



2026年6月10日

報道関係各位

リカレント強化学習により都市鉄道の省エネルギーな速度制御を実現 ～動的計画法で生成した参照軌道に近い走行軌跡を再現～

【本研究の要点】

- 都市鉄道の省エネルギーな速度制御に向けたリカレント強化学習フレームワークを開発。
- 一時点の観測だけでは把握しにくい列車の慣性・勾配・力行・制動応答の遅れなどを考慮するため、Long Short-Term Memory (LSTM) を組み込んだ Recurrent Soft Actor-Critic (RSAC) を採用。
- 約 2,000 m の駅間走行シミュレーションにおいて、提案手法が動的計画法 (DP) で生成した参照軌道に近い走行軌跡を示し、正確かつ省エネルギーな列車運転への適用可能性を確認。

【研究の概要】

上智大学 理工学部 機能創造理工学科の宮武 昌史教授と同大学大学院 理工学研究科 理工学専攻グリーンサイエンス・エンジニアリング領域の Mingyu Lyu 氏は、都市鉄道の省エネルギーかつ正確な速度制御に向けて、リカレント強化学習 (*1) を用いた自動列車運転制御フレームワークを開発しました。

都市鉄道では、短い駅間を走行しながら、速度制限・勾配・定時性・乗り心地・エネルギー消費を同時に考慮する必要があります。従来、動的計画法 (DP, *2) やモデル予測制御 (MPC, *3) などの最適制御手法により、省エネルギーな速度パターンを計算する方法が用いられてきました。しかし、これらの手法は車両特性・走行抵抗・線路勾配・運行条件などのモデル精度に依存するため、条件が変化した場合には再最適化が必要となり、制御性能の低下を招く恐れがあります。

これに対し、本研究では時系列記憶を組み込んだ強化学習手法である Recurrent Soft Actor-Critic (RSAC, *4) を採用しました。RSAC は Long Short-Term Memory (LSTM, *5) により過去の走行履歴を保持しており、列車の現在速度・残り距離・残り時間・現在および次の速度制限といった状態量を入力としつつ、一時点の観測では把握しにくい列車の慣性・勾配の影響・力行・制動の応答遅れなどを捉えながら操作を決定します。これにより、過去の走行履歴を踏まえた滑らかな速度制御の実現を図りました。

約 2,000 m の駅間走行を対象としたシミュレーションでは、提案手法である RSAC-BC (*6) が、DP で生成した参照軌道 (*7) に近い加速・惰行・制動パターンを示しました。エネルギー消費は、DP による参照解の 13.63 kWh に対して 15.18 kWh であり、フィードフォワード型の SAC-BC (*8) の 26.58 kWh より大幅に低い値を示しました。これにより、リカレント構造と行動クローン事前学習を組み合わせた強化学習手法が、都市鉄道の省エネルギーかつ正確な自動運転制御に有効であることが示されました。

本成果は、多様な路線・車両条件への展開を見据えた次世代の自動列車運転システムの基盤技術として位置づけられます。今後は、不確実な運行条件、複数列車の協調制御、回生ブレーキや地上蓄電システムを含む鉄道電力システムとの統合最適化へ発展させることで、鉄道のさらなる省エネルギー化と持続可能な都市交通の実現に貢献することが期待されます。

