

交通信号機広域制御における マルチエージェントモデルを用いた 自己組織化オフセット制御手法の提案

藤森立^{1,a)} 大野啓介² 神埼陽平² 栗原聡¹

概要: 交通渋滞は時間損失, 経済的損失, 環境汚染などを引き起こす, 都市部における主要な問題の1つである. 我々は交通渋滞解決へのアプローチとして交通信号制御に焦点をあてた. 現行の信号制御システムは中央集権型のシステムであり, 拡張性, 頑健性, 即応性に欠くものであると指摘されている. 本研究ではこの課題を解決するため, 自律分散型のマルチエージェントモデルを用いたオフセット制御を行うことで, 交通状況に応じて動的かつ柔軟にグリーンウェーブ形成を行うシステムを考案した.

Self-Organizing Offset Control Method based on Multi-Agent Model for Wide Area Traffic Signal Control

1. はじめに

日本国では戦後, 経済成長に伴い交通量が著しく増加した. そして現在に至るまで非常に高い水準を維持している. それに伴って交通渋滞が都市部などで深刻な問題となった. 交通渋滞は大気汚染や交通事故, そして経済損失など様々な問題を引き起こす要因となることから, その解消は喫緊の課題である.

交通渋滞を解消するアプローチの一つとして, 信号機制御による交差点の交通流円滑化が挙げられる. 一般的な道路の場合, 交通渋滞の起点となるのは, 主に車両が密集する交差点である. 交差点での交通流を滞留させることなく制御するには, 信号機を制御するパラメータを適切に操作する必要がある. しかしながら, 通勤時間帯のラッシュ時と昼間では交通流が大きく異なるように, 交通流は時間の経過とともに大きく変化する. また, 突発的な事故や工事などでも交通流に変化が生じる. そのため, 信号機の制御パラメータを交通流の変化に応じて動的に変動させることが, 交通流の円滑化には欠かせない.

日本の現行の交通管制システムは中央の交通管制センターと端末の交通信号機が通信回線で結ばれているか否かによって二つに分けられる. 一つは非集中制御方式と呼び, あらかじめ設定した制御パラメータで動作する方式である. この方式では, 事前に測定した交通流パターンから算出したパラメータを設定するため, 測定時のパターンと同様の交通流パターンにおいては制御の効果が見込めるものの, 事故や災害などによる突発的なパターン変動に対応することは困難である. もう一つは集中制御方式と呼び, 交通道路に設置された感知器等から得られる交通情報をもとに, 交通管制センターで制御パラメータを算出する. このパラメータは通信回線を通じて端末の交通信号機に伝えられる. 日本においては集中制御方式の一つであるMODERATO[1]が主要交差点の信号機制御に導入されている. 集中制御方式には国際的にいくつかの制御方式が存在するが, 基本的には事前に設定されたエリア毎に, 道路上に設置された感知器による感知情報やプローブ情報を一定の時間間隔で管制センターに収集し, 一括でパラメータを推定している. したがって交通状況に合わせた制御により非集中制御方式に比べて渋滞緩和の効果は大きくなることが期待できる.

一方で, 管制センターによる集中制御であるため, 交通情報を信号制御に反映するまでに5~30分前後の時間が

¹ 慶應義塾大学理工学部
Faculty of Science and Technology, Keio University
² 慶應義塾大学理工学研究科
Graduate School of Science and Technology, Keio University
a) ryufmr@keio.jp

かかり、動的に変化する個々の交差点の交通状況への即応性は低くなる。また制御エリアに関する事前設計が必要でありスケール性が低い。また事前設計においてシミュレーションから数種類のパラメータ集合を作成し、運用では感知器などの情報からパラメータを選択する手法であるためリアルタイムの情報をあまり活用できていないといった問題点がある。

