

グラフ列挙手法により容量制約付き総社会効用 最大配分を実現する動的アルゴリズム

早川 敬一郎¹・羽藤 英二²

¹正会員 (株)豊田中央研究所 (〒 480-1192 愛知県長久手市横道 41 番地の 1)

E-mail: kei-hayakawa@mosk.tytlabs.co.jp

²正会員 東京大学大学院 (〒 113-8656 東京都文京区本郷 7-3-1)

E-mail: hato@bin.t.u-tokyo.ac.jp

自動運転車両の普及に備えて、都市ネットワークにおける自動運転車両の交通制御手法の開発が望まれている。本研究では、交通を生活行動の一部として捉えるアクティビティベースモデルに基づいて、人の移動と滞在を状態遷移モデルで定式化し、時空間プリズム制約と交通容量制約の下で総社会効用を最大化するアルゴリズムを提案する。提案アルゴリズムでは、MSA(Multi Scenario Approach)とZDD(Zero-suppressed binary Decision Diagram)を用いて、各時刻に実行し得る交通制御に対する期待効用を算出し、時間割引率を考慮した総社会効用を最大化する交通制御を逐次的に実行する。自動車が全て自動運転車である将来のネットワークにおいて、提案するアルゴリズムは、移動の一般化コストを低減させるだけでなく、移動と滞在を繰り返す利用者が一連のトリップチェーンから得る効用を最大化するような交通管制システムを実現することができる。

Key Words: アクティビティベースモデル, 時空間プリズム制約, 動的総社会効用最大配分, ZDD

1. はじめに

近年、自動運転車両の開発やシェアカーサービスの普及が進んでいる。これに伴い、従来は個人に所有され、ドライバーの意思決定に従って動かされていた自動車のあり方が大きく変わろうとしている。通信機能を充実させた自動運転車両は、周辺車両や道路インフラとの通信によって、ドライバーに意思決定を求めることなく走行できる。また、シェアカーサービスにおいては、複数の利用者が所望するトリップの組み合わせを達成する効率的な配車を実現する。

一方で、このような自動運転車両やシェアカーサービスの利用者は、移動に対する「選好」と「制約」を有している。「選好」の例としては、「少しでも早く目的地に向かいたい」ということや、「景色を楽しみながらのんびり移動したい」ということが挙げられる。一方で、「制約」の例としては、「待ち合わせ時刻までに目的地に到着する」ことや、「夜までに自宅に戻って来る」ことが挙げられる。自動運転車が利用者の個人保有である場合には自動運転車のCPUが、また、シェアカーである場合にはシェアカー予約システムのサーバーが、あるときは利用者ごとにカスタマイズされた行動履歴などから、またあるときは利用者への直接的な問いかけにより、このような利用者の「選好」と「制約」を推定し、その移動をアレンジすることが考えられる。本

研究では、このようなCPUやサーバをエージェントと呼ぶ。自動運転車両が普及した将来の交通管理システムは、このような「利用者エージェント」からの要望に応じて、有限の交通資源を最適に配分することが求められる。また、その配分に当たっては、ただ単にある地点から別の地点の移動を単独で考えるのではなく、日単位や週単位のトリップチェーンを考慮することが求められる。

このような、個々の利用者の選好、制約やトリップチェーンを考慮することができる交通分析手法として、利用者の時空間プリズム制約¹⁾を考慮したアクティビティベースの交通分析²⁾が知られている。また、有限な交通資源を最適に配分する研究としては、Nieら³⁾が、リンク容量制約を考慮した利用者均衡配分問題を定式化し、その解法を示している。一方で赤松⁴⁾およびAkamatsu and Wada⁵⁾は、同様な問題に対して、交通資源をシステム最適基準で配分した上で、適切な課金制度によってその状態を利用者均衡と一致させる手法を示している。ただし、いずれの場合においても、利用者の選好や制約は静的なものとして考えられており、動的な意思決定については考慮されていない。

本研究では、交通管理者や利用者の動的な意思決定を考慮して、赤松・和田⁵⁾の枠組みを動的な環境に拡張し、アクティビティベースの交通モデルに適用し、その解法アルゴリズムを提案する。

本研究の貢献は、以下の2点である。

- 利用者の日単位のトリップを正規化された単位時間を持つ状態遷移モデルで定義し、交通ネットワークと利用者の時空間プリズム制約を考慮した利用者の行動選択枝集合を、時空間拡張ネットワーク上で表現する。
- 時間割引率を考慮した総社会効用最大化を目的とし、交通資源の容量制約を考慮する動的総社会効用最大配分問題を定式化し、グラフ列挙手法を応用した解法アルゴリズムを提案する。

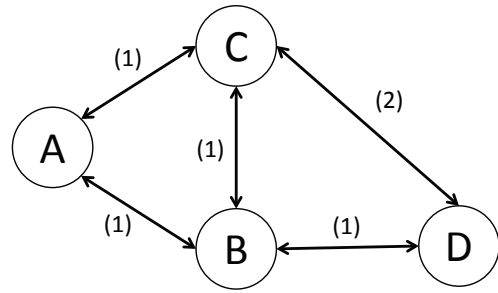


図-1 サンプルネットワーク

2. 交通管制システムモデル

本研究では、全ての自動車が交通管制システムの管制下にある自動運転車両である状況を想定する。エージェントは利用者が出発可能となったときに、交通管理システムに利用者の「選好」および「制約」を通知する。交通管理者はこの情報を用いてエージェントに適切な行動を指示し、エージェントはその指示の通りの行動を実行する。なお、以下のモデルでは、離散時刻 $t \in T = \{0, 1, \dots, \bar{T}\}$ を用いて時間を表現する。

