

科学研究費助成事業 研究成果報告書

令和 6 年 6 月 13 日現在

機関番号：13701

研究種目：基盤研究(C)（一般）

研究期間：2020～2023

課題番号：20K04719

研究課題名（和文）深層強化学習を内包した漸近的最適応答モデルと交通流シミュレーション分析への応用

研究課題名（英文）Basic framework of the asymptotic best-response model with deep-reinforcement learning in the traffic simulation applications

研究代表者

宮城 俊彦（Miyagi, Toshihiko）

岐阜大学・工学部・特任教授

研究者番号：20092968

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：交通システムの分析と最適化に関連する問題を解決するために漸近的最適応答アルゴリズムの開発と拡張およびその基礎理論を提案した。この成果は以下の問題に対処するために利用できる。1) 自動走行システムや分権型ネットワーク社会に適した交通システムの開発、2) 交通システムの効率化を図るための情報通信技術の発展を評価するための交通システム分析手法、3) 離散化された利用者が利用可能な経路の中から所要時間を最小化するような経路選択の行動モデルと学習実験の効果予測、3) 不完全情報のもとでの個人および個人と環境の相互作用を考慮したゲーム論的学習行動モデルの基礎理論の構築

研究成果の学術的意義や社会的意義

ABRアルゴリズムは、利用者が自律的に走行環境を学習し、漸近的に自己の効用を改善するように行動すると仮定した手法である。不確実な交通状況に応じた利用者の限界合理的な行動に基づく最適な経路選択の学習行動をモデル化している。また、ゲーム理論的なアプローチを取り入れることで、より複雑な状況下での利用者の行動を予測し、分析することが可能である。さらに、シミュレーションを通じて実際の交通流の動きを再現し、それに基づいた最適な交通量配分を行うことができる。これらの特徴により、ABRアルゴリズムは従来の交通ネットワーク均衡モデル、非集計行動モデルの拡張であると同時にゲーム論的行動モデルの拡張になっている。

研究成果の概要（英文）：The development of an asymptotically best response algorithm and its basic theory were proposed to address the following issues: 1) development of transportation systems suitable for automated driving systems and decentralized network societies, 2) transportation system analysis to evaluate the development of information and communication technologies to improve the efficiency of transportation systems, 3) the discretized behavioral model for route choice that minimize travel time among available routes for users and prediction of the effects of learning experiments, and 4) Development of a basic theory of game-theoretic learning behavior models that take into account interactions among individuals and between individuals and the environment under imperfect information.

研究分野：国土計画・交通工学

キーワード：離散的動的経路選択モデル 漸近的最適応答モデル Nash均衡 ゲーム論的行動選択モデル ミクロ交通流シミュレーションベース交通配分 心理学的学習行動モデル 深層強化学習 限界合理性

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

様式 C - 19、F - 19 - 1 (共通)

1. 研究開発当初の背景

高度情報化社会における CASE (Connected, Autonomous, Shared & Services, Electric)、MaaS (Mobility as a Service)、IoT (Internet of Things) といったモビリティ革命に対応した交通システムの分析には、従来のように交通利用者をマスとして捉える手法では限界があり、個別意思決定主体として捉える新しい分析手法が必要である。従来の非集計交通需要モデルは利用者間の相互作用の影響は取り扱われておらず、また、交通量配分の段階で集計化されるため交通情報が個々の利用者にどのような行動変容をもたらすかを分析できない。本研究では、トリップ主体を個別意思決定者として扱うゲーム論的粒子モデルとその拡張を提案しており、その基本フレームは非集計交通需要モデル、利用者均衡モデルの拡張である同時にゲーム論の混雑ゲームの拡張でもある。

2. 研究の目的

(1) 自発的動機付けによる行動変容モデル：研究の焦点はゲーム論的粒子モデルに交通均衡は存在するのかという理論的解析と交通均衡を求めるための計算手法の開発にある。その結果、漸近的再適応等 (ABR) アルゴリズムが提案されるが、ABR モデルをより広範囲に利用するための拡張モデルについても検討する。

(2) 相利共生的行動変容モデル：交通行動の変容分析においては、外部的なコントロールに基づく行動変容の有意性を検証するモデルも必要になる。このためには単に報酬に伴う行動分析のみならず、マイナスの利益が行動形態に与える影響を取り入れる必要がある。本研究では、認知科学の分野で研究されている社会的ジレンマを解決する学習行動モデルおよびゲーミフィケーション・アプローチの開発を目指す。

3. 研究の方法

(1) 自発的動機付けによる行動変容モデル：利用者は経路情報を全く持たず、自己の利得関数も知らない無知の状態にある (naïve user) であるが、日々の経路選択行動を通して経路情報を得て行動学習すると仮定する。この行動モデルの振舞いは、一般化弱仮想プレイとして定式化でき次のよう表される。

