

# 交通信号機の協調制御に向けた情報共有効果の評価

Specifying Useful Neighbors for Sharing Information to Improve Distributed Traffic Light Control

加本昌也 荒井幸代  
Masaya Kamoto Sachiyo Arai

千葉大学大学院工学研究科  
Graduate School of engineering, Chiba University

Traffic congestion causes many serious problems especially in urban areas. Importance of researches on intelligent traffic control in multiagent context is recognized recently, because limitation of current methods, which apply the pre-defined rules to control each traffic signal. However, the multiagent traffic signal control does not become in use, because the traffic network is a large-scale and has dynamic interactions among intersections.

In this paper, we focus on the interactions among intersections to put an agent which manages traffic signal effectively. First, we show the experimental results of the effects on traffic by using multiagent information sharing. Second, we show how to divide a large-scale traffic network to be handled by multiple agents' cooperation.

## 1. 本研究の背景と目的

交通渋滞緩和に対する解決策として信号機制御を考える。本研究では、知的エージェントを対象道路に配置し強化学習による信号機制御を行うが、あるエージェントが制御する信号群の決定、および周辺の信号群を制御するエージェントと何についての情報を共有するかについては、そのエージェント同士の関係の粗密によると考える。交差点ネットワークにおいて、関係の密な対象領域を決定することによって、相互作用を考慮する必要のある交差点の集合を特定することが本研究の目的である。まず第一段階として単純な問題設定で実験を行い、知的エージェントによる学習型制御と協調の効果を確認する。続いて、実在する道路網や交通データを用いたシミュレーション環境を構築し実験を行う。相互作用が密な区域の抽出には、コミュニティの内外の密度差を表現できるコミュニティ濃度を利用したコミュニティ抽出法を用いる。

## 2. 問題設定

### 2.1 基礎事項

信号制御に関する用語 [交通工学 06] 交差点には複数の交通流がある。そのうち、ある一組の交通流に与えられる通行権(青, 黄, 全赤)のことを現示という。信号制御のパラメータには、サイクル, スプリット, オフセットの3種類がある。サイクルは、信号制御が一巡することを指し、1サイクルに要する時間をサイクル長または周期という。スプリットは、サイクル長に対する現示時間の比である。オフセットは、道路沿いの線上に連なる複数の信号機の灯色を順次入れ替える場合の、信号機群に共通な現示開始の基準時点からの遅れ時間を示す。

本研究では上記のうち、スプリットの制御を扱う。

Q学習 [高玉 03] Q学習は次式にしたがってルールの価値Q値を更新する強化学習手法のひとつである。

$$Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma \max_{a' \in \mathcal{A}(s')} Q(s', a') - Q(s, a)]$$

上式において、 $Q(s, a)$  は状態  $s$  において実行可能な行動集合  $\mathcal{A}_s$  のうち行動  $a \in \mathcal{A}_s$  をとるときの価値、 $Q(s', a')$  は遷

連絡先: 加本昌也, 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

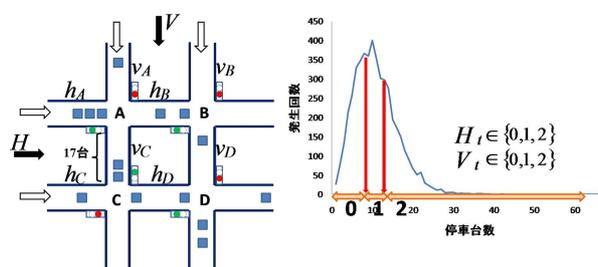


図 1: 道路環境

図 2: 停車台数の出現頻度

移先の状態  $s'$  において行動  $a'$  をとるときの価値を示している。また、 $r$  は状態  $s$  から  $s'$  への遷移において得られる報酬、 $\alpha$  ( $0 < \alpha \leq 1$ ) は学習率、 $\gamma$  ( $0 \leq \gamma \leq 1$ ) は割引率を表している。

### 2.2 環境設定

対象とする環境は、図 1 に示す格子状の道路環境で、4つの交差点 A, B, C, D が存在する。水平 (H), 垂直 (V) 方向の各道路は図 1 に示す矢印の向きだけの一方通行とする。また、車は外部から接続する道路上に発生し、右左折することなく動き、外部に接続する道路にたどり着いた時点で消滅する。