(1) エージェントモデル

エージェントの集合を $I = \{0, 1, \dots, \bar{I}\}$ で表す。時刻 $t \in T$ におけるエージェント $i \in I$ の状態を $s_{i,t} \in S_i$ で表し、その時刻にエージェント i が取る行動を $a_{i,t} \in \Gamma(s_{i,t}) \subset A_i$ で表す。ここで、 S_i はエージェント i が取り得る状態の集合であり、 $\Gamma(s_{i,t})$ は状態 $s_{i,t}$ の下でエージェント i が取り得る行動の集合である。また、 $A_i \subset S_i \times S_i$ はエージェント i が取り得る全ての行動の集合である。時刻 t に行動 $a_{i,t}$ を取ったとき、エージェント i は報酬 $R_i(s_t, a_{i,t})$ を得る。ただし、 s_t は、時刻 t における全エージェントの状態を表している。

a) 交通ネットワーク上の行動

本研究では、交通ネットワークを考慮し、利用者の状態として「時刻」および「位置」を考える¹。行動は状態間の遷移であり、行動前後の状態の「位置」が異なる場合は「移動」を、同一である場合は「滞在」を表す。エージェント i が取り得る全ての「移動」および「滞在」行動の集合をそれぞれ A_i^M および A_i^S で表す。ただし、 $A_i^M \cup A_i^S = A_i$ である。このとき、ノードの集合 N とリンクの集合 E からなる交通ネットワーク (N, E) が与えられれば、エージェント i の交通ネットワーク上の行動は、状態 $s_{i,t} \in S_i$ をノードとし、状態 $s_{i,t}$ の下で実行可能な行動 $a_{i,t} \in \Gamma(s_{i,t})$ をリンクとする時空間拡張ネットワーク上の有向グラフ G_i で表現できる。

¹ 本研究では簡単のため「時刻」と「位置」のみを考えたが、交通モードなどその他の情報も状態として導入可能である。

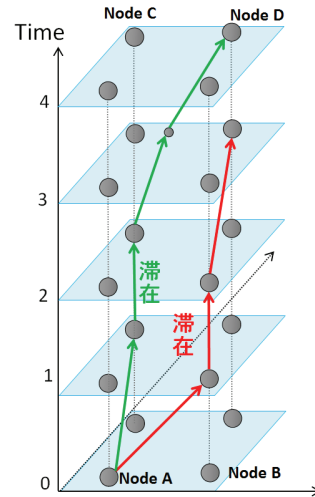


図-2 時空間拡張ネットワーク上の行動軌跡

例として、図-1に示すサンプルネットワークを考える。図中のかっこ内に示された数値は当該リンクを移動するのにかかる時間であり、リンク CD の移動には2時刻ステップ、その他のリンクの移動には1時刻ステップを必要とする。このとき、ノード A を出発し、ノード B またはノード C に1時刻ステップ滞在した後に、ノード D に向かうエージェントの軌跡は、図-2のように表現できる。

このように、交通ネットワーク上の行動を時空間拡張ネットワークで表す手法を用いれば、ネットワーク上の滞在を考慮したトリップチェーンを明示的に考えることができ、出発地と目的地が同一であるような周遊トリップも扱うことができる。また、時間を正規化し、その行動を一般的な状態遷移モデルで表すことで、動的計画法を含む様々な既往手法の適用や、時間割引率の影響の考察が可能となる。一方、この手法には、交通混雑によるリンク所要時間の変化を表現できないという短所がある。しかし、自動運転車両を最適に制御する交通管制システムの下では、例えば、需要が交通容量を超過する場合にはトリップを中止させることが

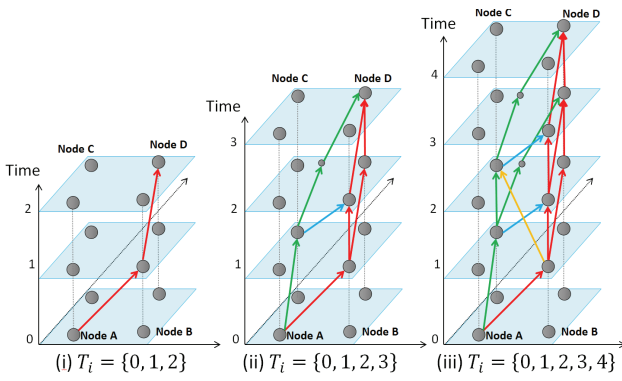


図-3 時空間プリズム制約を満たす行動軌跡

可能である。本研究では、このような技術の普及を見据えて、交通ネットワーク上に混雑が発生しないという条件の下での交通管制システムについて考える。また、本研究では「滞在」行動は全て交通ネットワーク上のノード $n \in N$ 上で発生するものとし、リンク途中での「滞在」は発生しないものとする。

b) 時空間プリズム制約

次に、時空間プリズム制約について考える。エージェント $i \in I$ は出発地 O_i 、目的地 D_i 、活動可能時間帯 $T_i = \{t_i^B, t_i^B + 1, \dots, t_i^E\} \subset T$ を有しているものとする。このとき、エージェント i は、時刻 t_i^B に出発地 O_i に、時刻 t_i^E に目的地 D_i に居る必要があるため、活動時間帯の間に実行可能な行動については、時空間的な制約が生じる。本研究では、Hagerstrand¹⁾ にならって、これを時空間プリズム制約と呼ぶ。