近年、現行の交通管制システムのような中央集権型のシステムに対して、即応性、スケール性の高い自律分散型の信号機制御システムが提案されている。これは個々の信号機が局所的な情報を用いて自律的に信号機制御パラメータを決定するシステムである。我々は交通流の円滑化を目的として、自律分散型交通信号機制御システムの構築を行ってきた。本論では、白井ら [2] の提案したセルオートマトンを用いた交通シミュレーション上でのオフセット制御システムを、より実環境に近いシミュレーションにおいても運用できるよう拡張し、また機能についてもさらなる拡張を行ったものを提案する。提案システムは、各交差点にその信号制御を担当するエージェントを配し、隣接するエージェントが連携するマルチエージェントモデルによるオフセット制御を行うものである。また、今回はシミュレーション上で基本的な有効性の検証を行った。

2. 着目する信号制御パラメータ

交差点では複雑な交通流が形成される。また、交差点は交通事故や交通渋滞が発生しやすい場所である。信号機はこの交通流を円滑に対処することを目的として設置される。一般的に信号機の制御は、現示の決定と制御パラメータの設定の2段階によって行われ、4枝交差点では2現示方式がしばしば採用される。また右折専用車線が設けられている場合、右折専用信号による現示が追加される場合もある。信号機制御パラメータは現示を切り替えるタイミングを決定する。制御パラメータは、サイクル長(信号1周期の長さ)、スプリット(1サイクル中で各現示に与えられる時間の比率)、オフセット(近隣交差点の信号機間における、青信号開始時刻のずれ)の3種類が使用される。なおオフセットは、ある共通な基準時間からのずれを絶対オフセット、隣接信号機間におけるずれを相対オフセットという。信号機制御は大きく2つの方向、スプリット制御とオフセット制御に分けることができる。スプリット制御は個々の信号機の現示の長さを制御し、オフセット制御は隣接する信号機同士での青信号が開始されるタイミングを制御するものである。特にオフセットの制御については複数の交差点が関わるパラメータであるため、制御に関わる交差点集合が与えられ、サイクル長を共通にするなどの制御が行われる。この交差点集合をサブエリアと呼ぶ。本研究では上記 MODERATO において自律制御化が遅れている後者のオフセット制御およびサブエリアの設定に着目する。

3. 関連研究

自律分散型の信号制御システムとは一般に分散知能を踏襲し、交差点単位で自律した信号制御を行うシステムのことを指す。このようなシステムは、交差点ごとに分散制御ユニットと必要なセンサ・通信設備を設置するだけで運用可能であり、それによって複雑かつ高コストなパラメータ設計・保守の作業を大幅に省くことができる。加えてシステムを統括する分散制御中央装置は従来の中央制御装置と比較して極め小規模かつ低コストであり、また装置故障によるリスクの分散化もできるなど様々な利点が指摘されている [3]。

自律分散型の信号制御手法には様々なアプローチが見られる。ヒューリスティックなアルゴリズムによる自律分散制御の例として、高橋ら [4] は周辺情報から物理的ばねモデルをベースにしてスプリットのとオフセットの二つを決定する手法を提案している。また、本研究におけるベースとなるシステムとして、白井ら [2] は各エージェントが自身と周辺交差点の情報を用いてルールベースにオフセットを決定を行う手法を提案している。

また、交差点を主体とした自律分散型信号制御手法の研究において近年注目されているのが強化学習を用いた手法である。交通状況に即した制御を強化学習によって自動的に獲得することが期待されている。このような手法は以前から Wiering [5] や Kuyler ら [6] などによって提案されていたが、特に近年、DQN [7] に代表される深層強化学習が登場した事により、信号制御がその応用先として高い注目を集めている。深層強化学習を応用した自律分散型信号制御システムの研究例として、Pol ら [8][9] や Wei ら [10] は転移学習による制御領域の拡大を提案し、Chu ら [11] は、マルチエージェントによる深層強化学習が広域制御に一定の効果があることを示した。

こうした強化学習による手法は交通状況に即応して現示切り替えを行うため、交通集中の解消に高い効果をもつことが示されているが、実際の導入にはいくつかの問題点がある。まずエージェントの行動が現示を切り替えるものであり、高速な現示の切り替えが起こりうるため安全性に懸念がある。学習と運用にはリアルタイムな交差点周辺の画像のような高度な情報が必要なため導入コストが高い。そして現状、強化学習による手法はスプリットを制御するものである。狙ってオフセットを制御することはできていない。