#### 2.2.1 集中型制御

決められたプログラムによる集中制御として全体を 1 エージェントで一括管理する場合について考える。

ここでは、図 1 の交差点 A, B, C, D に存在する全信号を 1 エージェントが強化学習を用い制御する。

(1) 状態集合  $S$ :  $s \in S$  は図 1 の水平, 垂直方向の各道路の停車台数に基づいて決定する。

停車台数は「赤信号から青信号に変わった直後に停止している車の台数」とする。

水平, 垂直方向の各道路の停車台数を  $h_i, v_i$  とし各々の総和  $R_h = \sum_{i \in A, B, C, D} h_i$  と  $R_v = \sum_{i \in A, B, C, D} v_i$  を計算する。予備実験から求めた図 2 に示す停車台数出現頻度を基に、 $R_h, R_v$  を停車台数が(少ない), (普通), (多い) 場合に分割し、各々 0, 1, 2 の値に対応させる。t サイクル時点での水平方向, 垂直方向の値を各々  $H_t, V_t$  とする。すなわち状態  $s_t$  は  $(H_t, V_t)$  の組で表現され、状態数は  $3^2$  となる。

(2) 行動集合  $\mathcal{A}$ : 各交差点における水平, 垂直方向のスプリット

の組み合わせ  $\mathcal{A} = \{a|(30\%, 70\%), (50\%, 50\%), (70\%, 30\%)\}$  から選択し, 4つの交差点の行動組み合わせ数は  $3^4 = 81$  通りとなる. なお行動選択手法は  $\epsilon$ -グリーディ手法を用いる.

(3) 報酬  $r$ : 状態  $s_{t+1}$  と  $s_t$  における停車台数の増減を意味する  $\Delta H$  と  $\Delta V$  から報酬値  $r$  を計算する. 具体的には以下の式を用いる.

$$r = 3 \times (H_t - 4/3 \times H_{t+1}) - H_t \times H_{t+1} + 3 \times (V_t - 4/3 \times V_{t+1}) - V_t \times V_{t+1}$$

以上により集中型制御における状態-行動の組み合わせは  $|\mathcal{S}| \times |\mathcal{A}| = 9 \times 81 = 729$  となる.

### 2.2.2 分散型制御

エージェントを各交差点に配置する. 図1の交差点に存在する信号を各々エージェント A, B, C, D が管理する. さらに, 各エージェントの情報共有関係を, 1. 各エージェントが独立であり情報共有しない, 2. 知覚情報を共有する, 3. 政策を共有する, の3つの場合について考える.

1. 独立に制御する場合各交差点の信号機  $i$  を以下の状態集合, 行動集合, 報酬によってモデル化する. なお, 行動集合, 報酬は情報共有する3段階において同じである.

(1) 状態集合  $\mathcal{S}$ :  $s \in \mathcal{S}$  は停車台数に基づいて決定される. はじめに図1の水平, 垂直方向の停車台数  $h_i$  と  $v_i$  を測定する. 停車台数の定義は集中型制御の場合と同様に「赤信号から青信号に変わった直後」の停車台数とする. つぎに予備実験から求めた図2のような停車台数の出現頻度を基に,  $h_i, v_i$  を停車台数が(少ない), (普通), (多い)に対応させて 0, 1, 2の値に対応させ, 各々  $H_t, V_t$  とし,  $(H_t, V_t)$  の組み合わせを  $s_t$  とする.

(2) 行動集合  $\mathcal{A}$ : スプリットを  $\mathcal{A} = \{a|(30\%, 70\%), (50\%, 50\%), (70\%, 30\%)\}$  から選択する.

(3) 報酬  $r$ :  $s_t$  から  $s_{t+1}$  に遷移した時の  $\Delta H, \Delta V$  を用いて報酬値を計算する. エージェント  $i$  は各々個別に政策を更新する.

状態-行動の組み合わせは各々のエージェントについて  $|\mathcal{S}| \times |\mathcal{A}| = 9 \times 3 = 27$  となる.

