

## CO<sub>2</sub>削減を目的とした車両移動予測に基づく 適応型信号機制御

富 樫 祐 二<sup>†1</sup> 梅 津 高 朗<sup>†1</sup> 東 野 輝 夫<sup>†1</sup>

現在、自動車による排気ガス量の削減が求められている。排気ガス量は自動車の挙動によって大きく変化するため、信号機による交通制御が排気ガス量削減に効果的であると考えられる。しかし、既存の信号機制御の多くは停止時間削減を目的としており、信号機における停止時間と停止回数のトレードオフ性から、排気量削減に効果的である信号停止回数の削減に効果的でないという問題点がある。本研究では、車車間通信によって得られた情報から信号周辺の車両の挙動を予測し、停止回数を削減するような信号機切り替えのプランを算出する分散制御手法を提案する。提案手法では、予想排気量を評価値とした強化学習を行うことで排気量最適化に適応する。シミュレーション実験を通して本手法によって既存手法と遅れ時間をほぼ同程度としたまま排気ガス量の削減効果を確認した。

### A self-learning traffic signal control based on prediction of vehicle behavior for CO<sub>2</sub> reduction

YUJI TOGASHI,<sup>†1</sup> TAKAAKI UMEDU<sup>†1</sup>  
and TERUO HIGASHINO<sup>†1</sup>

Recently, reduction of vehicle emissions is required. Controlling road traffic by traffic signals is effective to reduce vehicle emissions because amount of vehicle emissions is strongly affected by behavior of vehicles. However, most of major existing traffic signal control systems intend to reduce vehicle delays, and these systems are not effective against reduction of vehicle emissions because of the trade-off between stop duration and stop frequency at signals. In this paper, we propose a decentralized signal control technique for reduction of vehicle stops using vehicle information collected by inter-vehicle communication. The proposing technique controls signals based on reinforcement learning technique to predict emission amount as a evaluation value. Through simulation based evaluation, we show that the proposing technique is efficient to reduce vehicle emissions, with the similar average travel time.

#### 主な用語

車車間通信 信号制御 リアルタイム制御 協調制御 排気ガス量削減 強化学習

#### 1. ま え が き

現在、温室効果ガスによる地球温暖化が大きな環境問題となっており、温室効果ガスの排出量の削減は早急の課題である。東京都で2006年度に行われた調査によると、温室効果ガスの発生原因における運輸部門の割合は26.2%に上り<sup>1)</sup>、そのほとんどが自動車の排気ガスによるものである。自動車による排気ガス量の推計方法については様々な研究がなされており、自動車の走行における挙動から排気量を求める手法<sup>2),3)</sup>などが提案されている。これらのモデルによると、自動車による排気ガス量は、加速時に発生するものが大きな割合を占めており、同一の距離を走行する場合でも、一定の速度で走行する場合と、停止と発進を何度も行う場合では排気量に大きな違いが出る事が分かっている<sup>4)</sup>。よって、各交通流に対してどのように通行権を与えるかを決定し、交通流上の車両の挙動を制御する交通信号制御の動作を改善することによって、発生する排気ガス量を削減できると考えられる。

しかし現在提案されている信号機制御手法は渋滞や遅れ時間の縮小を目的としたものが大半であり、排気ガス量の削減という観点から制御を行うシステムは少ないのが現状である。渋滞や遅れ時間の縮小を目的とした場合、信号機による車両の停止時間の削減が重要となるが、排気量最適化においては信号機による停止回数の削減が重要となる。しかし、文献<sup>5)</sup>で述べられているように、信号機による停止時間と停止回数にはトレードオフ性があり、停止時間の削減を行う既存の手法が停止回数の削減に効果的であるとは必ずしも言えない。代表的なミクロ制御であるRHODES<sup>6)</sup>は、評価関数を予想排気量とすることで排気量最適化に対しても対応できると考えられるが、対象エリア内の車両列の衝突を予測し、全ての衝突の回避方法の組み合わせに対する評価を求めるため、広域のエリアでは計算量が爆発的に増加し、適用が困難であるという問題点がある。また信号機と車両の相互通信が可能な環境を想定し、車両に対して適切な走行速度情報を送信することで停止回数と排気量を削減する手法が提案されている<sup>7),8)</sup>、これらの手法では車両の信号機への到着時刻の調整のために信号機間で十分な道路長が必要であり、都市街路のように道路長が短いエリアに対して適用で

<sup>†1</sup> 大阪大学 大学院情報科学研究科

Graduate School of Information Science and Technology, Osaka University

きないという問題点がある。

本稿では、文献<sup>10)</sup>の手法を基に、交通需要情報に加えて、リアルタイムに収集した車両位置情報と車両速度情報から算出した各車両の信号機への予想到着時刻情報を用いることで車両の細かな挙動を考慮した制御手法を提案する。提案手法では、車両予想到着時刻情報を入力としたニューラルネットワークによる強化学習で制御パラメータの探索を行い、ニューラルネットワークの学習に予想排気量を評価値としたバックプロパゲーションを用いることで排気量最適化に適用する。また車両情報の収集については車車間通信を用いることを想定しており、既存の手法の多くで想定されている車両感知器による交通情報の把握に比べて、情報の収集率が車両感知器の設置状況に依存せず、車種や速度といった車両情報がより正確に得られるという利点がある。