時空間プリズム制約を考慮したエージェントの行動軌跡の例を図-3 に示す。この例では、出発地がノード A、到着地がノード D であるエージェントについて、3 種類の異なる活動時間帯を与えた場合の実行可能な行動軌跡を図示したものである。活動時間帯が短い (i) のケースでは、エージェントが取り得る行動軌跡は 1 通りであるのに対して、活動時間帯が増加した (ii) および (iii) のケースでは、それぞれ 4 通り、9 通りの実行可能な行動軌跡が考えられる。

時空間プリズム制約をさらに強める条件として、各ノード $n \in N$ 上の施設が利用可能である時間帯 $b_n \subset T$ と、各エージェントの必須立ち寄り施設条件を考える。例えば、上述のサンプルケースで、 $b_B = \{1\}$ 、すなわちノード B 上の施設が時刻 $t = 1$ のときのみ利用可能であり、ノード A からノード D に向かうあるエージェントにとって施設 B への立ち寄りが必須であると仮定すると、活動時間帯が上述の (i) のケースではこの制約を満たす行動軌跡は存在せず、(ii) および (iii) のケースで実行可能な行動軌跡はそれぞれ 1 通り、2 通りに絞られる。

このように、出発地、目的地、活動可能時間帯に加えて、必須の滞在地点などの制約を持つエージェントが実行可能な行動軌跡は、時空間プリズム制約によって大きく制限される。本研究では、エージェント i の時空間プリズム制約を満たす軌跡の集合を L_i で表し、実行可能な一連な行動を $l_i = \{a_{i,t_i^B}, \dots, a_{i,t_i^E}\} \in L_i$ で表す。このとき、実行可能な軌跡の集合 L_i は、状態遷移グラフ G_i 上の疎な部分グラフの集合となる。

(2) 交通管理者モデル

エージェントは、利用者の出発可能時刻 $t = t_i^B$ において、利用者の移動に対する選好と制約、すなわち、報酬関数 R_i と時空間プリズム制約 L_i からなる情報 $\theta_i = \{R_i, L_i\}$ を交通管理者に通知する。時刻 t までの交通管理者に情報を通知したエージェントの集合を $I_t \subset I$ とし、これら全てのエージェントから通知された情報の集合を $\theta_t = \bigcup_{i \in I_t} \theta_i$ とする。また、本研究では、交通管理者は将来の需要に対する確率的な推定モデル \tilde{D} を有しているものとする。

交通管理者は逐次的に受け取る利用者の情報に基づいて、これらのエージェントの当該時刻の行動 $\pi_t = \bigcup_{i \in I_t} \pi_{i,t}$ を決定する。ただし、 $\pi_{i,t} \in A_i(s_{i,t})$ は時刻 t にエージェント i に割り当てられる行動である。交通管理者の各時刻 t における意思決定は、

$$\pi_{i,t} = \pi(\theta_t, s_t, \tilde{D}) \quad (1)$$

と表すことができる。なお、本研究では交通管理者は当該時刻の全エージェントの状態 s_t を把握しているものとする。この仮定は、全ての自動運転車両が交通管理者のコントロール下にあるという前提の上では自然な仮定である。

a) 目的関数

本研究では、交通管理者は総社会効用の最大化を目的とする。ここで、総社会効用とは、全てのエージェントが得る報酬の合計で定義されるものとする。このとき、制御対象時刻 T 全体に渡って全エージェントに行動 $\pi = \{\pi_0, \dots, \pi_T\}$ を割り当てた場合の事後的な総社会効用は、以下の式で表される。

$$SW(\pi) = \sum_{i \in I} \sum_{t \in T_i} R_i(s_t, \pi_{i,t}) = \sum_{t \in \bar{T}} R(\pi_t) \quad (2)$$

ただしここで、 $R(\pi_t) = \sum_{i \in I_t} R_i(s_t, \pi_{i,t})$ である。これに対して、交通管理者の逐次的な意思決定を考えた場合、時間割引率⁶⁾ を β とすると、時刻 t における割引総社会効用は、以下の式で表される。

$$DSW(\pi, t) = \sum_{t'=t}^{\bar{T}} \beta^{t'-t} R(\pi_{t'}) \quad (3)$$

b) 制約条件

交通管理者の意思決定においては、時空間プリズム制約とリンク容量制約の二つの制約条件が考慮される。前節に述べたとおり、各エージェント $i \in I$ は時空間プリズム制約を考慮した実行可能な行動軌跡の集合 L_i を有する。このため、交通管理者が時刻 t にエージェント i に割り当てる行動 $\pi_{i,t}$ は、実行可能軌跡 $l_i \in L_i$ の一部である必要がある。一方で、交通ネットワークの時空間的な容量を超過する需要が発生した場合には、全てのエージェントに対して時空間プリズム制約する行動を割り当てることが難しくなる。本研究では、そのような場合の対策として、トリップをキャンセルする、すなわち出発地点から動かないという選択肢 l_i^C を実行可能軌跡に追加して考える。この選択肢の追加によって、交通ネットワークの容量が逼迫するような条件下では、移動に対する効用が低いエージェントのトリップをキャンセルする一方で、一度出発したエージェントに対しては、活動時間帯の間に目的地に到着することを保証することが可能である。以上をまとめると、行動 $\pi_{i,t}$ は以下の制約条件を満たす必要がある。

$$\forall i \in I_t, \exists l_i \in \{L_i \cup l_i^C\}, \pi_{i,t} \in l_i \quad (4)$$

