000	510,000	2,210,000
2012 年度	1,800,000	540,000	2,340,000
総計	6,300,000	1,890,000	8,190,000

研究分野：工学

科研費の分科・細目：土木工学・土木計画学・交通工学

キーワード：繰り返しゲーム、強化学習、交通行動理論、適応学習アルゴリズム、Nash 均衡、利用者均衡確率近似理論、動的離散的選択モデル

## 1. 研究開始当初の背景

経路選択行動の理論は、交通計画における公共路線計画や道路計画の需要予測に用いられる交通量配分の基礎を与える点で重要であり、長い間、交通工学の分野で研究されてきた。しかし、従来の交通量配分手法は、トリップを無限分割可能な変数として扱い、利用者を一様な意思決定者として扱ってきた。また、利用者は他者の行動や経路に関し完全な情報を得ているという仮定で構成されている。この結果、経路選択行動は完全合理的であり、数理最適化問題として定式化できるものとして理論が構成され、アルゴリズムが構築されてきた。しかし、こうした完全合理的トリップメーカーの仮定については、交通量配分モデルが提案された当初からの疑念であり、また、最近の室内実験結果では、人々は経路情報を得ている場合でも、当初予期されたような合理的な行動はとらず、したがって、Wardrop(利用者) 均衡は北位置的にしか実現しないことなどが明らかになり、実際の経路選択行動を理論とアルゴリズムの両面から見直す必要が生じている。

## 2. 研究の目的

本研究の目的は、利用者の交通情報は不完全であるという仮定に立ち、走行経験を通して環境を学習し、より良い結果を強化していく学習行動を基礎とした交通行動理論を確立し、その理論を基礎にした適応的選択行動アルゴリズムを開発することである。同時に、実験経済学的手法を用いた経路選択鼓動の室内実験を設計するとともに、実験によって得られる動的データを用いた行動モデルのパラメータ推定そして均衡計算の妥当性を実証することを目的にしている。要約すると、以下のようなものである。

(1) 離散的エージェント(atomic 仮定)、不

完全情報を仮定した経路選択行動モデルの構築

(2) 動的適応過程としての強化学習理論に基づくネットワークシミュレーションアルゴリズムの開発

(3) 経路選択行動の室内実験の設計とデータを用いた構造パラメータ推定法の提案

## 3. 研究の方法

### (1) 前提

利用者の経路選択問題を以下の2つに大別する。

① **Informed User (IU)問題**：実際には利用しなかった経路の情報を事後的に知ることができ、この情報に次期の選択に利用できる場合の経路選択問題。

② **Naïve User (NU) 問題**：利用した経路の情報しか知らない場合の経路選択問題。利用されなかった経路に対しエージェントは推測行動をとる。言い換えれば、NU問題は利用しなかった経路の推定値に基づく IU 問題ということができ、推定が漸近的に最良推定値に近づくよう改善できれば、収束した解は IU 問題で得られる解を十分近似することになる。

上記の利用者を想定し、利用者の交通行動を繰り返しゲームとして定式化する。また、室内実験法を設計し、動的データを収集する。得られたデータを用い、理論から誘導される経路選択モデルの構造パラメータを推定する。なお、構造パラメータ推定とは、システムの均衡とモデルパラメータ推定を同定に達成する手法を指す。

### (2) 経路選択行動の実験

被験者に対し、利用した経路の所要時間のみを提供する場合を NU 実験 (実験 I)、利用しなかった経路の上についても提供する場合を IU 実験 (実験 II) と呼び、それぞれ

のケースについて被験者18人を対象に150回の繰り返し実験を行なう。また、実験は2経路のケースと3経路のケースを行う。いずれの場合も被験者には道路の特徴を表す図の情報のみが提供され、走行時間関数や各経路の利用数は知らず、実験者から提供される所要時間情報のみで経路選択を繰り返す。被験者はコンピュータ画面で経路情報を得る事ができ、また、経路情報は、あらかじめ想定したリンク所要時間関数に基づき、利得に変換した値を提供する。被験者は、獲得した利得を実験後、実際に受け取ることができる。