提案手法の評価のため、ミクロ交通シミュレータ VISSIM<sup>11)</sup>上に本手法を実装し、都市街路を想定したマップで、既存のリアルタイム信号機制御手法に対して交通量や車車間通信の通信環境を変化させた場合の平均排気量と平均旅行時間を比較するシミュレーション実験を行った。実験結果より、本手法による排気ガス量の削減効果を確認した。また信号機による遅れ時間についても比較対象と同程度の短さとなることを確認した。

## 2. 関連研究

### 2.1 自動車の挙動の排気量に対する影響

文献<sup>2)</sup>や文献<sup>3)</sup>は、実際の走行データから回帰分析によって車種ごとの係数を算出し、算出した係数と車両の速度や加速度を用いて排気ガス量を計算するモデルである。文献<sup>2)</sup>のモデルでは、車両の速度や加速度によって、車両の状態を4つの走行モードに分類し、それぞれの場合で個別の式を用いて排気量の算出を行う。文献<sup>3)</sup>のモデルは、次の式(1)によって排気量を求めるモデルである。

$$E_n(t) = Max\{E_0, f_1 + f_2v_n(t) + f_3v_n^2(t) + f_4a_n(t) + f_5a_n^2(t) + f_6v_n(t)a_n(t)\} \quad (1)$$

ただし、 $E_n(t)$ はある時刻  $t$  における排気ガス量、 $E_0$ と  $f_1 \sim f_6$ は係数、 $v_n(t)$ は時刻  $t$  の車両の走行速度、 $a_n(t)$ は時刻  $t$  の車両の加速度を表す。走行状態における排気量は、ガソリン車やディーゼル車、バスといった様々な車種ごとに速度や加速度の二次式としてモデル化した場合の係数があらかじめ調査されており、それぞれの値は表1に示した通りと示されている。このモデルで計算した場合、例えばガソリン車が平均的な加速度で速度 50[km/h]の状態から減速して停止し、それから加速して再び速度 50[km/h]の状態に戻すために発生する排気量は、10秒間の停止状態に発生する排気量よりも大きくなるため、自動車の停止

表1 車種ごとの係数

車種	$E_0$	$f_1$	$f_2$	$f_3$	$f_4$	$f_5$	$f_6$
ガソリン車	0	5.53e-1	1.61e-1	-2.89e-3	2.66e-1	5.11e-1	1.83e-1
ディーゼル車	0	3.24e-1	8.59e-2	4.96e-3	-5.86e-2	4.48e-1	2.30e-1
バス	0	9.04e-1	1.13	-4.27e-2	2.81	3.45	1.22

が排気量に与える影響は大きいことが分かる。

### 2.2 既存の信号機制御手法

リアルタイムに信号機制御を行う手法の代表的なものとしては、SCOOT<sup>12)</sup>・TUC<sup>13)</sup>・RHODES<sup>6)</sup>といった手法が挙げられる。SCOOT・TUCは道路上の車両感知機から得られる各道路リンク上のリアルタイムな交通量を基に信号機の制御パラメータであるサイクル長・スプリット・オフセットの値を調節する手法であり、実際の道路上における旅行時間の削減効果が評価されている。特にTUCは系統道路上を走行する車両に対して、信号機の青時間が連続する”Green Wave”と呼ばれる効果をオフセット制御によって実現しており、系統道路上における旅行時間の削減はSCOOTよりも大きい<sup>14)</sup>。TUCによるオフセット制御は、隣接する信号機間の互いの方向に流入する車両台数の比によって決定される。例えば、図1のように、信号機1から右方向に流入する  $v_1$  台の車両列と信号機2から左方向に流入する  $v_2$  台の車両列が存在するとき、信号機1と信号機2の間のオフセット  $offset_{12}$  は次の式で決定される。

$$W_{12} = \frac{v_1}{v_1 + v_2} \quad (2)$$

$$W_{21} = \frac{v_2}{v_1 + v_2} \quad (3)$$

$$offset_{12} = W_{12}t_{12} + W_{21}t_{21} \quad (4)$$

ただし  $t_{12}$ 、 $t_{21}$  はそれぞれ信号機1から信号機2方向への最適なオフセット、信号機2から信号機1方向への最適なオフセットであり、通常は平均的な速度で信号機間を走行した場合の所要時間である。よって、 $v_1 = 0$  ならば  $v_2$  台の車両列が信号機2を通過した後、停止することなく信号機1も通過することが可能となる。すなわち、車両列に対して連続した青時間の割り当てが行われる傾向にある制御だと言える。このように車両列の大きさに従って待ち時間の負担を分散させ、全体の旅行時間の削減を実現している。

### 2.3 強化学習による分散型信号機制御手法

適応型手法を信号機制御に用いる場合は、制御によって変化した交通状況から評価値を求め、求めた評価値によって制御で選択した行動を評価しフィードバックをかける動作となる