本研究成果は、2026年3月25日に[国際学術誌「IEEE Access」](#)にオンライン掲載されました。

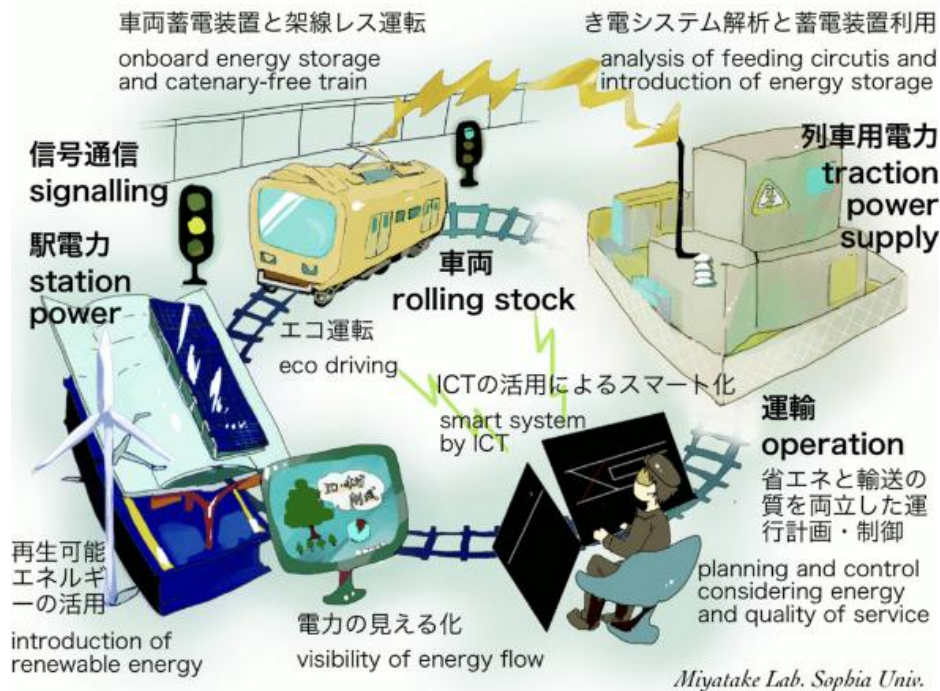


図: 本研究の概要

【研究の背景】

都市鉄道は、大都市圏の人流を支える基幹的な交通インフラであり、安全性や定時性に加えて、エネルギー消費の低減が重要な課題となっています。特に、短い駅間で加速・惰行・制動を頻繁に繰り返す都市鉄道では、速度制限・線路勾配・停止位置・目標走行時間・乗り心地といった多様なパラメータを同時に考慮した、最適な速度制御が求められます。

これまで、列車の省エネルギー運転には DP や MPC の他に遺伝的アルゴリズム、混合整数線形計画法などのモデルベース手法が用いられてきました。これらの手法は、線路条件や車両特性が既知である場合には優れた速度プロファイルを求められる一方、走行抵抗、勾配、信号条件、車両応答などのモデル精度に強く依存します。そのため、運行条件が変化した場合には、ロバスト性の低下や再最適化に伴う計算負荷が課題となります。

この課題に対し、近年は強化学習を列車制御に応用する研究が進められています。強化学習は、環境との相互作用を通じて長期的な報酬を最大化する方策を学習する手法であり、モデルを完全に明示しなくても、速度制限の遵守・省エネルギー性・定時性・乗り心地を考慮した運転操作を獲得できる可能性があります。

一方で、強化学習を鉄道のような安全制約の厳しいシステムに適用するには、試行錯誤による探索の安全性、報酬設計の難しさ、学習の不安定性が課題となります。また、列車の運動は、慣性・勾配・力行・制動応答の遅れなど、長い時間にわたる依存性を持つため、一時点の観測だけでは適切な操作を判断しにくい場合があります。

本研究では、このような部分観測性と時間的依存性を扱うため、LSTM を組み込んだ RSAC を用いました。さらに、DP で生成した省エネルギー運転の走行軌跡を行動クローン事前学習に用い、制約違反を防ぐ安全フィルターを組み合わせることで、モデルベース最適制御の知見と強化学習による適応的制御を接続する都市鉄道速度制御フレームワークの構築を目指しました。

【研究結果の詳細】

本研究では、都市鉄道の駅間走行を対象として、列車の速度制御を部分観測マルコフ決定過程として定式化しました。制御エージェントは、終端駅までの残り距離・目標走行時間までの残り時間・現在速度・現在の速度制限・次の速度制限変更地点までの距離・次区間の速度制限を観測し、力行・惰行・制動に対応する連続的な制御指令を出力します。一方で、線路勾配・走行抵抗・列車の慣性・力行・制動応答の遅れなど運動に影響する要因は、一時点の観測だけ

では十分に把握できません。そこで、過去の観測系列を内部状態として保持できる LSTM を組み込んだ RSAC を用いました。

提案手法では、RSAC に加えて、行動クローン事前学習と安全フィルターを組み合わせました。行動クローン事前学習では、DP で生成した省エネルギー運転の走行軌跡を教師信号として用い、AI に加速・惰行・制動の基本的な運転パターンを事前に学習させました。これにより、強化学習の初期段階で生じやすい不安定な探索を抑え、方策の収束を速めることを目指しました。また、安全フィルターは、AI が出力した操作が速度制限・制動距離・力行・制動制約を満たすように補正する仕組みです。これにより、学習済み方策だけに安全性を委ねるのではなく、方策と安全フィルターを組み合わせた閉ループシステムとして、制約に従う運転操作を実行できるようにしました。

シミュレーションでは、全長約 2,000m の駅間走行を対象としました。線路条件として、始端側 300m に-1.8%の下り勾配、中央部に平坦区間、終端側 300m に+1.8%の上り勾配を設定し、最高速度を 80 km/h、目標走行時間を 140 秒としました。この条件下で、提案手法である RSAC-BC を、従来の強化学習手法や DP+MPC 方式と比較しました。その結果、RSAC-BC は、比較した学習ベース手法の中で最も安定した学習収束を示しました。行動クローン事前学習を行わない RSAC も最終的には同程度の報酬水準に到達しましたが、RSAC-BC は学習初期から中期にかけて、より速く、ばらつきの小さい収束を示しました。これは、DP で生成した走行軌跡を用いた事前学習が無駄な探索を減らし、効率的な方策獲得に寄与したことを示しています。