(3) 繰り返しゲームと強化学習理論の統合  
本研究で提案する経路選択モデルは、ゲーム理論と強化学習理論を統合したものであり、従来の経路選択理論と比較して以下の特徴をもつ。

- ① 個々の利用者を独立した意思決定者と見なし、ネットワークフローを粒子の集団として扱う粒子モデルである。
- ② 完全情報を仮定しておらず、より現実的な交通行動モデルである。
- ③ Day-to-day のダイナミックな行動を想定している。
- ④ 粒子 SUE モデルである。しかし、NU の SUE モデルはこれまで提案されておらず、その特徴は分かっていない。
- ⑤ 確率的仮想プレイを基本に、確率近似理論および強化学習理論を統合したモデルである。
- ⑥ 動的データを前提に、システムの均衡と行動選択モデルのパラメータを同時決定する手法の提案。

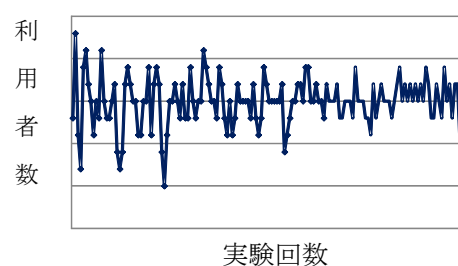
#### 4. 研究成果

##### (1) 室内実験

150回の繰り返し実験の結果について経路2の所要時間の変動を示したものを図

1に示す。図1はIUの2経路のケースを横軸に繰り返し回数、縦軸に経路2の選択者数で表示したものである。所要時間は理論値の30分周辺で変動し、収束しない。このことは、フローのほうも理論値の周辺で変動する事を表している。この状況は3経路のケース、また、NU ケースにも同様に観測された。しかし、それらのケースでは分散の幅は増大し、実験回数の増加に伴う分散の減少は、図1のケースに比較して小さかった。

図1 経路選択の変動（室内実験）



主な結論は、

- ① 平均値の周辺で変動するという結果は Selten らの室内実験結果と同じである。
- ② 繰り返し選択に伴う経路情報の提供方法は Selten らの累積和に加え、リグレット情報についても新たに加え実施したが、収束の傾向に大きな変更は見られなかった。ただし、効果は累積和とリグレットでは異なっており、リグレット情報のほうがより合理的経路選択を促すことを統計的に明らかにした。

##### (2) アルゴリズムと検証

本研究で開発されたアルゴリズムの特徴は以下のようなものである。

##### ① Informed User(IU)アルゴリズム

ドライバーは、自己の利得関数（あるいはコスト関数）を知っているが、他者の利得関数は知らない。ただし、前日の行動結果を利用する事ができる。この仮定は、既存の連続型 SUE と同じであるが、個人の利得関数は

異なってもよいという意味で、拡張されたモデルになっている。アルゴリズムは、基本的にクールノ（あるいはバートランダー）調整過程と同じであり、ランダムに抽出された利用者は、他者の行動が昨日と同じであるという仮定の下で最善行動をとる。アルゴリズムの詳細は割愛するが、結果は以下の補題に要約できる。

補題1：提案したIUアルゴリズムの下では、システムは確率1でNash均衡に収束する。

### ②Naïve User(NU)アルゴリズム

ドライバーは自己の利得関数を知らず、他者の利得関数も知らない。利用者が利用できる情報は、自己が経験した経路の実現利得のみである。この場合、利用者は利用しなかった経路の情報を推測する。推計値が実際に経験する実現値に漸近する場合、システムは安定する。しかし、経験値と推定値が一致しても、それがNash均衡を保証しているとは限らない。結果は以下の補題に要約できる。