2. 知覚情報を共有する場合エージェント  $i$  と  $j (i \neq j)$  が停車台数に関する情報を共有する場合を示す.

状態集合  $\mathcal{S}$ : エージェント  $i$  と  $j$  が各々お互いの状態  $s^i = (H_t^i, V_t^i)$ ,  $s^j = (H_t^j, V_t^j)$  を観測し, その組み合わせを各々状態入力  $s_t$  とする. したがって状態数は  $9^2$  となる. また行動集合  $\mathcal{A}$  は「独立に制御する場合」と同様に3種類の行動からなる.

エージェントは各々政策を持ち更新する.

ここでは状態-行動の組み合わせは各々のエージェントについて  $|\mathcal{S}| \times |\mathcal{A}| = 9^2 \times 3 = 243$  となる.

3. 政策を共有する場合エージェント  $i$  と  $j (i \neq j)$  が各々学習した政策を共有する場合を示す.

状態集合  $\mathcal{S}$  と行動集合  $\mathcal{A}$  は「独立に制御する場合」と同様であるが, エージェント  $i, j$  は同一の政策を更新し利用する.

ここでは状態-行動の組み合わせは各々のエージェントについて  $|\mathcal{S}| \times |\mathcal{A}| = 9 \times 3 = 27$  となる.

## 3. 計算機実験

### 3.1 実験設定

道路上の車の動きには ASEP モデル [渋滞学 03] を用いる. 交差点間に停車可能な車の台数は 17 台で全体で最大 136 (17 × 8) 台停車可能である.

車の流入率に差を設け, 水平方向  $H$  と垂直方向  $V$  の流入率の比  $H : V = 9 : 1$  とし, 集中型制御, 分散型制御による実験を行った.

また, 情報共有の効果を観察することを目的として, 政策を共有するエージェントの組み合わせを変え実験した. なお, シミュレーションには NetLogo 4.0.2 [NetLogo] を用いた.

### 3.2 実験結果

ランダム制御, 集中型, 分散型による制御を行った停車台数の変化を図3に示す.

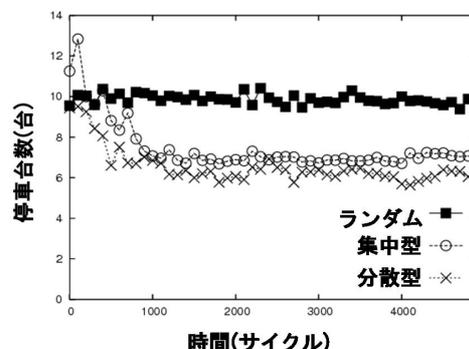


図3: 停車台数の変化

強化学習の結果, いずれの場合も停車台数が減少したことが確認できる. しかし, 集中型による学習では, 分散型に比べ学習が遅く, 停車台数が減っていない. これは集中型は分散型に比べ状態行動空間が大きいため学習に時間がかかり, かつ報酬を全体の停車台数の増減から与えているため, 交差点単位で報酬を与えることができる分散型に比べ精度の高い学習ができなかったためであると考えられる.

次に情報共有した実験として, 垂直方向の交差点 (A と C, B と D) が政策を共有し学習した場合と独立に学習した場合の停車台数の変化を図4(a)に, 垂直方向の交差点 (A と B, C と D) が政策を共有した場合を図4(b)に示す. 図4(a), 図4(b)から, 垂直方向 AC-BD の組で政策共有した場合は独立に学習した場合に比べ学習の収束が速い一方, 水平方向 AB-CD の組み合わせでは学習効率に変化がないことが確認できた.

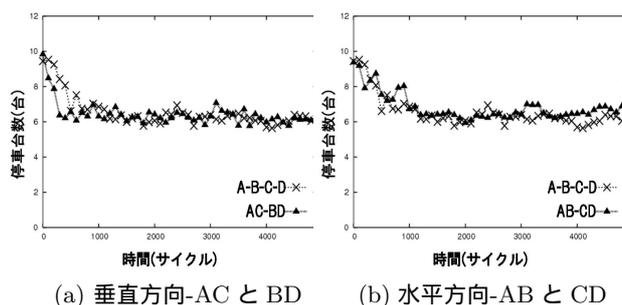


図4: 水平方向と垂直方向の協調の効果

ここで, エージェント A から D が各々独立に強化学習した過程で, 報酬を同時に得る回数をエージェントの組み合わせ別に図5に示す.