走行性能の比較では、RSAC-BC が、DP の参照軌道に近い加速・惰行・制動の構造を再現しました。エネルギー消費量は、DP による参照解が 13.63 kWh、DP+MPC が 14.12 kWh であったのに対し、RSAC-BC は 15.18 kWh でした。これは、完全なモデル知識を前提とする DP には及ばないものの、単一の学習済み方策によって参照解に近い省エネルギー性能を実現できたことを示しています。一方、リカレント構造を持たない SAC-BC ではエネルギー消費量が 26.58 kWh となり、時系列依存性を扱う RSAC の有効性が確認されました。

また、RSAC-BC は、DP+MPC のように、あらかじめ計算された速度参照をオンラインで追従する方式とは異なり、過去の観測履歴を含む情報から、力行・惰行・制動の操作を直接出力します。そのため、モデルベース制御の高い最適性と、学習ベース制御の適応性を接続する手法として位置づけられます。今回の結果は、多様な路線・車両条件への展開を見据えた、省エネルギーかつ正確な自動列車運転制御の基盤となり得ることを示しています。

今後は、実運用に近い条件での検証が重要です。具体的には、パラメータ変動・センサー雑音・速度制限の確率的变化などを含む条件でのロバスト性評価が課題となります。さらに、複数列車の協調制御・列車間隔制御・回生ブレーキ・変電所容量・地上蓄電システムを含む鉄道電力システムとの統合最適化へと発展させることで、鉄道のさらなる省エネルギー化に寄与することが期待されます。これにより、正確かつ省エネルギーな列車運行の実現可能性が高まり、持続可能な都市交通の構築と気候変動抑制における鉄道の役割強化につながることを期待されます。

※本研究は、日本学術振興会 科学研究費助成事業(JSPS KAKENHI、課題番号:23K03822)の助成を受けて実施したものです。

【用語】

*1 強化学習: AI が環境との相互作用を通じて、長期的な報酬が最大になるように行動方策を学習する機械学習手法。

*2 動的計画法(DP: Dynamic Programming): 複雑な最適化問題を小さな部分問題に分けて解く手法。列車速度制御では、線路条件や車両特性が既知である場合に、エネルギー消費を抑えつつ目標時刻に到着する速度プロファイルを求めるために用いられる。

*3 モデル予測制御(MPC: Model Predictive Control): 現在の状態から一定時間先までの挙動をモデルに基づいて予測し、その範囲で最適な操作を計算する制御手法。

*4 Recurrent Soft Actor-Critic(RSAC): 連続制御問題に用いられる強化学習手法 Soft Actor-Critic(SAC)に、時系列情報を扱うリカレント構造を組み込んだ手法。

*5 Long Short-Term Memory(LSTM): リカレントニューラルネットワークの一種で、時系列データに含まれる長期的な依存関係を扱うための構造。

*6 Recurrent Soft Actor-Critic with Behavior Cloning(RSAC-BC): 再帰型強化学習法 RSAC に、熟練者の行動を模倣する Behavior Cloning(BC)を組み合わせた手法。過去の情報を記憶しながら学習できるため、部分観測環境において安定した学習が可能となる。

*7 参照軌道: 性能比較や学習の基準として用いる走行軌跡。

*8 Soft Actor-Critic with Behavior Cloning(SAC-BC): フィードフォワード型ニューラルネットワークを用いた強化学習手法 Soft Actor-Critic(SAC)に、熟練者の行動を模倣する Behavior Cloning(BC)を組み合わせた手法。まず熟練者データを用いた教師あり学習によって方策を事前学習し、その後 SAC による強化学習によって方策を自律的に改善することで、学習初期の探索効率や安定性を向上させる。

【論文名および著者】

媒体名:	IEEE Access
論文名:	Recurrent Reinforcement Learning for Urban Railway Speed Control
オンライン版 URL:	https://ieeexplore.ieee.org/document/11455996
著者(共著):	Mingyu Lyu, Masafumi Miyatake

【研究内容に関するお問い合わせ先】

上智大学 理工学部 機能創造理工学科

教授 宮武 昌史 (E-mail: m@miyatake.main.jp)

※本資料の配布先記者クラブ: 文部科学記者会、科学記者会、国土交通記者会

報道関係のお問合せ: 上智大学広報グループ

TEL:03-3238-3179 E-mail: sophiapr-co@sophia.ac.jp