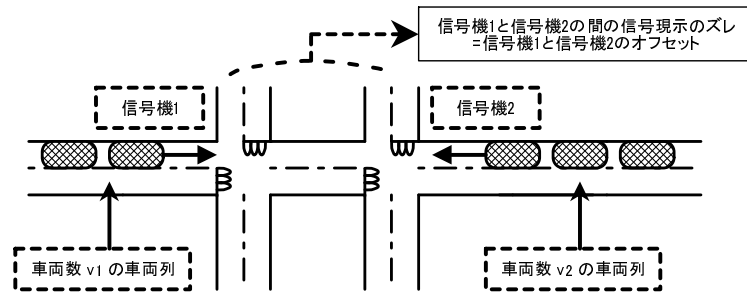


図1 TUCによる制御例

表2 信号機制御に用いる各パラメータ

$\delta$	制御サイクルを実行する時間間隔
$C_{min}$	1つのフェーズの長さの最小値
$C_{max}$	1つのフェーズの長さの最大値
PredictWindow	システムが何秒先の車両到着まで考慮するかという予想の範囲
$S_t$	システムの制御サイクル $t$ における状態
$a_t$	現フェーズ長さを $C_{min} + \delta \cdot i$ に設定する行動
$Q(S_t, a_i)$	状態 $S_t$ における行動 $a_i$ の評価の推定値

ため、教師なしの強化学習と呼ばれる手法を用いることが適切であると考えられる。一般的な教師なし学習では、システムの状態集合が有限であるならば、ある時刻  $t$  における状態  $S_t$  での行動  $a_t$  の評価値  $Q(S_t, a_t)$  は次の式によって更新される。ただし、 $r_{t+1}$  は行動選択後に発生したコスト、 $\alpha$ 、 $\gamma$  はそれぞれ慣性係数・減衰率と呼ばれる、強化学習の適用度合いを調整するパラメータである。

$$Q(S_t, a_t) \leftarrow Q(S_t, a_t) + \alpha[r_{t+1} + \gamma \min_a Q(s_{t+1}, a) - Q(s_t, a_t)] \quad (5)$$

ここで、無数の状態に対する厳密な関数  $Q$  を構築することは困難なため、内部パラメータを強化学習の手法を用いて調整する手法がしばしば用いられる。

### 3. 車両到着予想時刻情報を用いた機械学習による協調型信号機制御手法

提案手法はニューラルネットワークを用いた強化学習によって分散制御を実現している文献<sup>10)</sup>の手法を基に、リアルタイムに収集した車両情報から車両到着の分布を算出し、算出した車両到着の分布から信号機による停止回数を最小にするような制御を行う。

### 3.1 リアルタイム信号機制御における教師なし強化学習のモデル化

提案手法ではリアルタイム信号機制御を図2のような状態遷移モデルで表現する。ここで、各変数の意味は表2の通りである。ただし、信号現示の切り替えサイクルに合わせて制御を行うのでは制御を行うサイクルが長すぎて、時々刻々と変化する交通状況に合わせた制御が行えない可能性が高い。そこで、より短い時間の制御サイクルを設け、制御サイクルごとに集めた情報に基づいて信号の切り替えタイミングを調整することとした。また、制御サイクル間に増加した車両停止や車両遅れなどを制御行動によって発生した損失とみなし、各行動（ここでは信号切り替えタイミングの決定）によって発生する損失を推定し、損失が最小となるように信号を制御する。

提案手法では、式(5)で示したように発生した損失と遷移した先の状態の評価値の推定値の和を行動に対する評価値の目標値として教師なし強化学習を適用する。システムが制御対象とする領域の状態は図2のように信号交差点周辺の車両分布と現在の交通権割り当てが行われてからの経過時間で表現する。ここで、交差点の車両分布は、その時点で交通権が割当てられている道路を主道路、その他の主道路と交差する道路を副道路とし、主道路、副道路を交差点に向かって走行しているか停車している車両の交差点からの距離と交差点へ向かう速度で表すものとした。このようにすることで、交差点の形状によらず、また、信号現示の状態も考慮に入れた形で、交差点の車両分布を単純なデータにモデル化できる。

1つの方向に対して交通権を与える時間を1回のフェーズとしたとき、現在のフェーズを何秒先まで継続するかという現フェーズ長を制御対象とした学習を行うため、用いる目標値との誤差を図3のように算出する。状態遷移により直接的に発生した損失として制御サイクル間に増加した停止車両台数を用いる。次に、遷移後の状態の良さの指標としてはその状態から未来の事象の評価に当たり現実の測定結果を用いることは出来ないため、遷移後の状態から次の状態に遷移するときの各行動の評価値の最小値を用いる。

### 3.2 分散型信号機制御の概要

3.1節で強化学習による適用型手法の最適化の仕組みを示したが、マップ全体の車両分布の情報を一度に扱うと計算コストが非常に大きくなり、さらに信号機制御では近隣の信号機同士が制御結果に相互に影響し合うという性質があるため、マップ内の全ての信号機を集中的に制御することは難しいと考えられる。そこで、提案手法では強化学習による最適化の仕組みを実現するために、いくつかの信号機からなる信号機グループを制御単位とし、それぞれの信号機グループが独立して動作するような分散型の制御構造としている。