この結果から本設定では, エージェント A と C, エージェント B と D が他の組み合わせに比べて, 同時に報酬を得る回数が多いのがわかる. このことから, この2エージェントの学習

環境が類似しているため、政策を共有し一緒に更新することによって学習が加速したといえる。

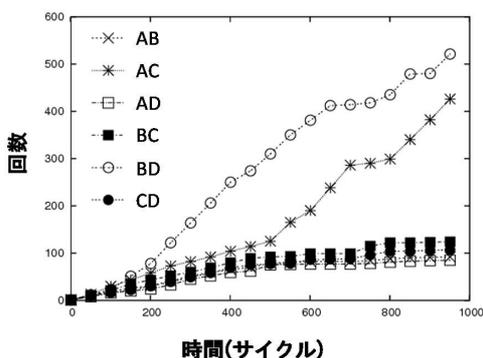


図 5: 報酬を同時に得た回数

## 4. 交通網への応用

### 4.1 交通網の例

前節で示した実験より、交通量に依存関係のある交差点の信号機間で情報を共有することの有効性を確かめた。以後の章では依存関係の強い交差点を抽出する方法を説明する。

対象環境は、JR 京葉線検見川浜を中心とした千葉県千葉市美浜区磯辺、および、真砂の国道 14 号線、海浜大通り、花見川、マリーナストリートを境界とする地区とし、図 6 のようにモデル化した。図 6 の中心の印を JR 検見川浜駅ロータリーとし、他の印は対象地区外へつながる短点とする。印は信号付き交差点を示す。実験には、知的マルチエージェント型交通流シミュレータ MATES (Multi-Agent based Traffic and Environment Simulator) [西川 02] を使用する。

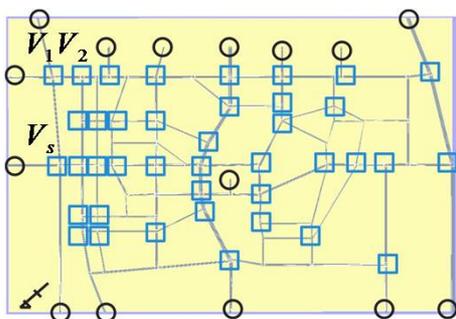


図 6: 対象とする交通網 (千葉市美浜区周辺)

### 4.2 協調対象区域の抽出

ここまでの基礎実験では、情報共有する相手を設計者があらかじめ決定し組み合わせを変え比較をしたが、大規模環境の実験を行う際には、設計者が組み合わせを考えるのではなく、関係の密な交差点エージェントを見つけ出し情報共有などの協調を行う方法をとるべきである。協調すべき交差点集合を見つける方法として、コミュニティ濃度に基づくコミュニティ抽出法 [高橋 08] が挙げられる。以下、この手法について説明する。

## 5. コミュニティ濃度に基づいた抽出法

### 5.1 コミュニティ濃度

コミュニティ濃度  $Cd$  とは、式 (1) で定義される評価尺度である。

$Cd \geq 1$  を満たす部分グラフ  $G_s$  は周囲より密であることを示し、このときその部分グラフ  $G_s$  をコミュニティとみなす。

$Cd$  値が大きいとき、部分グラフ内の連結度が部分グラフとその外部との連結度よりも密であること、つまりコミュニティ内外の相対的な密度差が大きく、かつコミュニティ内部の密度が外部の密度よりも大きいことを意味する。

$$Cd = \frac{k_{G_s}}{k'_{G_s}} = \frac{|E_s| \times 2}{|V_s|} \bigg/ \frac{(\text{境界リンク数})}{|V_s|}$$

$$= \frac{|E_s| \times 2}{(\text{境界リンク数})} \quad (1)$$

$k_{G_s}$ : 部分グラフ  $G_s$  内のノードの平均次数 (コミュニティ内リンク密度)

