

# 強化学習を用いた 高速道路における制限速度制御手法の構築

藤本 想<sup>1</sup>・塩見 康博<sup>2</sup>・花房 比佐友<sup>3</sup>

<sup>1</sup> 学生員 立命館大学大学院 理工学研究科 (〒525-8577 滋賀県草津市野路東 1-1-1)

E-mail:ce0011fe@ed.ritsumei.ac.jp

<sup>2</sup> 正会員 立命館大学 理工学部環境都市工学科 (〒525-8577 滋賀県草津市野路東 1-1-1)

E-mail:shiomi@fc.ritsumei.ac.jp

<sup>3</sup> 正会員 株式会社アイ・トランスポート・ラボ (〒101-0052 東京都千代田区神田小川町 3-10)

E-mail:hanabusa@i-transportlab.jp

高速道路における交通渋滞への対策として、動的交通流管理 (ATM) 及び ATM 内の可変制限速度制御 (VSL) の運用が進められている。しかし、現行の可変制限速度制御は必ずしも最適とはいえ、現実の交通流への対応度合いも高いものではない。そこで本研究では、高速道路での渋滞発生抑制及び旅行時間の短縮の実現に向けた、高速道路上における制限速度制御手法を提案する。提案する速度制御のシステムは、強化学習モデルと交通流シミュレーションを用いた交通状況予測モデルで構築されている。強化学習を適用することで、多様な交通状態での最適な制限速度値を学習させ、道路区間ごとに動的に最適な制限速度を算出する。構築した速度制御システムの妥当性を検証した結果、制御前と制御後で平均旅行時間が減少するなど交通状況の改善が確認された。

**Key Words:** Variable Speed Limit, Microscopic Simulation, Reinforcement Learning, Freeway

## 1. はじめに

高速道路における慢性的な交通渋滞は依然として重要な社会問題であり、移動の効率性を保証するためのマネジメント手法の開発が求められている。交通集中による渋滞対策として有効とされている手法が、動的交通流管理 (Active Traffic Management, ATM) である。中でも可変制限速度 (Variable Speed Limit, VSL) は、旅行時間や通過交通量の改善は難しいものの、事故率の低減やショックウェーブの減少、車線利用率の平準化の効果が報告されている<sup>1)</sup>。また、マクロな視点より Cell Transmission Model を用いた強化学習による可変制限速度制御によって、高速道路における移動時間が安定した需要シナリオにて 49.34%、変動する需要シナリオにて 21.84%短縮されたという研究結果も発表されている<sup>2)</sup>。

可変制限速度のアルゴリズムは様々なものが提案されているが、実際に運用されているものは必ずしも最適である保証はなく、理論に基づくアルゴリズムも、現実の交通流に完全に対応したものとは言い切れない。そこで、車両の追従挙動を考慮したマイクロ交通流シミュレーションを強化学習として用いることに本研究における新規性

が存在すると考え、強化学習によって渋滞リスク及び旅行時間を抑えられる可変制限速度制御アルゴリズムの検討に対して重要な意義をもつと考えた。

そこで本研究では、強化学習を用い、高速道路での渋滞リスクの抑制及び旅行時間を抑えられる制限速度設定手法を提案する。具体的には、まず、マイクロ交通流シミュレーションの 1 つである MicroAVENUE を用いた強化学習モデルを構築する。それに基づき、強化学習による制限速度設定の妥当性を検証する。

## 2. 本研究でのシナリオ設定

### (1) MicroAVENUE の概要

MicroAVENUE<sup>3)</sup>は株式会社アイ・トランスポート・ラボによって開発された高速道路及び公共交通に対応可能な追従型の交通シミュレーションソフトであり、縦断勾配による速度低下や渋滞のショックウェーブの状況などを表現することが可能である。追従走行には非線形追従モデルの一種である IDM+ (Intelligent Driver Model+) が用いられており、0.1 秒毎に現在の追従状態値から、次の

時刻での加速度を求められる。前方車両との速度差や安全車間距離、希望速度、最大加減速度などから加速度を算出する。その際、個別車両ごとにパラメータを分布させることで、挙動のばらつきを考慮できる。MicroAVENUE の入力データにおいて制限速度の設定を行い、車両が制限速度を超過しないように調整した。