例えば、最小フェーズ長が20[sec]、フェーズ間の切り替えにかかる時間であるクリアリン

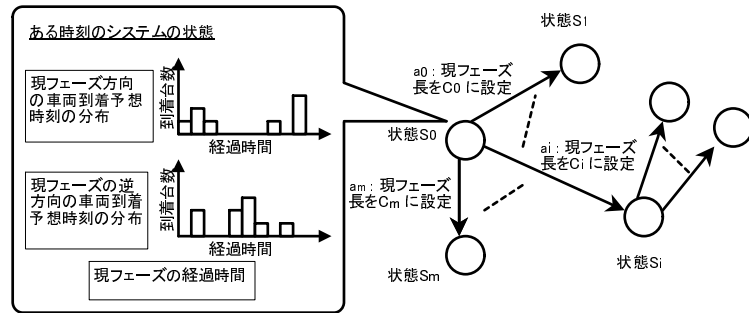


図 2 提案手法における状態例

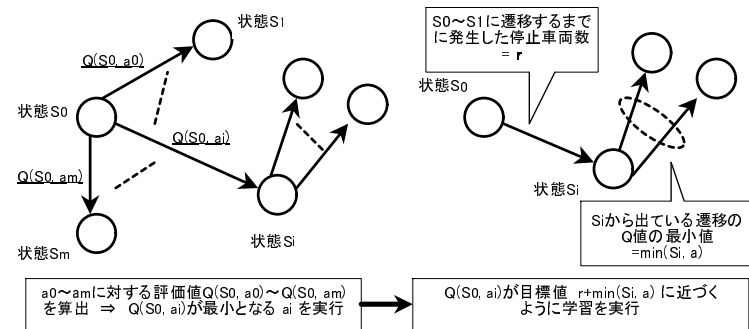


図 3 提案手法における学習誤差の算出

グ時間が 4[sec] であるならば、ある方向に交通権を与えるフェーズが終了してから、次にその方向に再び通行権を与えることが出来るまでの最短時間は 24[sec] である。よって信号機がどのような状態であっても、24[sec] 以内に目的の現示を行うことができるといえる。このようなパラメータの制約上では、例えば道路長が 200[m] で車両の平均的な速度が 15[m/sec] ならば、2 つの道路を車両が移動する間に 24[sec] 以上の時間が経過する。よって 1 つ先の信号機に存在する車両の存在を知ることが出来れば、急な交通量の変化が起こった場合にも対応することが可能であるため、隣接する信号機同士を信号機グループとすることで単体の信号制御と比較して大きな効果が得られるものと思われる。

提案システムは独立して信号制御を行った後、周囲にその情報を通知する独立制御信号機

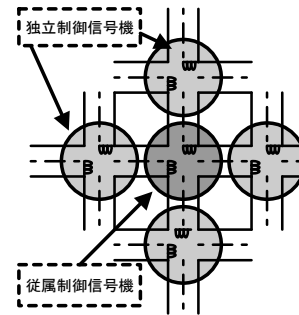


図 4 提案手法における信号機の配置

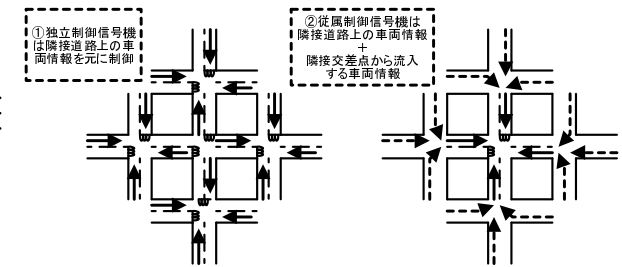


図 5 提案手法における階層的な制御

表 3 車両情報

車両 ID	車両を識別するための一意な ID
時刻	送信時刻 YYYY/MM/DD/hh/mm/ss
車両速度	0km/h ~ 100km/h
車両の位置情報	緯度経度情報 世界測地系 WGS-84
進行ベクトル	進行後方情報 16 方位系

と、周囲の独立制御信号機から通知された情報を用いて信号制御を行う従属制御信号機で構成する。例えば、格子状の道路網の場合には、図 4 のような 5 つの交差点からなる信号機グループを制御単位とし、それぞれの信号機グループが独立して制御サイクルを行う。制御は図 5 のように、まず独立制御信号機が制御を行い、次に従属制御信号機が制御を行う。このとき、独立制御信号機は自らと隣接する道路上の車両情報を用いた制御を行うが、従属制御信号機はそれに加えてすでに制御サイクルの終了した独立制御信号機を通過して自身に向かって来る車両についても考慮するため、より広範囲に存在する車両情報を用いた制御となる。

### 3.3 車車間通信による周辺車両情報の収集

提案手法では車両情報の収集に車車間通信を用いる。通信プロトコルには文献<sup>15)</sup>で提案されている CRCP を用い、車両から定期的に送信される車両情報のビーコンを信号機が受信することで車両情報を収集する。収集する車両情報を表 3 に示す。

車両速度、車両位置情報は車両が次の信号機に到着する時刻を算出するために必要であり、進行ベクトルは信号機に流入する車両であるかどうかを判別するために用いられる。加

表 4 車両の状態ごとの車両の重み

車両の状態	重み
隣接道路で走行中	1.0
隣接道路で停車中	0.1
隣接道路以外に存在	0.7

えて、各車両がどの交差点へと進むかを予測するために車両が左折・右折・直進レーンなどのどれに位置しているかというレーン番号も収集している。これらは、各信号交差点が自らの位置情報と、隣接する道路の形状の情報を知っているとすれば、車両の位置情報から求めることができる。

