

Q-routing アルゴリズムによって現れる経路選択行動の遷移

Route Selection Behavior Generated by Q-routing Algorithm

内田 英明^{*1}
Hideaki UCHIDA

藤井 秀樹^{*1}
Hideki FUJII

吉村 忍^{*1}
Shinobu YOSHIMURA

荒井 幸代^{*2}
Sachiyo ARAI

^{*1} 東京大学大学院工学系研究科
School of Engineering, The University of Tokyo

^{*2} 千葉大学大学院工学研究科
Graduate School of Engineering, Chiba University

In this paper, we describe the impact of changes in road network on drivers' route selection behavior. In the almost all traffic simulation model, driver agents choose the shortest route, while the drivers in the real world choose their routes through their own restricted knowledge. Therefore, we developed a reinforcement-learning-based routing algorithm and implemented it in a traffic simulator. This model is applied to simulate the tram railway extension plan in Okayama City.

1. はじめに

1.1 本研究の背景

道路交通は現代社会の基盤システムであると同時に多くの問題を生んできた。これらを解決するため様々な施策が提案されており、近年ではその評価に交通流シミュレーションが採用されることが多い。しかしシミュレーションにおいて再現される交通流にはいくつかの注意すべき性質が存在する。

多くのシミュレータにおいて運転者の経路選択はネットワーク全体のコスト情報に基づいて行われる。一方、現実の運転者は限定された周囲の状況と経路に対するそれまでの知識に基づいて経路を選択する。この差異はシミュレーションによって再現される交通現象に大きく影響する。例えば、交通施策の実施により、ある時点から道路ネットワークが変化する場合、ネットワークの全コスト情報を利用できるモデルでは、常に各運転者が最適経路選択を行うため、交通状況は即座に定常的な状態に落ち着く。しかし、過去の走行経験に基づいて経路選択が行われるとすれば、知識を修正していく過程で過渡的な交通状況が生じると考えられる。交通施策の決定する自治体などにとって、この過渡期における現象を事前評価することは、十分な時間経過の後収束する定常状態を予測することと同様に重要である。

1.2 本研究の目的

著者らはこれまで、マルチエージェント型交通流シミュレータ MATES[吉村 04, Fujii 11]に強化学習に基づいたルーティングアルゴリズムである Q-routing[Boyan 94]を導入し、過去の経験に基づいた経路選択の再現を試みてきた[内田 11]。

本稿では実問題への適用に主眼を置き、岡山県岡山市で検討されている路面電車の軌道延伸計画に関し、実施直後の過渡的な交通状況を再現すると同時に、定常状態に収束するまでシミュレートすることを目的とする。

2. 経路選択アルゴリズム

2.1 Q-routing の概要

Q-routing は強化学習の一つである Q 学習の枠組みに基づいた自律分散型の経路選択アルゴリズムである。もともとパケットルーティングの分野で提案されたアルゴリズムであるが、動的

なネットワーク環境への適応力が高いことで知られている。また、価値関数によって各エージェントの経験的知識を表現できるため、交通施策による経路選択行動の時間的遷移を確認できる点でも優れている。Q-routing を自動車の経路選択に応用する場合のアルゴリズムは次のように見なすことができる。

エージェントは、現在所属している交差点 x 、その隣接交差点 y 、目的地 d の組に対してそれぞれ価値関数 $Q_x(y, d)$ を持ち、この価値関数表を更新しながら経路を選択する。ここで $Q_x(y, d)$ は現在地 x から y を経由して d に到達するまでにかかる旅行時間の推定値を表し、交差点 x において、エージェントは価値関数 $Q_x(y, d)$ が小さな隣接交差点 y を優先して選択する。またエージェントは y に到着するとただちに xy 間の旅行時間 t_x と y から d までの旅行時間の推定値 q_y を用いて $Q_x(y, d)$ を更新し、次の選択に反映させる。

我々の過去の研究[内田 11]ではさらにパケットルーティングと自動車の経路選択との差異に着目した拡張を行っている。

2.2 行動選択手法

前節で定義した価値関数 $Q_x(y, d)$ を用いて、エージェントは次に進むべき隣接交差点を選択する。最もシンプルな手法としては価値関数が最小である隣接交差点を必ず選択する greedy 選択、価値関数を無視した乱数によって選択を行うランダム選択があるが、本研究ではその中間的な手法である ϵ -greedy 選択と Boltzmann 選択の 2 つについて検討した。 ϵ -greedy 選択は微小な確率 ϵ を与え、 $1-\epsilon$ で greedy 選択、 ϵ でランダム選択を行う。Boltzmann 選択は価値関数の値に応じた確率で選択を行う。

ネットワーク上に複数の最適経路があるケースでも、学習中の Q 値には若干の差が生じることがある。 ϵ -greedy 選択ではその差を過大に反映した結果、エージェントの経路選択から多様性が失われる場合がある。一方、ソフトマックス関数によって選択確率を算出する Boltzmann 選択では、価値関数が同程度の経路には同程度の選択確率が振り分けられるため、エージェントの経路選択の多様性を保ったまま学習を収束させられる。このことから、学習アルゴリズムを含んだ手法の中でも Boltzmann 選択が現実の運転者に最も近い挙動を示すと考えられる。

3. 路面電車軌道延伸シミュレーション

3.1 シミュレーション条件

提案する経路選択アルゴリズムを用いて岡山県岡山市の路面電車軌道延伸計画のシミュレーションを行う。シミュレーション対象領域を図 1 に示す。領域は約 3km 四方であり、262 ノード



