

第 IV 部門

強化学習を用いた高速道路における制限速度探索手法の構築

立命館大学大学院

学生員 ○藤本 想

立命館大学

正会員 塩見 康博

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¹⁾。可変制限速度のアルゴリズムは色々と提案されているが、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、現実の交通流に必ずしも対応したものではない。そこで、強化学習を用いて車両挙動などを考慮した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。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、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. MicroAVENUE の概要

MicroAVENUE²⁾は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。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し、個別車両ごとにパラメータを分布させ、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。

3. 本研究でのシナリオ設定

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。また、隣接する区間同士での制限速度差が 10km/h 以上存在した場合、図 1 に示すように学習区間内の次の区間に近い標識から制限速度を 5km/h ずつ調整し、段階的にする。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多い速度値であると考えたことが理由である。

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。

シミュレーション実行時に、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を強化学習で用いる交通量として設定した。30 分間交通量として小型車

（本線 300 台、合流 120 台）・大型車（本線 50 台、合流 20 台）・バス（本線 50 台、合流 20 台）となっている。

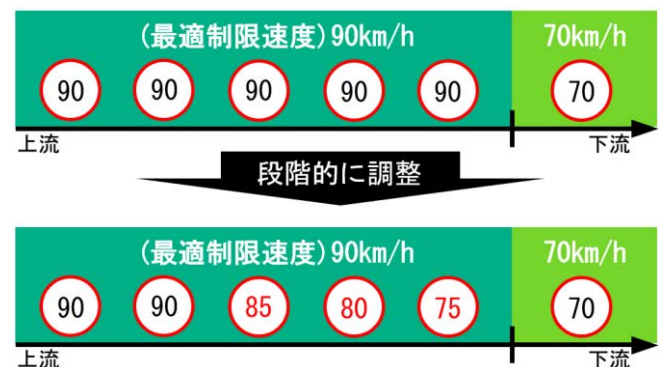


図 1 同一学習区間内の標識の速度調整

4. 強化学習モデルの構築

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2に示す。

本研究においては、交通状態（交通量・各区間で設定された制限速度）に応じてMicroAVENUEが再現度の高い交通流を示すため、総ペナルティの最小化を通じた最適な制限速度調整手法の探索が可能となる。

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方法を探索し決定することが可能となる。

5. シミュレーション結果

最適制御によって、図3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。

各区間における制限速度調整によって、合流部やサグ部でのショックウェーブの抑制・遅延に効果をもたらした。図4に最適制御前後の交通状態の変化を示す。赤の実線部が合流部で赤の破線部がサグ部である。

また、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。

6. 本研究の成果と今後の検討事項

本研究を通して、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。

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、制限速度の調整方法においても再検討を行う。

また、車線の増加や急カーブ、工事や交通事故といった、幅広い高速道路や状況に対応できるようにプログラムを補強することも考えていく。

【参考文献】

- 1) 塩見康博:高速道路における動的な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) 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/>

行動ポリシー『制限速度の調整』⇨最善策を探索

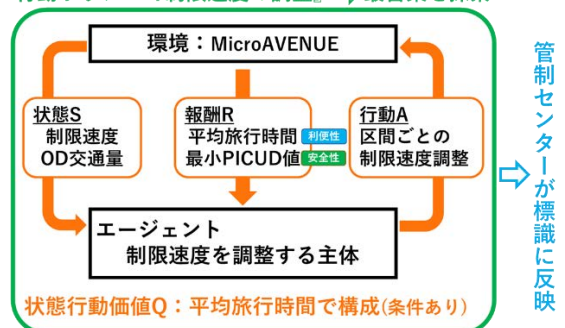


図2 VSL強化学習の概略図

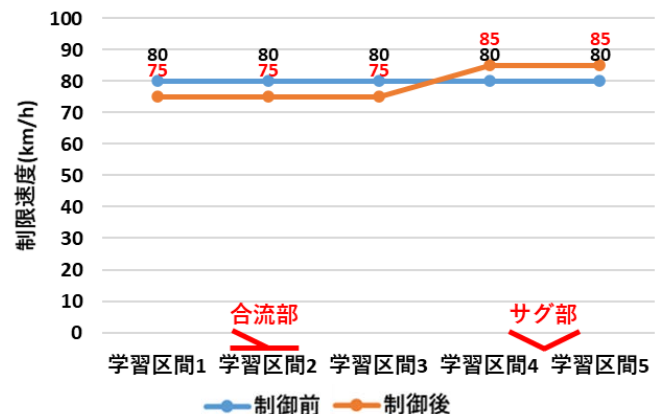


図3 最適制御前後の制限速度の変化

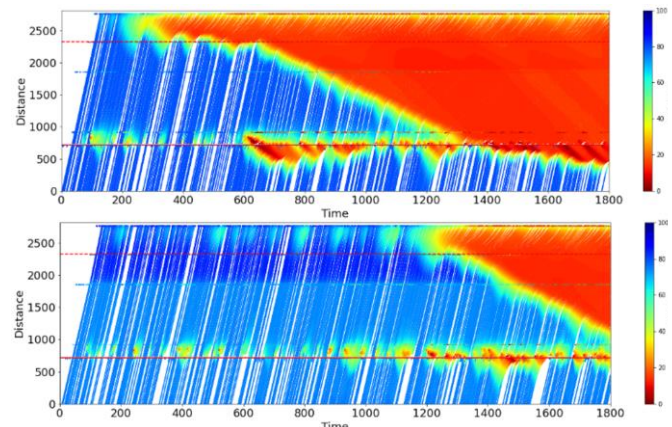


図4 最適制御前(上)後(下)の交通状態の変化