また図 5 で示したように従属制御信号機は独立制御信号機周辺の交通情報も考慮した制御を行うため、各制御サイクルで各独立制御信号機は、自らの属する各信号機グループの従属制御信号機に後述の手法で求めた車両分布情報を伝達することで信号機間の情報の共有を行う。

### 3.4 収集車両情報を基にした車両到着予想時刻の分布の算出

提案手法では車車間通信を利用して収集した車両情報を、車両の到着予想交通量の分布情報に変換した後、それらを現在の状態を入力として最適なフェーズ長の推定を行う。ここで、ある信号機に対する車両の到着予想交通量の分布とは、各秒ごとにその信号にどれだけの交通量が到着するかの予測値である。提案システムではまず信号機周辺の各車両に対して車両の状態と走行位置から予想到着時刻を求める。ここでは、それぞれの車両は現在の状態から法定速度まで一定割合で加速し、その速度を保って交差点へと到達する線型モデルで到着時刻を予測する。隣接の従属制御信号機より遠くの車両に関しては、走行レーン情報からその交差点での移動方向を予測し、従属制御信号機の信号切り替えタイミングからそこで見込まれる停車時間も加味して到着予想時刻を求める。

排気量最適化においては信号停止回数の削減が有効であるため、既に停止している車両に比べて走行中の車両の方が排気量に対する影響度が大きいと考えられる。そこで、提案システムでは到着予想台数に表 4 のように隣接道路を走行中の車両の重みが最も大きく、隣接道路で停止している車両の重みが小さくなるような係数を掛けて交通量としている。また、隣接道路以外に存在している車両の到着予想時刻は隣接道路に存在する車両に対する予想に比べて精度の低下が考えられるという点を考慮して、隣接道路以外に存在している車両の重みは小さく設定している。

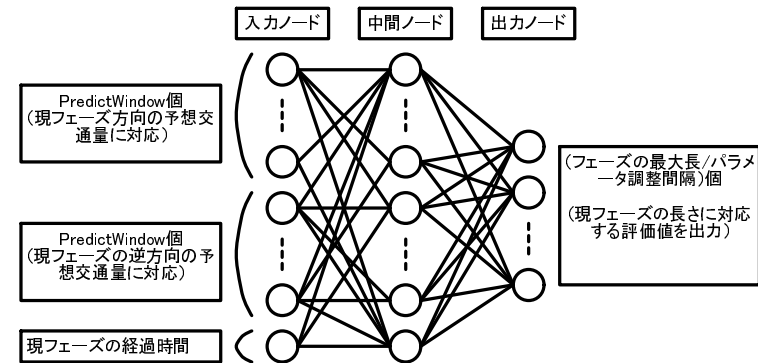


図 6 提案手法で用いるニューラルネットワークの構造

### 3.5 ニューラルネットワークによる制御パラメータ算出手法

3.4 節で求めた時刻ごとの車両到着予想の分布を用いて、ニューラルネットワークによって現フェーズの最適な長さを算出する手法について説明する。このとき使用するニューラルネットワークの構造を図 6 に示す。*PredictWindow* はシステムが何秒先の車両挙動まで予測するかを表す値であり、ニューラルネットワークの入力は現在青現示としている方向の到着予想交通量の分布、逆方向の到着予想交通量の分布、さらに現在のフェーズの経過時間を入力としており、入力ノード数は  $PredictWindow * 2 + 1$  である。また中間ノード数は入力ノード数と同数とした。出力ノード数はそれぞれのノードが現フェーズの青現示の継続時間に対応しているため、ノード数は最大フェーズ長から最小フェーズ長を引いた、フェーズ長の変化の幅をパラメータ更新間隔で割った値である。提案手法では教師なし学習を用いるため車両の停止台数の実測値と出力値の比較から出力値の学習データを生成するため、停止台数の実測値と比較可能な値が出力されるような機構とする必要があるためこのような出力構造としている。

ニューラルネットワーク内の各リンクはそれぞれ重みを持ち、各ノードはそれが中間ノードならば接続されている各入力ノードの、出力ノードならば接続されている各中間ノードの出力値と各接続リンクの重みによって重み付き和を求め、その値を励起関数に入力することでノードの出力値を算出する。提案システムでは出力値同士の比較が可能でなければならず、ニューラルネットワーク内の各ノードの出力値の算出にステップ関数のような出力が 2 値となるものを用いることができないため、励起関数として以下のような式で表されるシグ

モイド関数を用いる．

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

### 3.5.1 制御パラメータの決定

現制御サイクルにおけるシステムの状態を  $S_t$  としたとき，ニューラルネットワークの出力として，現フェーズ長を  $C_{min}, C_{min} + \delta, \dots, C_{min} + \delta \cdot i, \dots, C_{max}$  としたときのそれぞれの評価値  $Q(S_t, a_0), \dots, Q(S_t, a_i), \dots, Q(S_t, a_{(C_{max}-C_{min})/\delta})$  が得られる．現フェーズ長の決定では，現フェーズの経過時間より小さい値にフェーズ長を設定することはできないため，現フェーズの経過時間以上，フェーズの最大長以下に対応する各行動の評価値の中なら最小であるものを見つけ，その行動に対応するフェーズ長に現フェーズ長を決定する．よって，現フェーズの経過時間を  $\phi$  としたとき現フェーズ長  $C_{current}$  は次の式に従って決定される．

$$C_{current} = C_{min} + \delta \cdot \text{is.t.} a_i = \min_a Q(S_t, a), \phi \geq C_{min} + \delta \cdot i \quad (7)$$