(2) 設定した路線の全体図

本研究での学習環境として、渋滞の原因となるボトルネック（合流部・サグ部）の含まれている片側 1 車線の高速道路を MicroAVENUE を用いて表現した。その全体図を図-1 に示す。

(3) 制限速度

MicroAVENUE でシミュレーション実行するにあたって、本研究では、片側 1 車線の直線的な高速道路で可変速度標識を 5km/h 単位で表示できるように調整し、500m 程度の間隔で設置することを想定している。その際、図-2 に示すように隣接する区間の制限速度差は 5km/h を超えないとする制約を加えた。これは、特に先の区間で制限速度が大幅に低下した場合に車両が急減速してショックウェーブが発生することを抑制することを目的としている。

そして、合流する道路の制限速度は 40km/h、本線の道路の強化学習前の制限速度は 80km/h とする。これは、気象条件や工事等の原因によって最も設定されることのできる速度値であると考えたことが理由である。

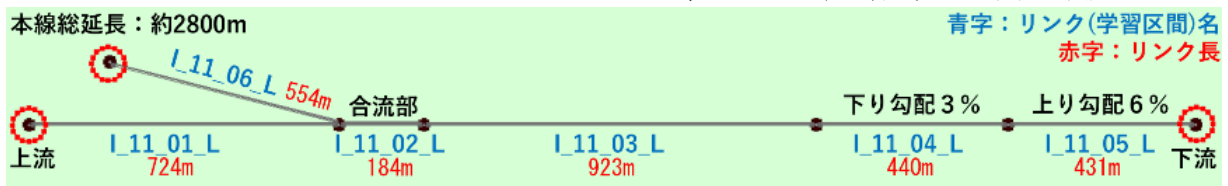


図-1 MicroAVENUE で作成した片側 1 車線の高速道路

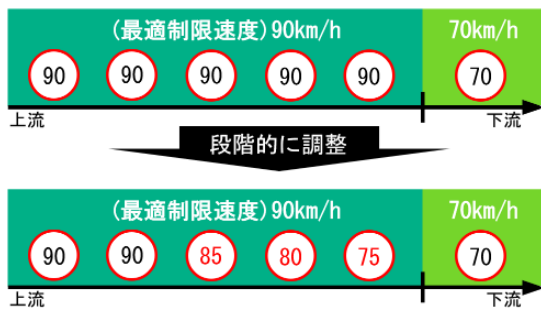


図-2 同一学習区間内の標識の速度調整

表-1 区間交通量の設定

交通量設定	本線 (台/30分)			合流 (台/30分)		
	小型	大型	バス	小型	大型	バス
A	300	50	50	60	10	10
B	300	50	50	120	20	20

(4) 実行時間と区間交通量

本研究の強化学習を効率的に遂行するために、シミュレーションの予測時間は 30 分としている。これは、シミュレーション開始時の高速道路上に車両が 1 台も存在しない状態から、渋滞状況が再現されるまでに最低限必要な時間が 30 分程度であると考えたことに由来する。

区間交通量の設定に関して、シミュレーション実行時にサグ部で渋滞現象が確認できた交通量設定 A と、合流部・サグ部の両エリアにて渋滞現象が確認できた交通量設定 B の 2 種類を強化学習で用いる交通量設定とした。30 分間における交通量設定 A・B の内訳を表-1 に示す。

3. 強化学習モデルの構築

(1) 強化学習の概要

渋滞を抑制するという目標の下、設定した交通量・配分に対する最適な速度制限手法を、MicroAVENUE を用いた強化学習によって導出するプログラムを構築する。本研究での VSL 強化学習の概略図を図-3 に示す。

MicroAVENUE 内の道路状況・交通状況の下で、設定された区間交通量と制限速度に対して得られた平均旅行時間と PICUD 値の結果を報酬化することにより、最適な制限速度調整手法を探索することが可能となる。図では「報酬」という言葉を用いているが、安全性が確保された上での平均旅行時間の短縮を学習目標としているので、数値の増減を揃えるように負の報酬（ペナルティ）とし、ペナルティの最小化を目指すよう設定している。