本論における提案システムは、上述のような学習は行わないものの、交通量という比較的単純な情報を用いてエージェントが互いに連携することで、大域的な組織を構成し、従来手法では適用領域が限定される系統制御手法を動的に実施することを狙ったシステムを提案する。

4. 提案手法

本研究では、自律分散型の信号制御を行うため、MODERATOにおいても限定的に用いられているグリーンウェーブという制御手法を応用した、以下のようなマルチエージェントモデルによる制御を提案する。

各交差点に設置されるエージェントは、自身が管理する交差点に流入する車両数に関する情報と、隣接する交差点に設置されるエージェントから得られる局所的な情報のみを用いて互いに協調を行い、それぞれが自律的に信号機制御パラメータを調整する。本研究ではオフセット制御に着目し、全ての信号機のサイクル長は共通の長さであるとする。

連続する交差点群において、オフセットを調整することで、その区間を走行する車が、赤信号によって停止すること無く走行できるようにする制御手法をグリーンウェーブと呼ぶ。現行のMODERATOではグリーンウェーブが実施されるサブエリアは事前に設定されている。事前のシミュレーションにより算出された状況と、実際の交通状況とを比較し、サブエリアの選定とオフセット値を決定している。人手による設定であることから、幹線道路や主要交差点以外はサポートされにくいという現状がある。本研究の提案手法ではエージェントが自律的に協調することで組織を形成し、組織内でグリーンウェーブを実施する。

4.1 エージェントの構成

本節では単一エージェントが持つ基本的な機能や取り扱い情報について説明する。

エージェントは自身の担当交差点への交通量を観測するセンサ、隣接エージェントと情報をやり取りする通信機能、それらの情報から制御方針を決定する行動決定機能を備える。

エージェントは常に、次の5つのうちいずれかの状態(モード)をとる。

独立エージェント 協調を行わないモード。

起点 (Root) エージェント 組織形成の起点となるモード。

従属 (Sub) エージェント Root エージェントに従い、組織の構成員となるモード。

組織解消起点 (ClearanceRoot) モード 組織解消を行う際、Root エージェントがこのモードになる。

組織解消従属 (ClearanceSub) モード 組織解消を行う際、Sub エージェントがこのモードになる。

ここで、交差点 i と j が隣接していると仮定する。交差点 i のエージェントを A_i で示す。以降、説明の主体を A_i とする。 A_i は事前に表 1 の情報を与えられる。

表 1 エージェントの事前情報

データ	記号
信号機サイクル長	c
スプリット	s
隣接交差点 j との距離	$l_{(i,j)}$
隣接交差点 j との間の道路の系統速度	$v_{(i,j)}$

A_i はセンサから表 2 の情報を取得することができる。なお、車両数はサイクル毎の値とする。

表 2 センサからの取得情報

データ	記号
$i \rightarrow j$ 方向への車両数	$p_{(i,j)}$
$j \rightarrow i$ 方向への車両数	$p_{(j,i)}$
交差点 i への車両流入総数	P_i
現在時刻	t

A_i は A_j との通信で表 3 の情報を取得できる。ただし、 A_j が Root エージェントに従っているとき、その Root エージェントを A_R 、また A_R の担当交差点を R で表すとす。

表 3 隣接エージェントから取得可能な情報

データ	記号
$j \rightarrow i$ 方向への車両数	$q_{(j,i)}$
$i \rightarrow j$ 方向への車両数	$q_{(i,j)}$
交差点 j への車両流入総数	P_j
交差点 j の状態	$state_j$
交差点 R の車両流入総数	P_R
交差点 j の次サイクル開始時刻	T_j
交差点 R の次サイクル開始時刻	T_R