### 3.5.2 決定した制御パラメータの評価

3.1 節で述べたように，提案手法では行動に対する評価値の推定がどの程度正確であったかを示す指標として，行動の評価値の真値と推定値の差を用いる．評価値の真値の算出は行動を実行してから次の制御サイクルまでに発生したコストと行動によって遷移した状態に対する評価値の和によって求めるため，ある制御サイクル  $t-1$  で選択された行動  $a_{t-1}$  に対する評価値  $Q(S_{t-1}, a_{t-1})$  の修正は次の制御サイクル  $t$  に得られる情報を用いて行われる．遷移によって発生するコストは信号機での停止車両の増加数を用いるため，システムは毎制御サイクルで停止車両数をカウントし，現制御サイクルの停止車両数と前制御サイクルの停止車両数の差によって算出する．よって，制御サイクル  $t-1 \sim t$  で発生するコスト  $r_t$  は，サイクル  $t$  の停止車両数を  $S_t$  とすると次の式で求められる．ここで用いる停止車両数は，信号機の交通権が与えられていない方向に存在する車両で，速度が一定の閾値以下であるような車両数をカウントすることで算出する．

$$r_t = S_t - S_{t-1} \quad (8)$$

また，遷移後の状態に対する評価値は，遷移後の状態  $S_t$  における各行動  $a_0 \sim a_{(C_{max}-C_{min})/\delta}$  の評価値  $Q(S_t, a_0), \dots, Q(S_t, a_{(C_{max}-C_{min})/\delta})$  の中の最小値とすることで，遷移後の状態における最適な行動に対する評価値としている．よって， $Q(S_{t-1}, a_{t-1})$  の真値は  $r_t + \min_a Q(S_t, a)$  であるため， $Q(S_{t-1}, a_{t-1})$  の修正に用いる誤差は次の式で求められる．

$$(r_t + \min_a Q(S_t, a)) - Q(S_{t-1}, a_{t-1}) \quad (9)$$

### 3.5.3 フィードバック動作

提案手法では，行動の評価値に対して誤差が算出された場合には，一般的な誤差修正手法であるバックプロパゲーションによってニューラルネットワークの内部パラメータの修正を行う．バックプロパゲーションでは，まず出力ノードから中間ノードに向かって，続いて中間ノードから入力ノードに向かってニューラルネットワーク内のリンクの重みを調整していくが，そのとき調整量の指標となる修正値は次の式で表される出力ノード  $i$  の目標出力値と実際の出力値の二乗誤差  $E_i$  である．ただし， $t_i$  は目標出力値， $o_i$  は実際の出力値とする．

$$E_i = \frac{1}{2}(t_i - o_i)^2 \quad (10)$$

提案システムでは，実際にシステムが選択した行動についてのみ修正を行うため，ある制御サイクル  $t-1$  で選択した行動を  $a_{t-1}$  とすると， $a_{t-1}$  以外の誤差は 0， $a_{t-1}$  の誤差は式 (7) で求められる値である．よって，各出力ノードに対する修正値は次のように表される．ただし  $\alpha$  は誤差の大きさを調整するための慣性係数である．

$$E_i = \begin{cases} 0 & (a_i \neq a_{t-1}) \\ \{\alpha((r_t + \min_a Q(S_t, a)) - Q(S_{t-1}, a_i))\}^2 & (a_i = a_{t-1}) \end{cases} \quad (11)$$

$E_i$  を用いた重みの調整は，まず中間ノードと出力ノード間のリンクについて行われる．誤差の発生した出力ノードに接続されている各リンクの重みを  $w_{ij}$  とすると，それぞれの重みの修正量は以下の式で求められる．ただし， $\eta$  は学習による影響を調整するための学習係数である．

$$\Delta w_{ij} = -\eta \frac{\partial E_i}{\partial w_{ij}} \quad (12)$$

式 (10) により，重みの変化量に対して修正値の変化が大きいような重みは調整量が大きくなるのが分かる．これにより修正値に対して影響力の大きい程リンクの重みが大きく調整され，効率的な出力値の修正が行えると考えられる．中間ノードと入力ノード間のリンクの重みの調整も同様にして行う．

## 4. シミュレーション実験による評価

提案する信号制御手法の有効性を評価するため，ミクロ交通シミュレータの VISSIM 上に本手法を実装し，シミュレーション実験を行った．VISSIM は，指定したマップ上に指定

表 5 シミュレーション環境

マップの形状	格子状
信号交差点間の間隔	200[m]
交通流平均の車両発生密度	1080[台/h]
最小フェーズ長	20[sec]
最大フェーズ長	60[sec]
シミュレート間隔	5[回/シミュレーション秒]
信号機のパラメータ決定の間隔	4[sec]
シミュレーション時間	1800[sec]
電波伝搬距離	200[m]
車車間通信の通信間隔	4[回/シミュレーション秒]