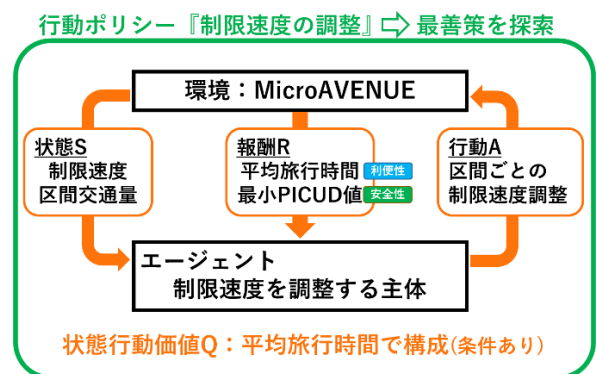


図-3 VSL 強化学習の概略図

本研究においては、交通条件（高速道路全体の区間交通量・各区間の制限速度）に応じて MicroAVENUE が再現度の高い交通流を示すことを利用し、総ペナルティの最小化を通じた最善な制限速度調整手法の探索を行った。

(2) 報酬の定義

本研究の強化学習では、評価指標の一つに PICUD 値を使用する。区間交通量が多い場合にむやみに制限速度を上げることによる衝突事故リスクを抑える意図がある。鈴木ら<sup>4)</sup>の研究内容をもとに本研究で用いる PICUD 値の設定を行う。

PICUD (Possibility Index for Collision with Urgent Deceleration) 指標は前方車両が急減速した場合、後方車両が反応遅れを伴って急減速をし、完全に停止した状況での車間距離を表したものである。したがって、PICUD 指標の値が 0 以下である時は衝突を意味し、急減速が発生した場合、両車両が衝突する潜在的危険性が高いことを示している。PICUD 値の算出方法は式(1a)で表される。

$$PICUD = \frac{V_2^2}{-2a} + s_0 + \left( V_1 \Delta t + \frac{V_1^2}{-2a} \right) \quad (1a)$$

- $V_1$  : 減速開始時の後続車両速度 [m/s] ,
- $V_2$  : 減速開始時の先行車両速度 [m/s] ,
- $s_0$  : 車間距離 [m] ,  $\Delta t$  : 反応遅れ時間 [s] ,
- $a$  : 減速時の加速度 [m/s<sup>2</sup>]

ここで、反応遅れ時間  $\Delta t$  は 1.0 秒と設定する。これは、制動停止時の空走距離算定において、反応動作に要するブレーキ反応時間が 1.0 秒と与えられるためである。また、減速時の加速度  $a$  は、0.3G ( $G=9.8\text{m/s}^2$ ) を用いる。これは、0.3G がヒヤリハットと呼ばれる現象を表す場合に用いる減速時の加速度を示しており、通常の自動車のブレーキよりも強めの減速が生じることに由来する。

MicroAVENUE では当該車両とその前後の車両それぞれに対する PICUD 値を拡張機能として算出することが可能である。ただ、当該車両の前方と前方車両の後方のデータが重複するため、当該車両とその後方追従車両との PICUD 値のみを用いた。

また、もう一つの評価指標として平均旅行時間を使用する。平均旅行時間の短縮を通して高速道路の利便性を向上させる意図がある。平均旅行時間は MicroAVENUE の学習区間全体（全長約 2800m）における数値であり、単位は秒である。

(3) ペナルティと学習目標

ペナルティ（負の報酬）に関して、PICUD 値においては各学習区間で最小値を用いるが、この値が 0 以上であればペナルティを与えない。0 未満となれば衝突危険性があるため、ペナルティとして +99999 を与える。平均旅

行時間においては高速道路全体の秒数を直接ペナルティの値とする。

本研究では、最小 PICUD 値によるペナルティ（5 区間）と高速道路全体の平均旅行時間によるペナルティの 6 項目を合計したものを総ペナルティとし、これが最小値となるように行動ポリシーを更新していくことで、安全性を確保した上で利便性が最大となる制限速度調整手法を探索し決定することが可能となる。本研究の強化学習における制限速度の調整回数は 500 回で設定し、最善な制限速度調整手法の探索を行った。

