

利用者均衡をシステム最適へと促す混雑時料金の設定

Inducing System-Optimum from User Equilibrium during Peak Periods of Traffic Flow

石川翔太 荒井幸代
Shota Ishikawa Sachiyo Arai千葉大学大学院工学研究科
Graduate School of engineering Chiba University

In order to induce the system-optimum situation from user equilibrium during peak periods of traffic flow, a road traffic policy of congestion toll pricing has been considered. In general, this policy is applied during peak periods, but it seems not rational time to start control. This paper proposes the method to find a starting time to control route allocation by means of the congestion toll pricing. In our method, an action trajectory of SO situation is observed, and find the common states by comparing the state transitions of system-optimum with that of user equilibrium. Then, the common states of both situations are defined as a set of states that is not necessary to control traffic flow. Through some experiments, we examine the performances of our approach.

1. はじめに

近年、道路システムの高度情報化によって、ドライバーはリアルタイムで精度の高い情報を入手できるようになった。この情報に基づいて、各ドライバーが自らの旅行時間を最小化するルートを選択する時、交通ネットワークは「利用者均衡 (User Equilibrium: 以下 UE と記述する)」と呼ばれる均衡状態に至る。しかし、道路の混雑時において、UE は交通ネットワークの「システム最適 (System Optimum: 以下 SO と記述する)」とは一致しない ($UE \neq SO$) 問題がある。

UE と SO を一致させるための交通施策として、混雑時料金の導入が考えられる。混雑時料金とは、混雑した道路に課金することによって交通量を調整する最も導入効果の高い交通施策とされる。現状では、混雑時料金は時間帯別に設定され、一定時間の需要予測に基づいている。一方、道路状況は時々刻々と変化するため、一定時間同じ状況が続くことは考えにくい。

そこで、本論文では、混雑時料金を導入するタイミングについて考察を行う。目的地に対して複数のルートが存在する分岐点 (ノード) で各ドライバーの選択ルートを指示することを制御と呼ぶ。従来、時間帯別に混雑時料金を導入する制御に対して、提案手法ではネットワークの状態を観測し、制御を行うタイミングを判定することができる。具体的には、UE と SO に至るまでの、それぞれの状態遷移に着目し、共通の状態において異なる制御を施した状態集合を抽出する。この状態集合に至った時点を制御のタイミングとなる「制御境界」とし、この状態から制御を開始した場合と、それ以外の任意の状態から制御を開始した場合の交通ネットワークの交通流をシミュレーション実験によって比較し、制御境界状態から制御を開始することの有効性を検証する。

2. モデリング

2.1 交通流モデル

本論文では、セルオートマトンを用いて交通流の各状態を離散化し、各時刻における交通ネットワークの状態と個々の自動車の挙動を時間ごとに観測する。また、走行ルールは、前方情報を考慮した拡張 NaSch モデル [増淵 09] に従う。

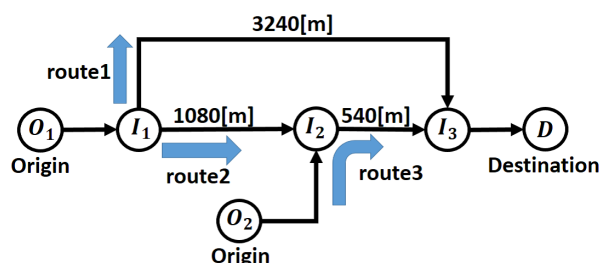


図 1: 対象ネットワーク

2.2 対象ネットワーク

本論文の対象とする交通ネットワークを図 1 に示す。リンクを道路、分岐ノード I_1 、分岐合流ノード I_2, I_3 を信号機、始点ノード O_1, O_2 を起点、終点ノード D を目的地とし、起点と目的地の間に三つのルートが存在する。ここで、route2 は O_1 から D までの最短経路である。route2 は O_2 から D までの route3 と一本の道路を共有しているため、route2 を選択するドライバーが増えれば、混雑による旅行時間の増加を引き起こす。そこでネットワークを SO にするためには、交通量を route1 と route2 に適切に配分しなければならない。