$k'_{G_s}$ : グラフ  $G' = (V_s, E \setminus E_s)$  におけるノード  $v_s \in V_s$  の平均次数 (コミュニティの境界リンク密度)

$V_s$ : 部分グラフ  $G_s$  内のノード集合

$E_s$ : 部分グラフ  $G_s$  内のリンク集合

### 5.2 コミュニティ濃度に基づいた抽出法

コミュニティ濃度  $Cd$  に基づいた探索型コミュニティ抽出法 (Cd-based-Method: 以下 CDB 法) を示す。はじめに、コミュニティの探索の始点となるシードノードを選択し、これをコミュニティ内の初期ノードとする。次にコミュニティ内のノードと接続している各ノードをコミュニティに加えた場合のコミュニティ濃度  $Cd$  を計算し、最もコミュニティ濃度  $Cd$  値が大きいノードをコミュニティ内に加える。終了条件を満たすと探索を終了する。

CDB 法はコミュニティ内のノードの加えられた順番とそのときのコミュニティ濃度  $Cd$  が算出される。コミュニティ濃度  $Cd$  の推移を計測し、コミュニティ濃度が極大値を示したときに、等高線を引きシードノード中心のコミュニティ図を作成することができる。

この手法は重み付きグラフ、重み付きではないグラフの双方に適用可能であるため、交通量を反映したコミュニティ抽出が可能である。重み付きグラフの場合はリンクの重みをそのまま使用し、重み付きではないグラフではリンクの重みを 1 として使用する。

### 5.3 交差点ネットワークへの適用

このコミュニティ濃度を用いたコミュニティ抽出を、交差点ネットワークに適用することによって、関係の密な交差点集合を特定する。例えば交差点ネットワークにおいて、ノードを交差点に対応させ、リンクを道路、重みを交通量に対応させることができる。ここで、図 6 の  $V_s$  の交差点をシードノードとすると、最初に  $V_1$ 、次に  $V_2$  の交差点といった順にコミュニティ内にノードが加えられていく。CDB 法によりノードの加えられた順番とそのときのコミュニティ濃度が得られる。重みを考えない場合のコミュニティ濃度の推移を図 7 に示す。このグラフの極大値からコミュニティ図を作成できる。ノードがコミュニティに取り込まれる様子を図 8 に示す。図中の黒く塗りつぶしたノードがコミュニティに取り込まれたノードである。本手法は、中心として考えた任意のノードをシードノードとすることが可能であることから、信号の制御対象地域に焦点をあてた

コミュニティが抽出できる。また、上述したように CDB 法は重み付きグラフに適用可能なので、交通量を重みとすることで交通状況を考慮に入れることができる。

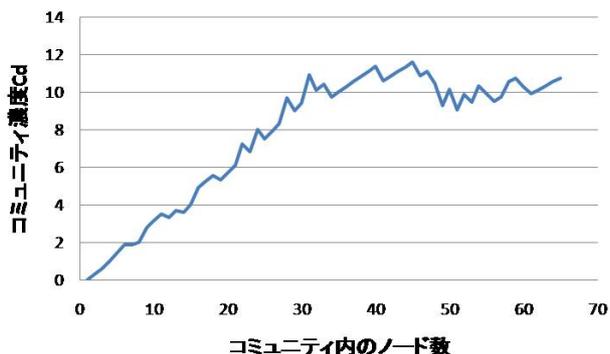


図 7: コミュニティ濃度  $C_d$  の推移例

## 6. まとめ

本研究では、協調すべき交差点を特定するための第一段階として、はじめに各交差点における停車台数に関する情報や学習結果を共有することの効果、集中型と分散型の強化学習実験によって評価した。集中型による強化学習では、状態行動数が増加すること、また報酬が信号機全体に対してしか設計できないことから、分散型に比べ学習に時間がかかり性能が悪化した。計算時間、報酬設計と収束値の観点から、本設定では集中型よりも分散型が有効であった。