次に、リンク容量制約について考える。ある時刻 t におけるリンク容量制約は次の式で表される。

$$\exists \mathbf{l} \in \mathbf{L}(\boldsymbol{\pi}_t), \forall t' \geq t, \forall i \in I_t, \forall e \in E \quad (5)$$

$$\text{Count}(\mathbf{l}, t', e) \leq \text{Cap}(e)$$

ここで、 \mathbf{l} は、当該時刻に存在する全てのエージェント $i \in I_t$ の行動軌跡 l_i の組み合わせであり、 $\mathbf{L}(\boldsymbol{\pi}_t)$ は、当該時刻の全エージェントの行動の組み合わせを $\boldsymbol{\pi}_t$ とした場合の全エージェントの実行可能軌跡の組み合わせの集合である。 $\text{Count}(\mathbf{l}, t', e)$ は、全エージェントの行動軌跡の組み合わせが与えられた場合に、時刻 t' にリンク $e \in E$ を通行する車両の台数であり、 $\text{Cap}(e)$ は、リンク e のリンク容量である。すなわちこの制約式は、時刻 t に行動 $\boldsymbol{\pi}_t$ を決定した場合に、全エージェントの時空間プリズム制約と全リンクのリンク容量制約を同時に満たすような、全エージェントの将来の行動軌跡の組み合わせが存在することを示している。

以上のように、交通管理者は、式 4 で表される時空間プリズム制約と式 5 で表されるリンク容量制約を満たす制御を実行する。

c) 逐次的意思決定モデル

交通管理者は、将来の需要モデル \tilde{D} を考慮して、式 3 で示す割引総社会効用の期待値を最大化する交通制御を行う。すなわち、交通管理者の意思決定は、

$$\boldsymbol{\pi}_t = \underset{\boldsymbol{\pi}'_t = \{\boldsymbol{\pi}'_{t, \dots, \boldsymbol{\pi}'_{\bar{T}}\}}{\text{argmax}} \mathbf{E} \left[\sum_{t'=t}^{\bar{T}} \beta^{t'-t} R(\boldsymbol{\pi}'_{t'}) \right] \quad (6)$$

と表される。ただし、 $\mathbf{E}[\cdot]$ は期待値を表す演算子である。Bellman の最適性原理⁷⁾ を用いて逐次的な意思決定に分解すると、

$$\boldsymbol{\pi}_t = \underset{\boldsymbol{\pi}'_t \in \Gamma(\mathbf{s}_t)}{\text{argmax}} [R(\boldsymbol{\pi}'_t) + V(\boldsymbol{\pi}'_t)], \quad (7)$$

と表すことができる。ここで、 $\Gamma(\mathbf{s}_t)$ は状態 \mathbf{s}_t の下で実行可能な行動の組み合わせの集合であり、 $V(\boldsymbol{\pi}'_t)$ は、以下の式で与えられる価値関数である。

$$V(\boldsymbol{\pi}'_t) = \beta \cdot \mathbf{E} \left[\max_{\mathbf{a}'_{t+1} \in \Gamma(\mathcal{T}(\mathbf{s}_t, \boldsymbol{\pi}'_t))} V(\mathbf{a}'_{t+1}) \right]. \quad (8)$$

ただし、 $\mathcal{T}(\mathbf{s}_t, \boldsymbol{\pi}'_t)$ は、時刻 t に状態 \mathbf{s}_t の下で行動 $\boldsymbol{\pi}'_t$ を実行したときの次の時刻 $t+1$ における状態を表すものである。

交通管理者は、前節に示した時空間プリズム制約とリンク容量制約の下で、式 7 で表される最適化問題を逐次的に解くことで、総社会効用を最大化することができる。

3. 解法アルゴリズム

前章では、本研究で考える交通管制システムモデルを定式化した。本章では、前章で示した交通管理者の逐次的な意思決定を実現するための解法アルゴリズムを示す。

(1) アルゴリズムの全体像

まず、アルゴリズムの全体像を、Algorithm 1 に示す。

Algorithm 1 Main

Require: $(N, E), T, \tilde{D}, I, \theta$

Ensure: $\boldsymbol{\pi}, \mathbf{x}$

- 1: $t \leftarrow 0$
 - 2: $\mathcal{Z} \leftarrow \text{ENUMERATE}((N, E), T, \tilde{D})$
 - 3: **for** $t \in T$ **do**
 - 4: $I_t \leftarrow \text{RENEWAGENT}(I, t)$
 - 5: $\mathcal{Z} \leftarrow \text{UPDATE}(\mathcal{Z}, \theta_t)$
 - 6: $\theta_t \leftarrow \text{RENEWTYPE}(\theta, t)$
 - 7: $\mathbf{s}_t \leftarrow \text{GETSTATE}$
 - 8: $\boldsymbol{\pi}_t \leftarrow \text{EFFICIENTALLOCATION}(I_t, \theta_t, \mathbf{s}_t, \tilde{D}, \mathcal{Z})$
 - 9: $\mathcal{Z} \leftarrow \text{NARROWDOWN}(\mathcal{Z}, \boldsymbol{\pi}_t)$
 - 10: **end for**
-

初期状態におけるアルゴリズムへの入力は、交通ネットワーク (N, E) 、制御対象時間帯 $T = \{0, 1, \dots, \bar{T}\}$ 、および需要予測モデル \tilde{D} である。これに加えて、エージェントからの逐次的な通知を受けて、利用者の集合 I とその選好および制約に関する情報 θ が逐次的に更新される。

提案するアルゴリズムでは、制御対象時間帯 T に発生が予想される全ての実行プランを、Zero-suppressed binary decision diagram (ZDD)⁸⁾ を用いて事前に列挙し、 Z に保存する (2 行目). 列挙された実行プランは、エージェントからの新たな通知によって更新され (5 行目), また、各時刻に実際に実行されるプランの確定に伴って絞り込まれる (9 行目). 列挙された実行プラン Z に加えて、エージェントから通知された情報 (6 行目) と各時刻の交通ネットワークの状態 (7 行目) に基づいて、アルゴリズムは各エージェントの最適な行動を逐次的に決定する (8 行目).