対象ネットワークにおける UE を、横軸を一時間当たりの自動車流入量、縦軸を平均旅行時間とし、図 2 に示す。流入量の増加に伴い、最短経路の route2 の旅行時間が増加していることがわかる。

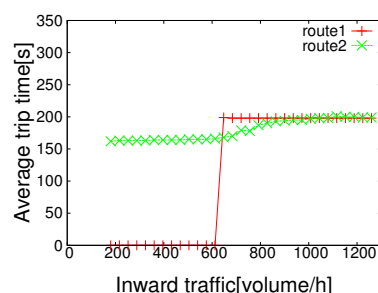


図 2: 利用者均衡 (UE) での平均旅行時間

連絡先: 石川翔太, 千葉大学大学院工学研究科, 千葉市稲毛区 弥生町 1-33, 043-251-1111(代表)

3. 提案手法

3.1 システム最適行動の抽出：Q 学習

対象ネットワークの SO を実現する制御方を Q 学習を用いて獲得する。Q 学習は、環境がマルコフ決定過程であることを条件として $\langle S, A, P_a, R \rangle$ (S : 状態集合, A : 行動集合, P_a : 状態遷移確率, R : 報酬関数) のとき, R を所与として最適な方策 π^* を獲得する。本論文では、複数の出力ルートが存在するノード I_1 にエージェントを配置し、交通流を配分する方策を学習させる。また、状態を「各ルートの車両密度と交通流入量」、行動を「ルート割当て」とする。エージェントは状態 $s \in S$ を知覚し、方策 π に基づいてルート割当て $a \in A(s)$ を行う。エージェントが各車ごとにルート割当てをし、目的地に着いた時、「旅行時間」を負の報酬 r として受け取ることで、旅行時間を最小化する制御方策が学習できる。

Q 学習は状態 s , 行動 a の価値 $Q(s, a)$ を式 1 により更新する。ここで $\alpha (0 \leq \alpha \leq 1)$ は学習率, $\gamma (0 \leq \gamma \leq 1)$ は割引率を表し, k は s において a を選択し, $Q(s, a)$ を更新した回数である。

$$Q_{k+1}(s, a) = Q_k(s, a) + \alpha[r + \gamma \max_{a' \in A(s')} Q_k(s', a') - Q_k(s, a)] \quad (1)$$

3.2 特徴期待値から制御境界を求める

交通ネットワークの UE と SO が一致している状況ではルート配分は不要である。この状況では、UE に至る状態遷移を構成する状態集合と、SO に至る状態遷移を構成する状態集合は一致している。一方、両者が一致しない状況では、両者の間で状態遷移が異なる状態集合が存在するはずである。そこで、本研究では、UE と SO の間で異なる状態遷移を構成する状態集合のうち、両方の状態遷移に含まれる状態以外の集合を、「制御対象状態」と呼ぶことにする。制御対象状態を抽出し、制御不要な状態での課金を避け、必要な状態に対する課金によって交通ネットワークの SO 維持が期待できる。

本論文では、制御対象状態を、特徴量 ϕ と特徴期待値 μ [Abbeel 04] を用いて識別する。各状態は、特徴量と呼ばれるベクトル $S \rightarrow [0, 1]^k$ で表す。また、方策 π に従ったときのエージェントの行動軌跡 (状態遷移) から得られる期待割引累積特徴量は式 (2) によって計算される。

$$\mu(\pi) = E\left[\sum_{t=0}^{\infty} \gamma^t \phi(s_t) | \pi\right] \quad (2)$$

ここで状態 $s \in S$ において UE と SO の特徴期待値をそれぞれ、 $\mu(\pi_{UE}(s))$, $\mu(\pi_{SO}(s))$ とする。以下に二つの特徴期待値の関係と制御対象状態の識別法を定義する。