補題2：提案したNUアルゴリズムの下では、システムは確率1でNash均衡に収束することはない。

補題2は提案アルゴリズムの欠点を示すものではない。実際、不完全記憶のもとでは確率1でNash均衡に収束することはないことが証明されている。本研究で提案されたアルゴリズムも、それを裏付けている。実際、アルゴリズムはかなり高い確率でNash均衡に収束するが、補題2は混雑ゲームでも、そのことが常に保証されるものではないことを示している。

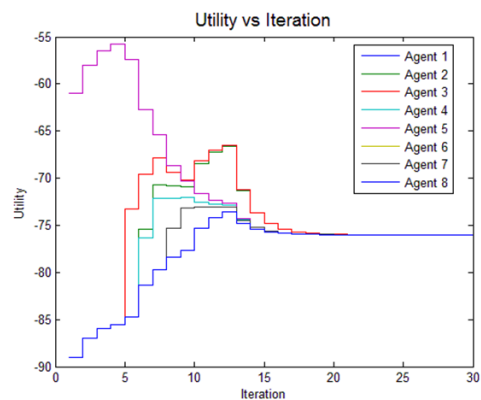
### (3) 数値実験

数値実験は、Braess ネットワーク (4 ノード、5 リンク) および Sioux-Falls ネットワーク (24 ノード、76 リンク) で検証した。また、Braess ネットワークの例では、リンクコスト関数が非対称ヤコビ行列をもつ場合、

階段関数など、従来のUE、SUEアルゴリズムでは扱えなかった関数を使った検証も行っている。以下の結果を得た。

- ① いずれの場合もIUアルゴリズムは確率1でNash均衡に収束するが、NUについては、高々90%程度の収束率であった。
- ② 収束効率性は非粒子モデルと比較して高い。図2は、Braess ネットワークで線形コスト関数の場合を示している。非粒子モデルが均衡値の周辺で変動し、その変動幅が小さくなるのに多くの反復回数を必要とするのと対照的である。
- ③ 大きなネットワークを対象にする場合には、利用可能経路をどのように限定するかが問題になる。本研究ではK経路法を利用した経路抽出を行った。

図2 アルゴリズムの収束性



### (4) 経路選択関数の構造パラメータ推定

室内実験で得られた動的データを用いて、ロジットモデルのパラメータ推定と均衡計算を同時に行う。本研究では、最適応答を元に逐次、擬似尤度関数法を改良していくNPL法を採用した。パラメータの推定と均衡計算が同時進行するので、均衡した結果とパラメータ推定の整合性がとれるという利点をもつ。なお、仮定した効用関数は次のようである。

経路1の効用 =  $-\alpha$  (経路1の所要時間)

経路 2 の効用 =  $-\beta$

表 1 は、静学モデルと同額モデルの結果を比較したものである。動学モデルとは、過去の走行経験に加え、先読みをして経路選択を行うケースを指す。いずれの場合も理論的に計算される均衡値（ルート 1 (R1) の利用者数 = 12、ルート 2 (R2) の利用者数 = 6）に収斂するが、パラメータの  $t$  値に大きな差異があり、先読みモデルの係数が統計的には安定していることが示された。

表 1 構造パラメータの推定結果

パラメータ	静学モデル		動学モデル	
	$\alpha$	$\beta$	$\alpha$	$\beta$
値	-1.103	-1.224	-0.965	-1.196
T 値	-7.656	-13.61	-8.203	-14.20
尤度比	0.4181		0.4184	
R 1	11.859		12.031	
R 2	6.141		5.970	

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 7 件)

- ① Miyagi, T., G.C. Peque, Jr and J. Fukumoto: Adaptive Learning Algorithms for Traffic Games with Naive Users, the 20th International Symposium on Traffic Theory and Transportation, 2013.(in print)
- ② 池田 愛, 宮城俊彦: 経路選択行動に関する室内実験, 交通工学, Vol. 48, No.2, pp.53-62, 2013.  
(<http://www.jste.or.jp/Books/kikan-cont-4802.html>)
- ③ Kato, H. and T. Miyagi: Estimation of Structural Parameters of the Economic Growth Model with Subspace System Identification, *Journal of China-US Business*