4. シミュレーション結果

(1) 交通量設定 A

交通量設定 A においては、最適制御によって、図-4 に示すように、合流部での制限速度を引き下げ、合流部の前後区間での制限速度を引き上げる結果となった。

図-5 に最適制御前、図-6 に最適制御後の交通状態を示す。タイムスペース図内、赤の実線部が合流部を指し、赤の破線部がサグ部を指している。各区間における制限速度調整によって、サグ部を起点とするショックウェーブが抑制されたことに加え、合流部を起点とするショックウェーブの抑制の効果も見られた。

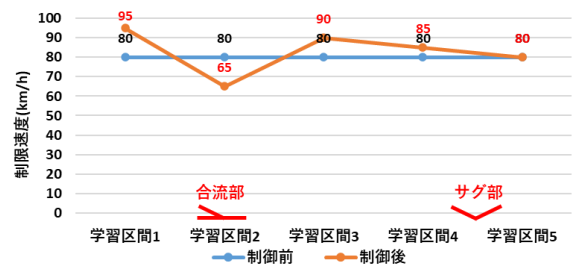


図-4 最適制御前後の制限速度の変化（交通量設定 A）

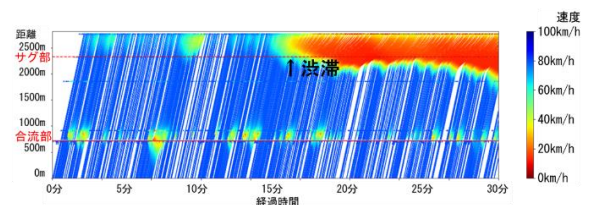


図-5 最適制御前の交通状態（交通量設定 A）

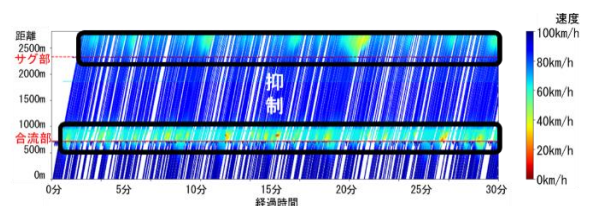


図-6 最適制御後の交通状態（交通量設定 A）

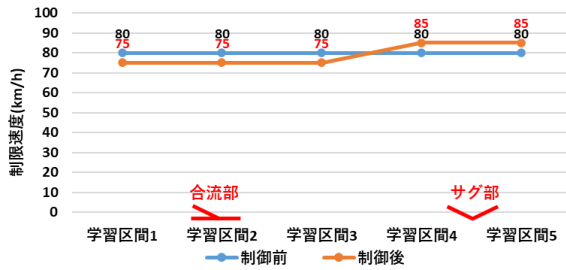


図-7 最適制御前後の制限速度の変化 (交通量設定 B)

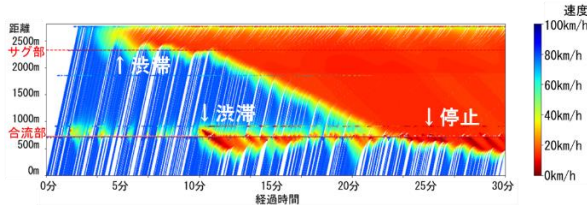


図-8 最適制御前の交通状態 (交通量設定 B)

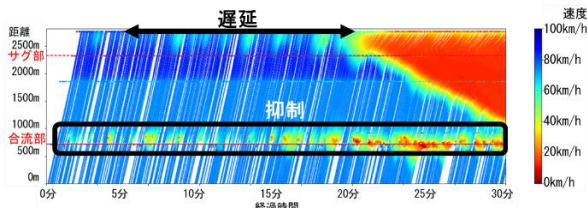


図-9 最適制御後の交通状態 (交通量設定 B)

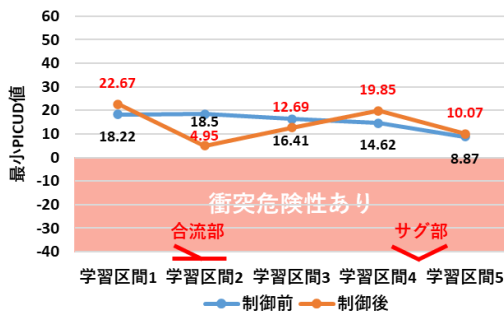


図-10 最適制御前後の PICUD 値の変化 (交通量設定 A)