以下では、実行プランを列挙する部分と最適行動を決定する部分について、それぞれ詳細に報告する.

(2) 実行プランの列挙

ここでは、Algorithm 1 に示した ENUMERATE, UPDATE および NARROWDOWN の各関数の詳細を示す.

関数 ENUMERATE この関数は、初期状態における入力情報を用いて予想される実行プランを列挙する事前処理を行う関数である. まず最初に、時空間プリズム制約の集合 $STP = \{STP_1, STP_2, \dots, STP_K\}$ を考える²⁾. 次に、式 4 を考慮して、それぞれの時空間プリズム制約 $STP_k \in STP$ を満たす全ての実行可能行動軌跡を生成する. その後、需要モデルから各時空間プリズム制約ごとにエージェント数 \tilde{N}_k を見積もり、その数の行動軌跡の集合を生成する. 最後に、これら全ての利用者の実行可能行動軌跡の組み合わせを考え、リンク容量制約式 5 を満足しない組み合わせを選択肢から除去する. このようにして、需要モデルに基づいて予想される各エージェントの実行プランを列挙する. 各々の時空間プリズム制約に対応する実行可能軌跡の集合は、時空間拡張ネットワーク上のグラフ G_i における疎なグラフの集合となっているため、行動軌跡の組み合わせや条件に合わない選択肢の除去は ZDD のライブラリ³⁾ を用いて高速に計算できる.

関数 UPDATE この関数は、エージェントからの実際の通知を受けて、列挙した実行プラン Z を更新するプログラムである. 各時空間プリズム制約 STP_k に対して、実際に通知をおこなったエージェント数 N_k が見積もり値 \tilde{N}_k よりも小さかった場合は、余剰の行動軌跡を削除する. 逆に見積もり値 \tilde{N}_k よりも大きかった場合には実行可能な行動軌跡を追加し、既に列挙した他のエージェントの行動軌跡との組み合わせのうち容量制約を満足するものを求め、 Z を更新する.

²⁾ この集合は、需要モデルに基づいて生成された典型的な時空間プリズム制約の集合であり、各要素 $STP_k \in STP$ は、例えば利用者の層別 (会社員, 学生, 主婦など) ごとの時空間プリズム制約を表す.

³⁾ <https://github.com/takemaru/graphillion/wiki>

関数 NARROWDOWN この関数は、交通管理者による意思決定を受けて、列挙した実行プラン Z を更新するプログラムである. 各時刻に各エージェントが実行する行動が決定すると、エージェントの実行可能軌跡の選択肢集合は縮小する. この関数は、決定された行動を含む実行可能軌跡のみを残し、その他の選択肢集合を削除する.

(3) 最適行動の決定

続いて、列挙された実行プラン Z を用いて、最適行動を決定するアルゴリズムの詳細を示す.

関数 EFFICIENTALLOCATION この関数は、利用者から通知された情報 θ_t , 現在の状態 s_t , 需要予測モデル \tilde{D} および列挙済みの実行プラン Z を入力とし、multi-scenario approach⁹⁾ を用いて、当該時刻における最適行動 π_t を決定する. アルゴリズムにおいては、まず、現在の状態 s_t に基づいて各エージェントが実行可能な行動の組み合わせ A_t を抽出する. 次に、需要予測モデル \tilde{D} に基づいて、将来のエージェントに関する M 通りのサンプルシナリオを作成し、それらを現在時刻までに既に存在するエージェントの情報と統合する. サンプルシナリオ毎に、実行可能なアクション $\pi_t \in A_t$ を実行した場合の割引総社会効用の最大値を求め、それらの値を平均することで、それぞれのアクション $\pi_t \in A_t$ に対する価値関数 $V(\pi_t)$ を求める. この値を用いて、アルゴリズムは、式 7 で表される意思決定を実行する.

4. 数値実験

本章では、前章までに示した提案手法を評価する数値実験について示す. 実験では、エリア内を周遊する観光客によるトリップと朝夕の通勤トリップが互いに干渉する観光地エリアの交通マネジメントを想定し、既存手法と比較した提案手法の有効性を示す.

(1) 実験条件

今回の実験では、制御対象時刻は $T = \{0, 1, \dots, 8\}$ の 8 タイムステップとし、割引率は $\beta = 1$ とした. 以下に、交通ネットワークおよびエージェントに関する実験条件を示す.

a) 交通ネットワーク

実験では、Fig. 1 に示すネットワークを考えた. このネットワークは、 $N = \{A, B, C, D\}$ の 4 ノードと、10 本の片方向リンクを有している. 図中のかっこで示す値はリンクコストである. 実験においては、ノード A を住居エリア、ノード D をオフィスエリアと考え、ノード B およびノード C は娯楽エリアと考えた. さらにノード B に位置する娯楽施設の営業時間は $b_B = \{3, 4, 5\}$

とした。なお、ノード B 以外の施設については営業時間を限定せず、 $\forall n \in N \setminus B; b_n = T$ とした。

b) エージェントの条件設定

エージェント $i \in I$ の報酬関数については、以下のように設定した。

$$\forall a_{i,t} \in A_i^M : R_i(s_t, a_{i,t}) = 0 \quad (9)$$

$$\forall a_{i,t} \in A_i^S : R_i(s_t, a_{i,t}) = \begin{cases} v_i^n & (t \in b_n) \\ 0, & (\text{otherwise}) \end{cases} \quad (10)$$