4.2 協調機能

本節では、エージェントが交渉と協調を行い、グリーンウェーブを形成するまでの流れを説明する。

まず、道路ネットワーク中の全てのエージェントの初期状態は独立モードである。各エージェントはサイクル毎に、取得した情報から自身の Root への状態変更判定と隣接エージェントとの交渉による協調可否の判定を行う。

この節でも主体となるエージェントは A_i とし、その隣接エージェントは A_j とする。

4.2.1 起点発生

独立エージェント、あるいは Sub エージェントはサイクル毎に自身が Root になるか否かを判定する。独立エージェントの場合は式 (1) を、Sub エージェントの場合は式 (2) を満たすとき Root エージェントへと変化する。なお、(2) における P_R は A_i が属する組織の Root エージェント担当交差点への車両流入数を指す。

$$P_i > \alpha \quad (1)$$

$$P_i > P_R \quad (2)$$

α は事前に設定される定数である。今回実験に用いた値については後述する。

4.2.2 交渉

あるエージェント A_i が Root エージェントに変化したとき、 A_i は隣接エージェント A_j に対して協調提案を行う。式 (3) が満たされている場合、提案が A_j に渡され、 A_j はその受入可否を判定する。

$$(p(i, j) > \beta) \vee (q(j, i) > \beta) \quad (3)$$

β は事前に設定される定数である。今回実験に用いた値については後述する。

A_j が独立モードのときは無条件に提案を受け入れる。 A_j が Sub エージェントのとき、 A_j が属する組織の Root エージェントを A_{jR} として、式 (4) を満たすとき提案を受け入れる。 A_j が Root エージェントのとき、式 (5) を満たすとき提案を受け入れる。

$$P_i > P_{jR} \quad (4)$$

$$P_i > P_j \quad (5)$$

提案を受け入れない場合、 A_j に変化はない。提案を受け入れた場合、 A_j は A_i を親とする Sub エージェントに変化する。このとき二者間で相対オフセットが算出される。算出方法については後述する。

4.2.3 伝播

A_j が Sub エージェントに変化した時、Root エージェントのときと同様に、 A_j も親以外の隣接エージェントに対して協調提案を行う。ここで、 A_j 属する組織の親を A_{jR} とし、 A_j の親ではない隣接エージェントを A_k とする。式 (6) が満たされている場合、提案が A_k に渡され、 A_k はその受入可否を判定する。

$$(p(j, k) > \beta) \vee (q(k, j) > \beta) \quad (6)$$

A_k が独立モードのときは無条件に提案を受け入れる。 A_k が Sub エージェントのとき、 A_k が属する組織の Root エージェントを A_{kR} として、式 (7) を満たすとき提案を受け入れる。 A_k が Root エージェントのとき、式 (8) を満たすとき提案を受け入れる。

$$P_{jR} > P_{kR} \quad (7)$$

$$P_{jR} > P_k \quad (8)$$

提案を受け入れない場合、 A_k に変化はない。提案を受け入れた場合、 A_k は A_j を親とする Sub エージェントに変化する。このときも、二者間で相対オフセットが算出さ

れる。以降、提案が完全に受け入れられなくなるまで協調提案は伝播し、起点となったエージェントを中心に、提案を受け入れた Sub エージェントたちによって組織が構成される。この組織内でグリーンウェーブを実施する。

4.3 オフセット算出

本節ではオフセットの算出方法について述べる。まず協調する隣接交差点間で相対オフセットが算出される。その後、組織の起点のサイクル開始時刻 T_R を基準に相対オフセットは絶対オフセットに換算され、導入される。

隣接交差点 i, j 間の双方向交通の車両数 $p(i, j), q(j, i)$ のうち、大きい方を p_l 、小さい方を q_s とする。 p_l 側の系統速度を v_l とする。

$$r = \frac{p_l}{q_s} \quad (9)$$

式 (9) で求まる双方向交通量比 r としきい値 $\gamma, \delta (\gamma > \delta \geq 1)$ を用いて、式 (10) によって相対オフセットの絶対値 O_r を算出する。

$$O_r = \begin{cases} \frac{l(i, j)}{v_l} & (r \geq \gamma) \\ \frac{l(i, j)}{v_l} \frac{r - \delta}{\gamma - \delta} & (1 \geq r \geq \gamma) \\ 0 & (r \geq 1) \end{cases} \quad (10)$$