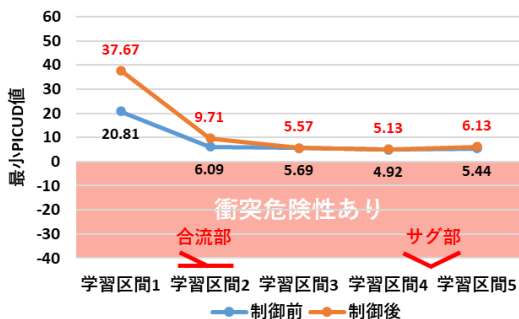


図-11 最適制御前後の PICUD 値の変化 (交通量設定 B)

## (2) 交通量設定 B

交通量設定 B においては、最適制御によって、図-7に示すように、サグ部に至るまでの制限速度を抑え、サグ部での制限速度を上げる結果となった。

図-8に最適制御前、図-9に最適制御後の交通状態を示

す。各区间における制限速度調整によって、サグ部でのショックウェーブの発生を遅らせることに成功しており、合流部でのショックウェーブの抑制の効果も見られた。

## (3) 最小 PICUD 値の変化

図-10に交通量設定 A、図-11に交通量設定 B の最適制御による最小 PICUD 値の変化を示しているが、最小 PICUD 値が 0 を下回ることによるペナルティを設定することで、最適制御によって衝突危険性のある制限速度の組み合わせが棄却されることとなった。これは、安全性を確保した上で利便性を向上させる制限速度制御が行われていることを証明している。

## 5. おわりに

### (1) 本研究の成果

本研究では、強化学習を用い、高速道路での渋滞リスクの抑制及び旅行時間を抑えられる制限速度設定手法を構築し、妥当性の検討を行った。本研究の主な成果は以下の2点である。

まず、構築した最適制御モデルを用いて最適制御を行った結果、制御前と制御後で平均旅行時間の結果に差が生じており、交通状況の改善が見られた。最適制御モデルが道路状態と交通状態を考慮し、制限速度の引き下げをサグ部手前・合流部・上り勾配部で切り替えるなどの調整を行うことで、変化する状況下において需要の分散を通したショックウェーブ発生抑制・遅延、交通流の円滑化に貢献できることが判明した。

次に、最小 PICUD 値によるペナルティの報酬設定を行ったことによって、衝突危険性のある制限速度の組み合わせを棄却でき、安全性が確保された上で利便性を向上させる制限速度制御が可能となることを明らかにした。

これらの成果より、交通状態に応じて、強化学習による可変制限速度制御により、安全性を向上させた上で旅行時間を短縮することが可能であることを明らかにした。

### (2) 今後の検討事項

今後の改善点として、まずは所与の交通量を設定するのではなく、動的な交通状態の変動に対応できるように拡張する。そのためには、シミュレーションコストの削減(計算時間の削減)にも注力していく必要がある。加えて、最適制御による交通流改善の効果を向上させるために、制限速度の調整方法においても再検討を行う。

謝辞：本研究は JSPS 科研費 19H02268, および 20KK0334 の助成を受けたものです。

参考文献

- 1) 塩見康博：高速道路における動的速度マネジメントについて，国際交通安全学会誌，Vol.45，No.3，pp.172-181，2020.
- 2) Zhibin Li, Pan Liu, Chengcheng Xu, Hui Duan, and Wei Wang：Reinforcement Learning-Based Variable Speed Limit Control Strategy to Reduce Traffic Congestion at Freeway Recurrent Bottlenecks, IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL.18, NO.11, 2017.
- 3) 株式会社アイ・トランスポート・ラボ『MicroAVENUE』，2022年1月24日閲覧，<https://www.i-transportlab.jp/index/products/microavenue/>
- 4) 鈴木弘司，松村悠貴：都市高速道路合流部の錯綜事象における衝突危険性評価に関する研究，土木学会論文集 D3(土木計画学)，Vol.71，No.5(土木計画学研究・論文集第 32 卷)，pp.L\_493-L\_502，2015.

(??.?受付)

Development of reinforcement-learning-based variable speed limit using microscopic traffic flow simulation on freeways

So FUJIMOTO, Yasuhiro SHIOMI and Hisatomo HANABUSA