つまり、各エージェント i は「移動」行動においては報酬を得ず、各ノード $n \in N$ 上の施設の営業時間内に当該ノードに滞在したときに報酬 v_i^n を得るものとした。

今回の実験ではエージェントとして、滞在型エージェント I_C と通過型エージェント I_P を考えた。ただし、 $I_C \cup I_P = I$ である。観光トリップを想定した滞在型エージェントの目的は、エリア内を周遊し各地点における滞在を楽しむことである。一方で、通勤トリップを想定する通過型エージェントは、出発地と目的地の間を短い時間で移動することを目的としている。

以下、それぞれについて、今回の数値実験で設定したパラメータを示す。

滞在型エージェント 滞在型エージェントはノード A を出発してノード A に戻る周遊トリップを行うものとする。これは、ノード A に宿泊する観光客の観光行動を想定した設定である。滞在型エージェントの主目的は、ノード B に位置する施設を訪問することとする。すなわち、全ての滞在型エージェントは、トリップがキャンセルされた場合を除き、ノード B の施設の営業時間 $b_B = \{3, 4, 5\}$ の間に 1 時間ステップ以上ノード B に滞在する必要がある。

滞在型エージェントについては、各ノードで得る報酬を以下の正規分布に従って設定した。

$$\begin{aligned} \hat{v}_i^A &\sim \mathcal{N}(50, 50), & \hat{v}_i^B &\sim \mathcal{N}(200, 200), \\ \hat{v}_i^C &\sim \mathcal{N}(100, 100), & \hat{v}_i^D &\sim \mathcal{N}(50, 50) \end{aligned} \quad (11)$$

ただし、 $\mathcal{N}(\mu, \sigma^2)$ は、平均 μ 、分散 σ^2 の正規分布を表している。この報酬関数によると、滞在型エージェントは、ノード B およびノード C に滞在することで、ノード A やノード D に滞在するよりも大きな報酬を得ることができる。また、滞在型エージェント $i \in I_C$ の活動可能時間帯 $T_i = \{t_i^B, \dots, t_i^E\}$ については、活動開始時刻 t_i^B が一様分布 $\{0, 1, 2\}$ に従い、活動終了時刻 $t_{i_d}^E$ については、 $\forall i \in I_C; t_i^E = t_i^B + 6$ と設定した。

通過型エージェント 通過型エージェントについては、ノード A を出発してノード D に向かうエージェントと、逆にノード D を出発してノード A に向かうエージェントの二通りを考えた。それぞれのエージェントの集合を I_P^{AD} および I_P^{DA} で表す。ただし、 $I_P^{AD} \cup I_P^{DA} = I_P$

である。ノード A からノード D へ向かうエージェント $i \in I_P^{AD}$ については、各ノードで得る報酬を、以下の正規分布に従って設定した。

$$\begin{aligned} \hat{v}_i^A &\sim \mathcal{N}(200, 200), & \hat{v}_i^B &= 0, \\ \hat{v}_i^C &= 0, & \hat{v}_i^D &\sim \mathcal{N}(500, 500) \end{aligned} \quad (12)$$

また、逆にノード D からノード A へ向かうエージェント $i \in I_P^{DA}$ については、各ノードで得る報酬を、以下の正規分布に従って設定した。

$$\begin{aligned} \hat{v}_i^A &\sim \mathcal{N}(500, 500), & \hat{v}_i^B &= 0, \\ \hat{v}_i^C &= 0, & \hat{v}_i^D &\sim \mathcal{N}(200, 0) \end{aligned} \quad (13)$$

すなわち、通過型エージェントは早い時刻に目的地に到着することで大きな報酬を得る。通過型エージェント $i \in I_P$ の活動可能時間帯 $T_i = \{t_i^B, \dots, t_i^E\}$ については、ノード D へ向かうエージェント $i \in I_P^{AD}$ の活動開始時刻 t_i^B は一様分布 $\{0, 1, 2\}$ に、また、ノード A へ向かうエージェント $i \in I_P^{DA}$ の活動開始時刻 t_i^B は一様分布 $\{2, 3, 4\}$ に従うものとし、活動終了時刻 t_i^E については、 $\forall i \in I_P; t_i^E = t_i^B + 4$ と設定した。

なお、今回の実験においては、全てのエージェントについて、同一リンクの同方向への移動は活動時間帯の間に一度しか実行しないものとした。

エージェント間の干渉と交通管制 上述した設定では、滞在型エージェントは、ノード A から出発するときとノード A へ戻ってくるときに、それぞれノード D へ向かう通過型エージェントおよびノード A へ向かう通過型エージェントとの間で交通リソースの競合が発生するため、適切な交通管制が必要となる。本実験では、交通管理者は、本節に示したエージェントの需要に関する統計モデルを事前情報として有しているものとした。ただし、各エージェントの出発時刻および報酬に関する確定値については、通知を受けるまで未知とする。

c) ベンチマークメカニズム

数値実験のベンチマークメカニズムとしては、オフライン最適と早い者勝ちの両メカニズムに加えて、提案手法において需要予測モデル \hat{D} を用いない場合 (model free) を採用した。オフライン最適メカニズムは、交通管理者が将来の全てのエージェントの情報を事前に正確に知っていると仮定するものであり、このメカニズムで得られる総社会効用は、与えられた実験条件においてメカニズムが達成可能な総社会効用の上限値を表す。