また、政策共有が効果的なのは、制御対象とする交差点の水平方向と垂直方向の停車台数の比が類似している場合に限り有効であることが示された。

次に、交通網への応用として、つながりの密な交差点を特定する手法として利用できるコミュニティ濃度に基づくコミュニティ抽出法を紹介し、実在する道路網において協調すべき交差点を抽出する指針を示した。

今後の課題として、スプリット以外のパラメータ制御、協調対象の考察、重み付きでのコミュニティ抽出などが挙げられる。

## 参考文献

- [交通工学 06] 社団法人 交通工学研究会：改訂 交通信号の手引き、丸善株式会社、2006.
- [高玉 03] 高玉圭樹：マルチエージェント学習-相互作用の謎に迫る-、コロナ社、2003.
- [渋滞学 03] 西成活裕：渋滞学、新潮社、2003.
- [NetLogo] <http://ccl.northwestern.edu/netlogo/>
- [西川 02] 西川紘史、式安智、吉村忍：知的マルチエージェント交通流シミュレータ MATES の開発、第 21 回計算電気・電子工学シンポジウム、pp151-154、2002.

- [高橋 08] 高橋篤、荒井幸代、須貝康雄、”コミュニティ概念の定式化、” Joint Agent Workshops and Symposium 2008、電子情報通信学会、日本ソフトウェア学会、情報処理学会、人工知能学会、CD-ROM、2008.

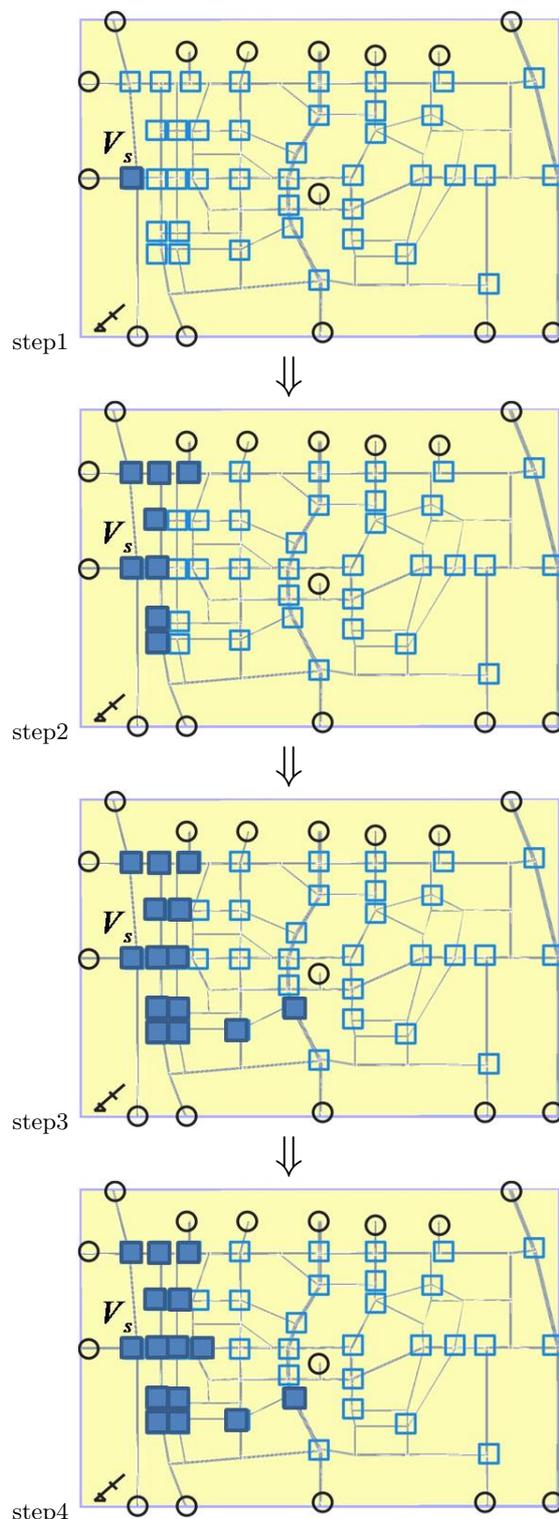


図 8: 図 6 の  $V_s$  を中心としたつながりの強い交差点の抽出過程