● 実現利得

$$\text{Online: } U_t^i(a^i), a^i \in A^i, i \in I$$

$$\text{Offline: } U_t^i, i \in I$$

交通情報は V2x コミュニケーションシステムも対象

$$U_t^i = u^i(a_t^i, a_t^{-i}) + e^i(a_t^i)$$

利得は未知の確率分布に従う誤差を含む

- 学習に伴う利得の更新

$$Q_{t+1}^i(a^i) = Q_t^i(a^i) + \alpha_{t+1} I\{a_t^i = a^i\} [U_t^i - Q_t^i(a^i)], a^i \in A^i$$

非同期モデル

- 経路選択確率の更新

$$\pi_{t+1}^i(a^i) = \pi_t^i(a^i) + \gamma_{t+1} [BR_t^i(a^i) - \pi_t^i(a^i)], a^i \in A^i$$

最善案を漸的に更新

$$BR_t^i(a^i) = \frac{\exp\{q_t^i(a^i) / \mu_t^i\}}{\sum_{a^i \in A^i} \exp\{q_t^i(a^i) / \mu_t^i\}}$$

最善の選択肢を選択する確率

一般化仮想プレイは、再帰的、連鎖的 Nash 均衡に収束することが分かっている。しかし、交通流の確率の変動が小さければ利得の変動はリップシツ連続になり、一意的な Nash 均衡に収束する。上式で表される経路選択行動を表現するモデルを漸近的最適応答 (ABR) モデルと呼び、上記のような特性を数値計算で確認する。

(2) 深層強化学習の導入：漸近的最適応答モデルと深層強化学習を組み合わせることで、ドライバーのより複雑な挙動をモデル化する。この試みにより、追い越し行動などのより詳細な交通行動をシミュレーションモデルに取り入れることが可能になる。利得の推定には LSTM モデルなどを用いる。

(3) 相利共生的な経路選択モデルの構築：ABR モデルは交通システムの利用に伴う利得 (報酬) のみを対象に、Nash 均衡を実現する手法である。道路空間を共有する利用者の挙動は互いの利益が相反する場合もある。このため、社会システムのコンフリクト緩和する心理学的ゲーム論アプローチが必要である。また、ゲーミフィケーションを用いた道路管理システムは外的な動機付けに基づく行動変容モデルの一種であり、道路利用者の走行データを道路管理者が収集し、それに基づく混雑緩和のガイダンスを与える場合、それがどの程度行動変容を促すかを実験的に確認する。

4. 研究成果

(1) ABR アルゴリズムの収束性：ABR は粒子モデルであり、利用者は自己の利得関数を知らないことを前提にしている。この仮定を表現できる交通流モデルは、セルラーオートマトン (CA) を用いた Naschr モデルである。図 1 は、CA モデルでランダム変動が小さい場合、図 2 は確率変動が大きい場合の ABR モデルの収束を単体上で表したものである。確率変動が小さい場合、Nash 均衡に一意的に収束する。しかし、確率変動を大きくすると、図 3 のよう経路利得は平均値の周辺でボラタイルに変動する。その結果、均衡はある一定のサイズの集合に収束するが、一定の場所に留まることはない。これらの結果は理論そして室内実験の結果に一致する。ABR モデルを実用規模のネットワークにまで拡張する試みは、計算に膨大な時間を要するため実行できず大きな課題として残された。

(2) 路側センサー・データを用いた状態変化予測：オキュパンシーの観測値を用いてその変化

を予測する LSTM モデルを作成し、実際のデータを利用して検証行った。先行データを教師データとし、30分先を予測する試みではかなり良い結果を得た。このモデルと ABR モデルを結合することを研究の目的の一つとしていたが、ABR モデルを実用規模のネットワークにまで拡張するには至らずこの試みは完結していない。

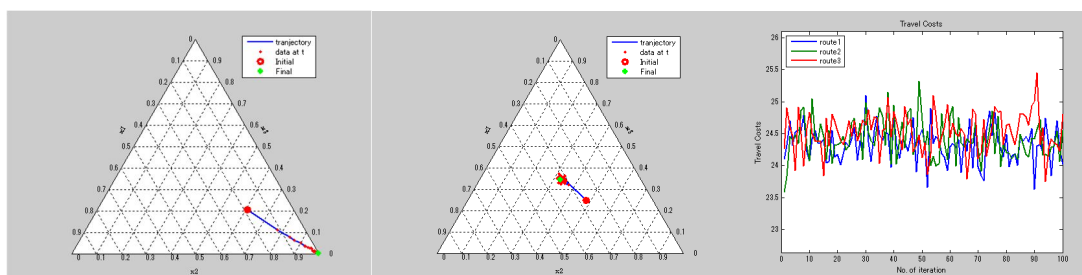


図1 隔離変動が小

図2 確率変動が大

図3 走行時間の変動

(3) 道路空間でのコンフリクトを緩和するモデル：この目的のため、利用に伴う〈報酬：R〉に加え、追い越しなどにより高い利益を得ようとする〈誘惑：T〉そして交通ルールを逸脱することによる〈罰則：P〉、他者に迷惑を与える〈迷惑：S〉などの要素を導入し、これらの相互関係から決定される選択性向を用いて満足度を高めるような行動変容モデルを考え検討した。モデルの概要は図4のようである。