図1 対象の道路ネットワーク(岡山市中心部)

(交差点)・381リンク(道路)から構成される。入力する交通量は2001年に行われた社会実験のデータを基に設定した[Fujii 11]。図1中の破線は現状の路面電車の営業区間であり、実線は延伸計画が提案されている区間を表す。

実験では延伸前と延伸後のそれぞれを計算した。その際、軌道延伸の影響として、新規軌道敷設による主要道の車線数減少、路面電車への乗換による発生交通量の微減、の2点を想定した。

延伸前のシミュレーションは価値関数を[0,1]で正規化したユークリッド距離で初期化し、定常状態がどの程度再現されるかを確認するため、A*アルゴリズム[Hart 68]との比較を行なった。A*アルゴリズムはネットワーク全体の情報を利用して最適解を算出する経路選択アルゴリズムであり、ここでは参照解として用いる。

延伸後は、延伸前のケースで学習済みの価値関数を初期値として利用し、延伸前の走行経験を持ったエージェントを作成した。価値関数を完全に読み込む learned, 半分に割引いて読み込む defective, 全く考慮しない reset の3者を比較する。

3.2 延伸前のシミュレーション

軌道延伸前のシミュレーション結果を図2に示す。図1中の清輝橋から岡山駅へ向かうエージェントの平均旅行時間を表している。Q-routingは、シミュレーション開始直後の学習がほとんど進んでいない段階では旅行時間が長くなっているが、最終的にはほぼ収束している様子が確認できる。収束した旅行時間はA*アルゴリズムと一致した。A*アルゴリズムは旅行時間を最小化することが数理的に保証されているため、Q-routingを用いても適切な定常状態も再現できていることを示している。

3.3 延伸後のシミュレーション

続いて、軌道延伸後のネットワークでシミュレーションを行った結果を図3に示す。図2と同様に清輝橋から岡山駅へ向かうエージェントの平均旅行時間を示している。ここでは読み込む価値関数の効果を割引くことで、エージェントが過去の走行経験を保持することの効果を観察した。

価値関数を完全に読み込む learned の旅行時間が最も安定し、走行経験の有用性を示した。しかし、learnedにおいても、過渡状態では学習終了時点と比較し最大30%程度の旅行時間の増大が見られた。結局、延伸前の走行経験は延伸後のバイアスにも成り得ることを示している。実際、延伸によって図1の

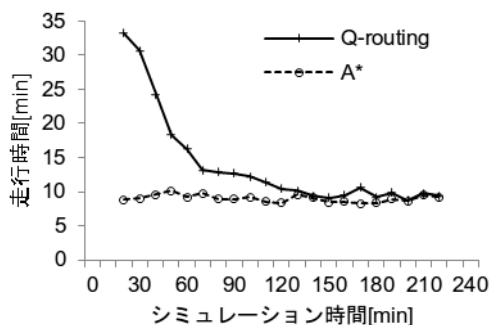


図2 軌道延伸前のシミュレーション結果

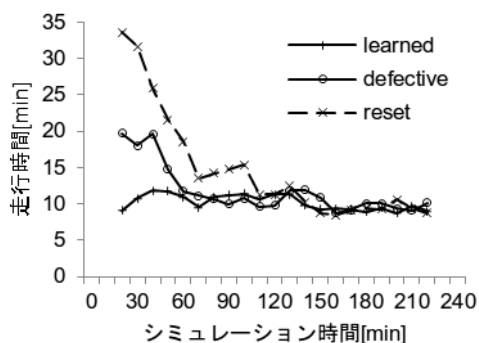


図3 軌道延伸後のシミュレーション結果

岡山駅～大学病院間などで渋滞が悪化し、それに合わせてエージェントが再度学習を行った様子を確認している。

4. おわりに

本稿では運転者の過去の走行経験と多様性を経路選択に反映させるため、交通流シミュレーションにエージェントの学習機能を実装し、岡山県岡山市の路面電車軌道延伸計画の問題に適用した。従来のシミュレーションでは再現できなかったエージェントのバイアスによる延伸後の過渡状態の渋滞現象を確認することができた。

今後の課題として、経路選択の遷移の詳細な解析や遷移を制御するための手法の検討などが考えられる。

参考文献

[Boyan 94] J. Boyan, M. L. Littman: Packetrouting in Dynamically Changing Networks: a Reinforcement Learning Approach, *Advances in Neural Information Processing Systems*, Vol. 7, pp.671-678, 1994.

[Fujii 11] H. Fujii, T. Sakurai, S. Yoshimura: Virtual Social Experiment of Tram Railway Extension Using Multi-Agent-Based Traffic Simulator, *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol. 15, No. 2, pp. 226-232, 2011.

[Hart 68] P. E. Hart, N. J. Nilsson, B. Raphael: A Formal Basis for the Heuristic Determination of Minimum Cost Paths, *IEEE Transactions on Systems Science and Cybernetics*, Vol. 4, No. 2, pp. 100-107, 1968.

[内田 11] 内田英明, 藤井秀樹, 吉村忍: 道路ネットワークの変化に対する経路選択の学習, 合同エージェントワークショップ&シンポジウム2011(JAWS2011)予稿集, 2011.

[吉村 04] 吉村忍, 西川紘史, 守安智: 知的マルチエージェント交通流シミュレータ MATES の開発, *シミュレーション*, Vol. 23, No. 3, pp. 228-237, 2004.