そして、式 (10) で求めた値から、 A_i から A_j に対する相対オフセット $O_{(i, j)}$ を式 (11) で算出する。

$$O_{(i, j)} = \begin{cases} O_r & (p(i, j) > q(j, i)) \\ -O_r & (p(i, j) \leq q(j, i)) \end{cases} \quad (11)$$

なお、現行の MODERATO では、オフセットの変動は 1 サイクルにつきサイクル長の 25% までという制限があるため、本研究における実験ではこの制限を守り、それ以上の変動が起こる場合には数サイクルを要して行った。

4.4 組織解消

本節では、一度形成された組織を解消する機能について述べる。この機能においては、組織の Root エージェントは直近数サイクルの自身への車両総流入数を記録するものとする。Root エージェントは記録している車両総流入数の平均値 \bar{P}_i をもとに、式 (12) を満たすとき、交通集中が緩和されたと判断し、組織の解消を実施する。

$$\bar{P}_i < \theta \quad (12)$$

なお、 θ は事前に設定される定数である。今回実験に用いた値については後述する。

組織を解消するにあたって、Root エージェントは ClearanceRoot エージェントに変化し、組織全体に対して組織

解消の命令を発する。この命令は伝播のときと同様にバケツリレー式で伝播する。この命令を受けて、Root エージェントの組織に属する Sub エージェントは ClearanceSub エージェントに変化する。ClearanceSub エージェントは ClearanceRoot エージェントを基準にしてオフセットを解消、その後、独立エージェントへと変化する。ただし、ClearanceRoot エージェント、ClearanceSub エージェントともに、独立エージェント同様に、Root への変更判定を行い、協調提案は無条件に受け入れる。

5. 評価実験

本研究における提案手法の有効性を検討する実験として、次の3つの手法に関して実験を行った。

- 定周期制御 (比較対象)
- 提案手法+組織解消なし
- 提案手法+組織解消あり

本論の提案手法はオフセット制御のみの手法であるから、サイクル長とスプリットに関してはどの手法でも共通の値を使用した。定周期制御においては全ての交差点の信号が同位相とした。またオフセット導入は、サイクルの最後に全て赤信号の時間を設けることで導入した。これは、青信号を含む現示を用いてオフセットを導入することで、スプリット制御の効果が意図せず現れ、提案手法が有利になることを防ぐためである。

5.1 交通シミュレータ

本研究では実験において、マイクロ交通流シミュレータ「Simulation of Urban MObility」[12][13]を用いた。「Simulation of Urban MObility」はドイツ航空宇宙センターが中心となって開発しているオープンソースのシミュレータであり、通称「SUMO」と呼ばれている。以下、本シミュレータを SUMO と表記する。

SUMO では、道路ネットワークに対して外部から供給される車両の台数、どこから流入するかという確率、交差点での直進右左折の確率を設定することができる。

5.2 使用パラメータ

式(1), (3), (10), (12)などで用いたパラメータ $\alpha, \beta, \gamma, \delta, \theta$ の値について述べる。

α, β は交差点の交通容量をもとに設計した。交差点の交通容量 Cap_i とは、ここでは信号1サイクルの間に通過することができる最大車両数とする。 α, β はそれぞれ Cap_i の 0.4 倍、0.2 倍とし、実数値は 20, 7 である。 θ は α の 8 割とした。 γ, δ についてはそれぞれ 1.5, 1.1 とした。

5.3 実験環境

今回の実験では 3x12 の道路ネットワークを用いた。東西方向を主要道とし、交差点では直進しやすいというシナ

リオとした。

車両流入確率に関しては、横3本の中央にある道路の東西それぞれが 0.16、残りの2本に関しては東西それぞれが 0.12。南北方向の道路に関してはすべて南北それぞれ 0.0085 とした。交差点での進行方向の確率については、直進が 0.8。右左折はともに 0.1 とした。これを示したのが図1である。