1. 領域 1: $\mu(\pi_{UE}(s)) > 0$, $\mu(\pi_{SO}(s)) = 0$ の場合、ネットワークは $UE \neq SO$ となり、制御対象状態である。
2. 領域 2: $\mu(\pi_{UE}(s)) = 0$, $\mu(\pi_{SO}(s)) > 0$ の場合、ネットワークは $UE \neq SO$ となり、制御対象状態である。
3. 領域 3: $\mu(\pi_{UE}(s)) > 0$, $\mu(\pi_{SO}(s)) > 0$ の場合、ネットワークは $UE = SO$ であるため、制御が不要な状態である。
4. 領域 4: $\mu(\pi_{UE}(s)) = 0$, $\mu(\pi_{SO}(s)) = 0$ の場合、UE と SO それぞれの均衡時に一度も出現しない状態である。この状態も制御対象状態とする。

図 3 に、交通流入量が 900[volume/h] の時の交通ネットワークの状態を route1 と route2 の各車両密度の組合せで表した時の、各状態の分類を示す。図 3 の領域 3 に属する状態が制御不要で、それ以外の領域へ状態遷移する時点から制御を必要とする状態と判断する。

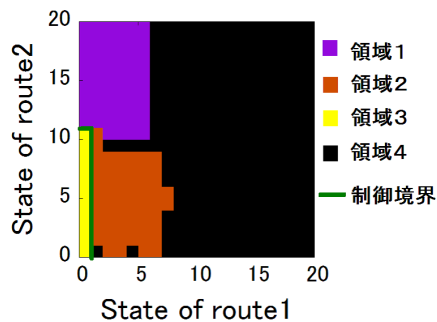


図 3: 流入量 900[volume/h] における状態の分類：制御要/不要の領域

4. 実験方法

対象ネットワークを流れる一日の交通量の変化を再現するため、1 時間毎に交通流入量を変化させる。本提案手法で分類した各状態の制御要/不要に応じてルート割当てを実施した場合とヒューリスティクスに基づいて制御を開始する 2 つの方法による場合とを比較し、提案手法の有効性を示す。

以下は、比較に用いた制御を開始するタイミングに関するヒューリスティクスである。

- はじめから制御……予測交通量のピークの交通量を最適に配分できる割合を求め、これに基づいて、シミュレーション開始時から終了時まで制御を続ける。道路が常に有料になっている状況。
- 時間帯別に制御……予測交通量のピークの交通量に至った時点で制御を開始して、ピーク時の終了時に合わせて制御を終了する。混雑する時間帯を予想して道路を有料化する状況。
- 提案手法：制御対象状態に限定した制御……道路の状態が制御境界外の状態に限り、制御を行う。道路状況に応じた混雑料金を導入している状況。

5. 結論および今後の課題

目的地へのルートが複数存在する交通ネットワークでは、交通量の増加に伴い UE と SO が一致しない状況が生じ、これを回避するために各車を適切なルートに割当てる制御が必要になる。本論文では、各状態の制御の要/不要の状態を識別する方法を提案し、制御開始と終了の適切なタイミングを決定できることを示した。

現状では、制御開始と終了のタイミングと、各ルートの交通流を制御する方策を示した。現状ではトップダウンに迂回ルートに車を割り当てる今後の課題として、提案手法で得られるルート配分を実現する課金額を決定する方法が必要となる。

参考文献

- [増淵 09] 増淵達也, 荒井幸代: 前方情報を考慮した走行ルールによるメタ安定相の発生と特徴の解析, 電子情報通信学会論文誌, NO.11, pp.1935-1944(2009)
- [Abbeel 04] Pieter Abbeel, Andrew Y. Ng: Apprenticeship Learning via Inverse Reinforcement Learning, In Proceedings of the 21st International Conference on Machine Learning, (2004)