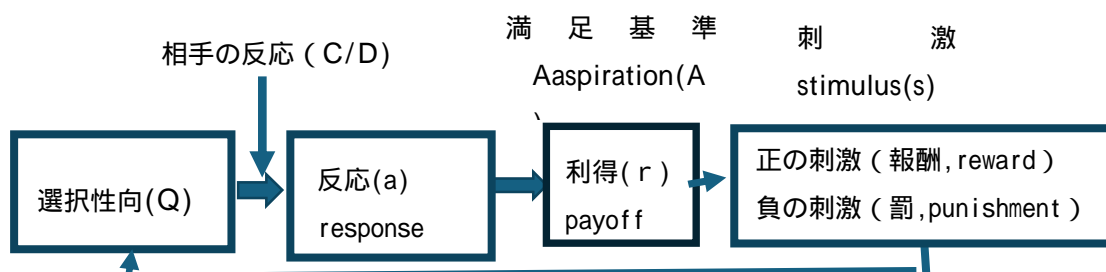


図4 コンフリクトモデルの概要

$$s(a) = \frac{r(a) - A}{\sup[|T - A|, |R - A|, |P - A|, |S - A|]}$$

強化関数

$$Q_{t+1}(a) = \beta Q_t(a) + s_{t+1}(a) \rightarrow \begin{cases} s_{t+1}(a) = \beta s_t(a) + \omega, & \text{if } a_t = a \\ s_{t+1}(a) = s_t(a), & \text{otherwise} \end{cases}$$

選択性向 (評価値)

$$p_{t+1}(a) = \frac{Q_t(a)}{\sum_{a \in A} Q_t(a)}, A = [C, D]$$

行動選択確率

このゲームは、仮定された利得ベクトルの下では収束したが、満足基準、誘惑、迷惑、罰則などのパラメータをどう決定するかは課題を残す。ゲーミフィケーションによるアプローチを広島での社会実験に適用しその有効性を確認したが、この結果をコンフリクトモデルのパラメータ推定に利用する方法は実現できていない。

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 4件/うち国際共著 0件/うちオープンアクセス 2件）

1. 著者名 荒木咲良、倉内文孝	4. 巻 97
2. 論文標題 高速道路における行動変更提案に向けた交通状況フォーキャストモデルの構築	5. 発行年 2023年
3. 雑誌名 第43回交通工学研究発表会論文集	6. 最初と最後の頁 659,666
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Kaewklueklom, R, Kurauchi, F. and Iwamoto, T.	4. 巻 14
2. 論文標題 Applying the Hyperpath Concept to Public Transit Accessibility Evaluation	5. 発行年 2022年
3. 雑誌名 Journal of EASTS	6. 最初と最後の頁 374,389
掲載論文のDOI（デジタルオブジェクト識別子） 10.11175/easts.14.374	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Kaewklueklom R., Kurauchi, F and Iwamoto T.	4. 巻 -
2. 論文標題 Discrepancy of route choice between theoretical hyperpath and actual choice based on smart card data in Shizuoka, Japan	5. 発行年 2022年
3. 雑誌名 Proceedings of the 6th Conference on Sustainable Urban Mobility	6. 最初と最後の頁 251,265
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Kaewklueklom, R., Kurauchi, F. and Iwamoto, T.	4. 巻 19
2. 論文標題 Investigation of Changes in Passenger Behavior Using Longitudinal Smart Card Data	5. 発行年 2021年
3. 雑誌名 International Journal of Intelligent Transportation Systems Research	6. 最初と最後の頁 155, 166
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s13177-020-00232-3	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計2件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 柴垣太郎, 倉内文孝, 中村俊之, 東義朗
2. 発表標題 高速道路における渋滞緩和をめざしたゲーミフィケーションデザインとその要素評価
3. 学会等名 令和5年度土木学会中部支部研究発表会
4. 発表年 2024年

1. 発表者名 杜然, 倉内文孝
2. 発表標題 深層学習による高速道路の渋滞ナウキャストに関する研究
3. 学会等名 令和2年度土木学会中部支部研究発表会
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

本研究に関連する内容は、研究討論会・セミナー等でも報告している。

- 宮城義彦: "ミクロ交通流シミュレーションモデルを基礎とした交通均衡モデルの理論枠組み", 北大・熊本大・岐阜大合同セミナー、北海道大学、3月19、20日、2022年。
- 宮城俊彦: "交通ネットワーク均衡を求めるためのゲーム論的学習アルゴリズム", 高知工科大・北大・熊本大・岐阜大合同セミナー、高知工科大学、9月13-15日、2022年。
- 宮城俊彦: "動的確率的行動選択モデルの構造比較分析", 科学研究費研究：三大学合同研究討論会、琉球大学、11月13日、2023年。
- 宮城俊彦: "強化学習モデルに関する話題提供", 北大・熊本大・岐阜大合同セミナー、熊本大学、3月28、29日、2023年。

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分担 者	倉内 文孝 (Kurauchi Fumitaka)	岐阜大学・工学部・教授	
	(10263104)	(13701)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------