した交通量を発生させたとき、現実のドライバーに近い性質で各車両の挙動をリアルタイムにシミュレートすることができるシミュレータである。シミュレーション実験の環境を表 5 に示す。マップは都市街路を想定した道路長 200[m] の格子状のマップとし、車車間通信プロトコルは CRCP を用いた。

評価指標には車両の排気量と遅れ時間の 2 つの指標を用いる。排気量は各車両の速度・加速度から式 (1) の排気量モデルによって算出する。遅れ時間については、車両が信号機周辺で一定速度以下になった時間を信号機によって停止した時間とし、各信号機での車両の停止時間の合計によって求める。

#### 4.1 既存の信号機制御との比較

提案手法と既存の信号機制御比較との比較を行った。比較対象として、既存の信号機制御手法である TUC を用いた。TUC は車両キューが系統道路上の各信号機を通過する際に連続して青時間となるような制御を行うことで旅行時間を削減する手法である。

提案手法と比較対象の手法のそれぞれを、6×6 の格子状のマップ上でシミュレーションを行い、100[sec] ごとに発生したマップ上の全車両の総排気量と総遅れ時間を算出し比較した。ここでは車両情報を送信するための車載機の搭載率は 100[%] とした。

まず、総排気量の時間変化は図 7 のようになった。図 7 より、提案手法によって比較手法に対して排気量が最大で 20[%] 程度の削減されていることがわかる。また時間経過に伴って徐々に排気量の削減効果が現れており、学習による効果が確認できた。

続いて、総遅れ時間の時間経過による変化は図 8 のようになった。図 8 より、提案手法では時間経過に伴って総遅れ時間が徐々に減少し、ある程度の時間が経過すると比較手法と同程度まで削減されていることが分かる。しかし、500[sec]~700[sec] の間、1300[sec]~1400[sec] の間のように、逆に前の時間よりも遅れ時間が大きく増加している部分もあり、時

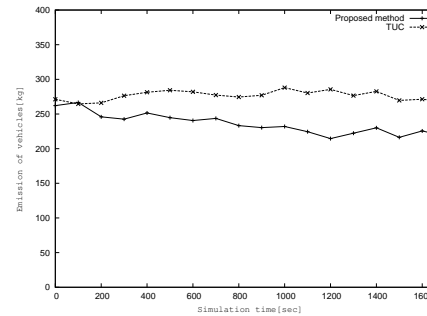


図 7 提案手法と比較手法の総排気量の時間変化

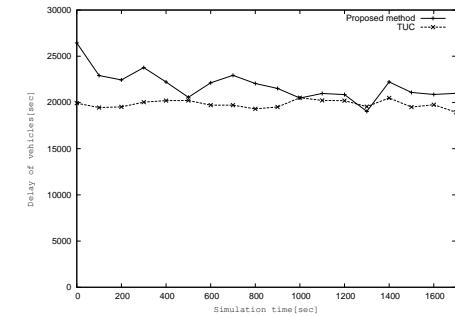


図 8 提案手法と比較手法の総遅れ時間の時間経過による変化

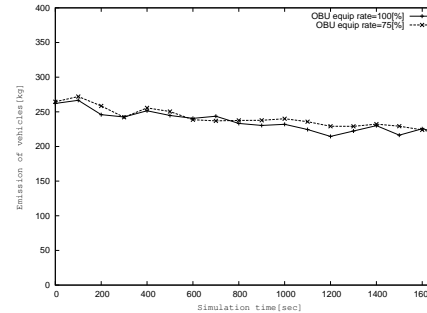


図 9 車載機の搭載率ごとの総排気量の時間変化

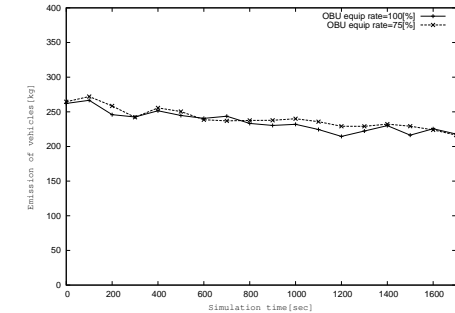


図 10 車載機の搭載率ごとの総遅れ時間の時間変化

間経過によって必ず遅れ時間が削減されているとは言えないことが分かる。これは、全体的に少しずつ最適化されていくが、時間的に局所的な入力の変化に対して即座に最適化されないというニューラルネットワークの性質によるものであると考えられる。

#### 4.2 通信状態による比較

続いて通信状態による比較として、車車間通信に必要な車載機の各車両への搭載率が異なる場合の制御結果の比較を行った。実験に用いるマップは 6×6 の格子状のマップとした。車載機の搭載率を 100[%]、75[%] とした場合のそれぞれの総排気量と総旅行時間の時間経過による変化は図 9、図 10 のようになった。