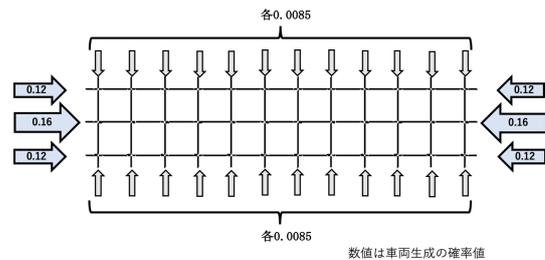


図1 車両流入イメージ

このような条件で 10 万台の車両の走行ルートを作成し実験を行った。すべての手法の実験において、車両は同じルートを通じた。

5.4 実験結果

本実験における評価は、各車両の信号によって生じた遅れ時間をもって行った。まず、到着台数あたりの累積遅れ時間を示したものが図2である。

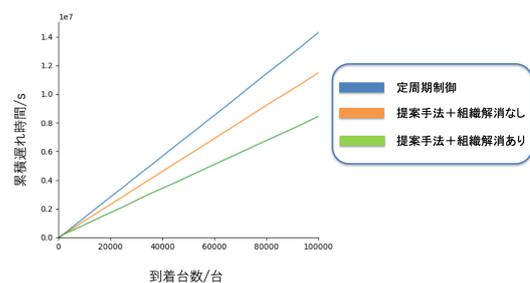


図2 累積遅れ時間

図2から提案手法において、累積遅れ時間の傾きが小さい事がわかる。また、累積遅れ時間において、組織解消ありの手法の方が、より傾きは小さい。

次に、車両1台ずつの遅れ時間の分布を示したものが図3である。

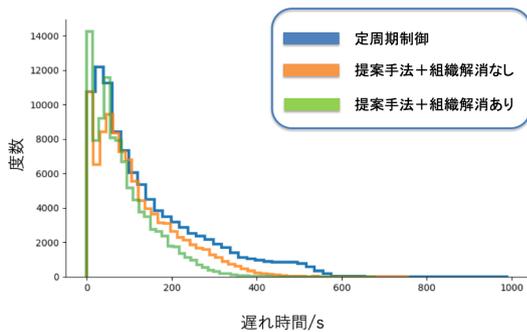


図 3 遅れ時間分布

図 3 から、提案手法において、遅れ時間分布の裾が軽くなっていることが分かる。また、組織解消ありの手法については遅れ時間 0 付近が著しく増加しているということも分かる。

最後に、遅れ時間の分布に関する統計量を表 4 に示す。

表 4 遅れ時間の分布の統計量

手法	最小値	最大値	平均値	中央値
定周期制御	0	992	143.20	100
組織解消なし	0	758	114.99	89
組織解消あり	0	679	84.62	66

表 4 から、提案手法によって最悪遅れ時間が小さくなり、平均値と中央値からも遅れ時間の分布がより小さい値の方へと移動したことが分かる。また統計量からも、組織解消ありの方がなしに比べて遅れ時間軽減の効果が高いことが分かる。

以上のことから、提案手法には遅れ時間軽減に一定の効果があると考えられる。また、組織形成をより柔軟にするために組織解消を行うことで、その制御効果はより高まるものであると考えられる。

6. 結論

本研究では渋滞問題に対するアプローチとして信号制御に着目し、交通流の円滑化を目指した自律分散型の信号制御システムという観点から、マルチエージェントシステムを用いた自己組織化信号制御システムを提案した。

本提案システムは、信号制御の中でも特にオフセット制御に主眼を置き、現行のシステムにおいても限定的に活用されているグリーンウェーブと呼ばれるオフセット制御手法を用いている。局所的な交通状況を監視する交差点が、周囲の交差点と交渉・協調を行うことで組織化し、その組織内でグリーンウェーブを形成することで、従来システムと比較してより動的かつ柔軟に実現することを目指したものである。