*Review*, Vol.11, N0.7, pp.877-855, 2012. (ISSN 1537-1514)

- ④ Miyagi, T. and G.C. Peque, Jr. : Informed-user algorithms that converges to Nash equilibrium in traffic games, *Procedia-Social and Behavioral Sciences* 54, Elsevier, pp.438-449, 2012. (ISSN 1877-0428)
- ⑤ 加藤裕人, 宮城俊彦: DSGE モデルによる公共投資の効果分析とモデルの時変パラメータ推定, 土木学会論文集 D3, Vol. 68, N0.5, pp. L121-L130, 2012. (<https://www.jsce.or.jp/committee/lp/monograph/file/ipv29-2012.pdf>)
- ⑥ 宮城俊彦・遠藤雅人: 適応的経路選択モデルにおける経路集合の限定手法と経路分散パラメータの推定法, 土木学会論文集 D3, Vol.67, No.5, pp. L541-L552, 2011. (<https://www.jsce.or.jp/committee/lp/monograph/file/ipv28-2011.pdf>)
- ⑦ Miyagi, T. : An adaptive learning algorithm for a route choice problem in uncertain traffic environments, *Urban Transport*, Vol.17, pp. 43-52, 2011. (ISBN 978-1-84564-520-5)

[学会発表] (計 9 件)

- ① Miyagi, T., G.C. Peque, Jr and J. Fukumoto: Adaptive Learning Algorithms for Traffic Games with Naive Users, the 20th International Symposium on Traffic Theory and Transportation, Noordwijk, the Netherlands, July 17, 2013.
- ② Miyagi, T., and G.C. Peque, Jr. : Informed-user algorithms that converges to Nash equilibrium in traffic games, the 15th meeting of the Euro Working Group on Transportation, Sept. 11, 2012, Cite

Descartes, France.

- ③ 宮城俊彦：多数エージェントの動学的離散的選択行動モデルに対する定常均衡アプローチについて，日本地域学会学術発表論文集 49，日本地域学会、2012 年 10 月 7 日、立正大学。
- ④ Miyagi, T., and G.C. Peque, Jr. :Traffic Games with Incomplete Travel Information, 土木計画学研究講演集、Vol. 45, 土木学会、2012 年 6 月 2 日、京都大学
- ⑤ 池田 愛・宮城俊彦，経路選択行動モデルの構造パラメータ推定について，土木計画学研究・講演集，Vo. 46， 2012 年 11 月 3 日、埼玉大学
- ⑥ 張 洋・G.C. Peque, Jr. ・宮城俊彦：混雑ゲームにおける Non-atomic モデルの数値計算特性について，土木計画学研究・講演集、Vo. 46， 土木学会、2012 年 11 月 3 日、埼玉大学
- ⑦ 池田 愛・宮城俊彦，動的経路選択行動の室内実験による検証と分析，土木計画学研究・講演集，Vol. 44， 土木学会，2011 年 11 月 27 日、岐阜大学。
- ⑧ 加藤裕人・宮城俊彦，動学的確率的一般均衡モデルの時変パラメータ推定による財政政策の計効果分析，土木計画学研究・講演集，土木学会、Vo. 44， 2011 年 11 月 27 日、岐阜大学。
- ⑨ Kato, H. and Miyagi,T.: Estimation of Structural Parameters of the Economic Growth Model with Subspace System Identification, the 22<sup>nd</sup> Pacific Regional Science Conference, RSAI, July 5, Seoul KyoYuk MunHwa HoeKwan, Seoul, Korea 2011.

## 6. 研究組織

### (1) 研究代表者

宮城 俊彦 (MIYAGI TOSHIHIKO)

東北大学・大学院情報科学研究科・教授

研究者番号：20092968

(2) 研究分担者

福本潤也 (FUKUMOTO JYUNYA)

東北大学・大学院情報科学研究科・准教授

研究者番号：3032344