図 9、図 10 より車載機の搭載率が変化しても総排気量や総遅れ時間はほぼ変化しないこ

とが分かる。これは、入力の値の大きさが変化しても、入力のパターンが変化しなければ得られる出力の値のパターンも変化しないというニューラルネットワークの特性のためであると考えられる。この実験では車載機が搭載されていない車両の発生割合は一様であり、システムの把握する車両到着の分布のパターンが車載機の搭載割合によって大きく変化しなかったため、車載の搭載率が高い場合と同様の学習が行われたと考えられる。

## 5. ま と め

本稿では、都市街路における自動車の排気ガス量の削減を目的とした適応型のリアルタイム信号制御手法を提案した。提案手法では、対象区域内の各信号機が車車間通信を用いて周辺の車両情報を収集し、位置情報と速度情報から予想される車両の信号機への到着時刻情報を入力とするニューラルネットワークを用いた強化学習により制御パラメータの算出を行う。到着予想時刻情報を用いることで車両の挙動を考慮した制御を行い、停止回数と排気量の削減を実現する。また、複数の信号機からなる信号機グループを制御の単位とし、信号機グループを並べることで広域のマップに対応し、分散制御を実現する。

ミクロ交通流シミュレータ上でのシミュレーション実験では、都市街路のように信号機が小さい間隔で格子状に配置されているマップにおいて、提案手法は既存の交通量情報を用いた手法と比較して自動車の排気ガス量を削減でき、信号機による遅れ時間も同程度に抑えられることが確認された。

今後の課題としては、系統道路に対する制御など、既存のマクロな制御モデルを組み込むことで、より効果的な制御を行うことが挙げられる。また、今回は通信状況の比較として通信の成功確率の変化を対象としたが、通信プロトコルによる比較実験を行うことが考えられる。

## 参 考 文 献

- 1) 東京都環境局：“都における温室効果ガス排出量総合評価(2006年度実績)”，2009.
- 2) 横田 久司, 鹿島 茂, 國領 和夫, 田原 茂樹：“走行動態に基づく自動車からの環境負荷推計モデルの開発”，大気環境学会誌, Vol.40, No.2, pp.67-82, 2005.
- 3) L. I. Panis, S. Broekx, R. Liu：“Modelling Instantaneous Traffic Emission and the Influence of Traffic Speed Limits”，*Science of the Total Environment*, Vol.371, pp.270-285, Elsevier Science, 2006.
- 4) H. Rakha, Y. Ding：“Impact of Stops on Vehicle Fuel Consumption and Emissions”，*American Society of Civil Engineers, Journal of Transportation Engineering*,

Vol.129, No.1, pp.23-32, 2003.

- 5) S. K. Zegeye, B. De Schutter, J. Hellendoorn, E. A. Breunese：“Reduction of travel times and traffic emissions using model predictive control”，*Proceedings of the 2009 American Control Conference*, pp.5392-5397, 2009.
- 6) P. Mirchandania, L. Head：“A real-time traffic signal control system: architecture, algorithms, and analysis”，*Transportation Research Part C*, Vol.9, pp.415-432, 2001.
- 7) S. Mandava, K. Boriboonsomsin, M.Barth：“Arterial velocity planning based on traffic signal information under light traffic conditions”，*Proceedings of the 12th IEEE Intelligent Transportation Systems Conference*, pp.1-6, 2009.
- 8) K. J. Malakorn, P. Byungkyu：“Assessment of Mobility, Energy, and Environment Impacts of IntelliDrive-based Cooperative Adaptive Cruise Control and Intelligent Traffic Signal Control”，*Proceedings of the 2010 IEEE International Symposium*, pp.1-6, 2010.
- 9) 西川 郁子：“振動同期を用いた交通信号機制御法について”，システム制御情報学会誌 システム/制御/情報, Vol.52, No.5, pp.163-168, 2008.
- 10) I. Arel, C. Liu, T. Urbanik, A. G. Kohls：“Reinforcement learning-based multi-agent system for network traffic signal control”，*Proceedings of the 14th IET Intelligent Transport Systems*, Vol.4, pp.128-135, 2010.
- 11) PTV system：“VISSIM: A microscopic Simulation Tool to Evaluate Actuated Signal Control including Bus Priority”，*Proceedings of the 64th ITE Annual Meeting*, 1994.
- 12) D. I. Robertson, R. D. Bretherton：“Optimizing networks of traffic signals in real time-the SCOOT method”，*IEEE Transactions on Vehicular Technology*, Vol.40, pp.11-15, 1991.
- 13) V. Dinopoulou, C. Diakaki, M. Papageorgiou：“Applications of the urban traffic control strategy TUC”，*European Journal of Operational Research*, Vol.175, pp.1652-1665, 2006.
- 14) E. Kosmatopoulos, M. Papageorgiou, C. Bielefeldt, V. Dinopoulou, R. Morris, J. Mueck, A. Richards, F. Weichenmeier：“International comparative field evaluation of a traffic-responsive signal control strategy in three cities”，*Transportation Research Part A: Policy and Practice*, Vol.40, pp.399-413, 2006.
- 15) Masashi Saito, Jun Tsukamoto, Takaaki Umedu, Teruo Higashino：“Design and Evaluation of Inter-Vehicle Dissemination Protocol for Propagation of Preceding Traffic Information”，*IEEE Transactions on Intelligent Transportation Systems*, Vol. 8, No. 3, pp. 379-390, 2007 .