そして本研究では、手法の提案とともにシミュレーションを用いた提案手法の有効性の検討を行った。そして実験

から、車両の累積遅れ時間の減少を確認した。また、1台ずつの遅れ時間の分布から、分布全体についての遅れ時間の減少効果と、最悪の場合の遅れ時間の大幅な減少効果を確認した。加えて、Root エージェントが継続して交通状況を監視し、協調制御の必要がなくなった場合においては協調解除を判断し、形成された組織を解消することによって、新たな組織形成が迅速かつ柔軟に行えるようになり、さらなる遅れ時間減少の効果が期待できるということが分かった。

一方で、ネットワーク全体における主要な方向が存在しない場合やある特定の閉路の交通量が多い場合については有効な組織形成を行うことができるかについての考察が必要である。加えて、通勤ラッシュと帰宅ラッシュのように、時間帯の変化に伴って交通流のトレンドが変動するような状況についても現在評価実験を検討している。

また、オフセット算出式は方向別交通流に偏りがある場合についてのみオフセットが発生するような式となっているが、双方向が均衡を保つような場合については、担当エージェントが組織内に含まれるもののオフセット制御を実施しない状況が発生する。そのような状況においてどのようなオフセット制御が好ましいかという検討が必要である。協調において用いている判定式等も単純な大小比較によるルールベースであるから、組織化におけるルールの設計については今後も考察・検討を加えていく所存である。

参考文献

- [1] Sakakibara, H., Usami, T., Itakura, S. and Tajima, T.: MODERATO (Management by Origin-DEstination Related Adaptation for Traffic Optimization), 電気学会研究会資料. RTA, 道路交通研究会, Vol. 1999, No. 26, pp. 47-52 (1999).
- [2] 白井嵩士, 矢野純史, 西村茂樹, 香川浩司, 森田哲郎, 沼尾正行, 栗原聡: マルチエージェントモデルによる信号機オフセット制御法の提案, 人工知能学会論文誌, Vol. 26, No. 2, pp. 324-329 (2011).
- [3] 竹内久治, 佐竹徳己, 井手幹生, 穂坂重孝: 分散型交通信号制御システム, 三菱重工技報, Vol. 40, No. 3, pp. 174-177 (2003).
- [4] 高橋光紀, 篠田孝祐, 諏訪博彦, 栗原聡ほか: マルチエージェントモデルによる信号制御シミュレーションの考察, 研究報告知能システム (ICS) (2014).
- [5] Wiering, M.: Multi-Agent RL for Traffic Light Control, *Physical Review D*, Vol. 96, No. 1 (2000).
- [6] Kuyer, L., Whiteson, S., Bakker, B. and Vlassis, N.: Multiagent reinforcement learning for Urban traffic control using coordination graphs, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 5211 LNAI, No. PART 1, pp. 656-671 (2008).
- [7] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D. and Riedmiller, M. A.: Playing Atari with Deep Reinforcement Learning (2013).
- [8] Pol, E. V. D.: Deep Reinforcement Learning for Coordination in Traffic Light Control, *Master thesis* (2016).
- [9] van der Pol, E. and Oliehoek, F. A.: Coordinated

- deep reinforcement learners for traffic light control, *30th Conference on Neural Information Processing Systems (NIPS 2016)* (2016).
- [10] Wei, H., Chen, C., Wu, K., Zheng, G., Yu, Z., Gayah, V. V. and Li, Z.: Deep Reinforcement Learning for Traffic Signal Control along Arterials (2019).
- [11] Chu, T., Wang, J., Codeca, L. and Li, Z.: Multi-Agent Deep Reinforcement Learning for Large-Scale Traffic Signal Control, *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–10 (2019).
- [12] Behrisch, M., Bieker, L., Erdmann, J. and Krajzewicz, D.: SUMO – Simulation of Urban MObility: An Overview, *SIMUL 2011*, ThinkMind (2011).
- [13] (DLR), G. A. C. and others.: Simulation of Urban MObility (2020). <https://sumo.dlr.de/docs/index.html>.