早い者勝ち (first-come first-serve; FCFS) メカニズムは、本研究のような状況、で一度出発したエージェントが活動可能時間帯の間に目的地に到着することを保証するために、一般的に用いられるメカニズムである。交通管理者はエージェントの希望に応じてその時点で最も報酬が大きい行動軌跡を確定する。一度確定した

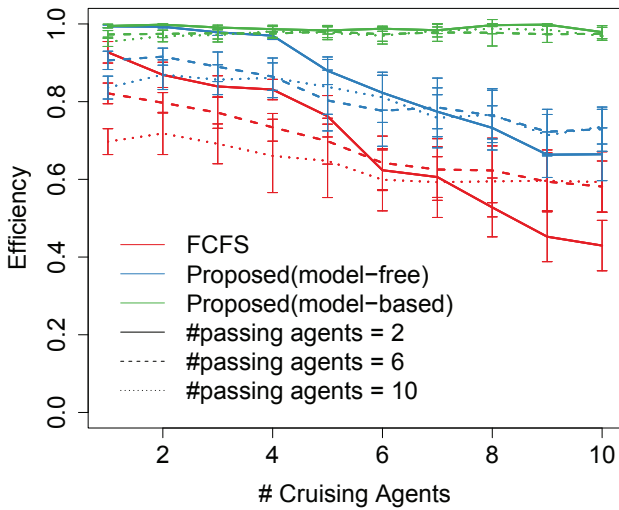


図-4 メカニズムの効率性評価

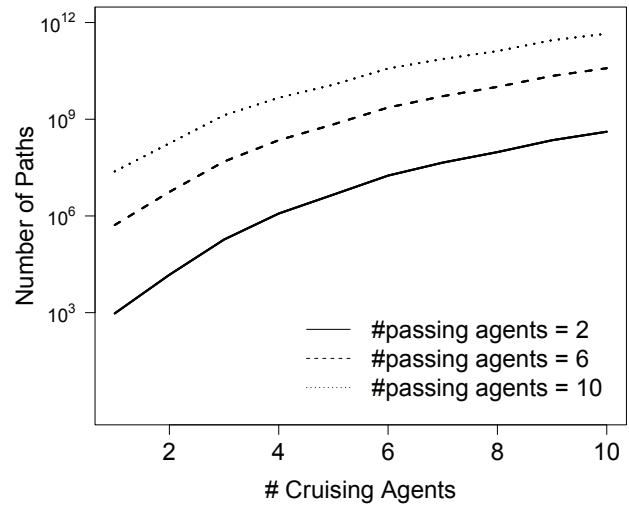


図-5 実行可能選択枝集合の数

行動は、その後通知を行ったエージェントの状況に関わらずに実行される。

提案手法は、3章に示したアルゴリズムによって割引総社会効用を最大化するものであり、アルゴリズム中で発生させるシナリオの数は、 $M = 10$ とした。

(2) 実験結果

まず、提案手法と早い者勝ちメカニズムの効率性を評価した。ここで、効率性とは、それぞれのメカニズムで得られた総社会効用の値をオフライン最適で得られる総社会効用で正規化したものである。実験では、通過型エージェントの数を2, 4, 6の三通りに設定し、それぞれのケースにおいて、滞在型エージェントの数を1から10まで変化させて実験を行った。実験は乱数を変化させながら、各条件における計算時間に応じて10回から100回試行した。

実験結果を図-4に示す。図中のプロットは各試行の平均値を示し、エラーバーは95%信頼区間を示している。図から分かるように、エージェント数が増加すると、早い者勝ちの効率性は大きく低下する。これに対して、提案手法では需要予測モデル \hat{D} を用いない model free の場合であっても早い者勝ちよりも良い効率性を示しており、需要予測モデル \hat{D} の情報を活用した場合にはオフライン最適に近い効率性を維持している。提案手法では、予約を受け付けた全てのエージェントに対して時空間プリズム制約の範囲内のサービスレベルを保証しつつも、後から来た報酬が大きいエージェントを優先的に取り扱うため、高い効率性を維持できていることが分かる。

ここで、各エージェントの時空間制約式4および容量制約式5を満たす実行可能解の総数を、図-5に示す。

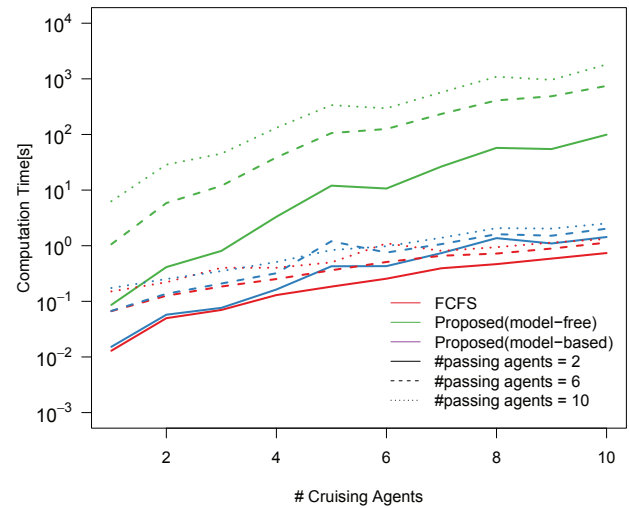


図-6 提案メカニズムの計算時間

図に示すとおり、今回の条件下で考える実行可能解の総数は最大で約1兆通りに達する。提案手法 (model-free) においては、ZDDを用いてこの全ての可能性を列挙して最適解を導出しており、このことが高い効率性の維持につながっていると考えられる。

次に、提案手法の計算時間を図-6に示す。将来のエージェントを考慮しないFCFSおよび提案手法 (model-free) の両メカニズムの計算時間と比較すると、提案手法 (model-based) の計算時間は大きい。しかし提案手法 (model-based) は、上述した全ての実行可能解を考えた上で、Algorithm 1に示す各行動プランごと、各サンプルシナリオ毎の多数の最適化計算を含んでいることを考えると、計算時間は許容範囲であるとも考えられる。これは、列挙した疎グラフに対する最適化処理を高速に実行可能なZDDの特徴を活かしたものである。

5. まとめ

本研究では、有限な交通資源の容量制約を考慮して総社会効用を最大化する交通配分を実現する動的な交通管制アルゴリズムを提案した。まず、交通ネットワーク上のエージェントモデルを正規化時間を持つ時空間拡張ネットワーク上での状態遷移モデルで表した。この時空間拡張ネットワークにおいてエージェントが取る行動は、出発地から目的地までの軌跡で表され、取り得る行動軌跡の選択肢集合は、出発地、目的地、活動可能時間帯および施設への立ち寄り条件などによって設定される時空間プリズム制約を満たす必要がある。交通管理者は、それぞれの時空間プリズム制約によって制限されたエージェントの選択肢集合に、交通容量制約を加味して、割引総社会効用が最大となる行動プランを逐次的に実行する。そのために、交通管理者は事前に需要モデルに基づいて現在時刻から制御対象時刻の終端までの間に起こり得る行動を列挙し、各時刻に実際に得られた情報に基づいて列挙された行動プランを更新しながら最適な行動を決定する。この一連の計算は、ZDD によって効率的に実行できる。

今後の方針としては、まず、数値実験によるパラメータスタディを充実させ、提案手法の有効性の更なる検証を行う必要がある。特に、今回数値実験では深く検討しなかった、報酬関数のばらつきの大きさや時間割引率の影響については、十分に評価する必要がある。

また、今回の研究では、エージェントが交通管理者に正確な情報を通知することを仮定したが、実際のサービスにおいては、エージェントは利己的な振る舞いを行い、場合によっては交通管理者に正しくない情報を通知することが想定される。このため、利己的なエージェントが交通管理者に進んで正確な情報を通知するようなメカニズムデザイン¹⁰⁾が必要である。

一般的に、動的な環境下においてこのような交通管制システムを考える場合には、早い者勝ちメカニズムが用いられることが多い。早い者勝ちメカニズムにおいては、早い時間帯に予約する利用者は、それ以後に予約する利用者に対して完全な優先権を保有する。これに対して Hayakawa ら^{11),12)}は、EV 充電の予約システムを対象として、早い時間帯に予約した利用者の利益を保証しつつ、遅い時間帯に予約した利用者の状況に応じて実行する制御を柔軟に変更する浮動型予約シ

ステムを提案している。交通サービスにおいても、同様の浮動型予約システムの設計が求められる。

上述の課題に加えて、これらの研究を実社会に実装するためには、利用者の報酬関数を推定するアルゴリズムの開発や、より高速に最適計算を実行するアルゴリズムの開発が求められる。

参考文献

- 1) Hägerstrand, T.: What about people in regional science?, *Papers in regional science*, Vol.24, No.1, pp.7–24, 1970.
- 2) Kitamura, R., Pas, E. I., Lula, C. V., Lawton, T. K., and Benson, P. E.: The sequenced activity mobility simulator (sams): an integrated approach to modeling transportation, land use and air quality, *Transportation*, Vol.23, No.3, pp.267–291, 1996.
- 3) Nie, Y., Zhang, H., and Lee, D.-H.: Models and algorithms for the traffic assignment problem with link capacity constraints, *Transportation Research Part B: Methodological*, Vol.38, No.4, pp.285–312, 2004.
- 4) 赤松隆: 一般ネットワークにおけるボトルネック通行権取引制度, *土木学会論文集 D*, Vol.63, No.3, pp.287–301, 2007.
- 5) Akamatsu, T. and Wada, K.: Tradable network permits: A new scheme for the most efficient use of network capacity, *Transportation Research Part C: Emerging Technologies*, Vol.79, pp.178–195, 2017.
- 6) Rust, J.: Structural estimation of markov decision processes, *Handbook of econometrics*, Vol.4, pp.3081–3143, 1994.
- 7) Bellman, R.: *Dynamic programming*, Princeton University Press, 1957.
- 8) Minato, S.-i.: Zero-suppressed BDDs for set manipulation in combinatorial problems, *Proceedings of the 30th international Design Automation Conference*, pp. 272–277, ACM, 1993.
- 9) Chang, H. S., Givan, R., and Chong, E. K.: On-line scheduling via sampling, *Proceedings of the Fifth International Conference on Artificial Intelligence Planning Systems*, pp. 62–71, AAAI Press, 2000.
- 10) Bergemann, D. and Välimäki, J.: The dynamic pivot mechanism, *Econometrica*, Vol.78, No.2, pp.771–789, 2010.
- 11) Hayakawa, K., Gerding, E. H., Stein, S., and Shiga, T.: Online mechanisms for charging electric vehicles in settings with varying marginal electricity costs, *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 2610–2616, April 2015.
- 12) Hayakawa, K., Gerding, E. H., Stein, S., and Shiga, T.: Price-based online mechanisms for settings with uncertain future procurement costs and multi-unit demand, *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, (accepted), July 2018.

(2018. 4. 27 受付)

Sequential algorithm using graph enumeration for traffic assignment aiming to
maximize social welfare

Keiichiro HAYAKAWA and Eiji HATO