

強化学習に基づく先見のかつ協調的な輸送システム

向 直人 渡邊 豊英

名古屋大学大学院 情報科学研究科 社会システム情報学専攻
〒 464-8603 名古屋市千種区不老町

近年、オン・デマンドな交通システムが注目されている。既存研究の多くは、顧客の要求発生後に、顧客満足度を最大にするような輸送車両や走行経路を割り当てる。一方、本稿では、強化学習手法(Q-学習)を用いることで、都市に潜む輸送要求の傾向を経験から学習し、先見のかつ協調的な運行サービスを提案する。車両は先天的に顧客発生傾向に関する知識を持つ必要はなく、輸送経験から得られる報酬を基に運行経路を獲得する。また、車両は、獲得した顧客発生傾向や、サービス領域のトポロジカルな特徴から各自の担当領域を獲得する。最後に、提案する運行システムのシミュレーションを行い、その有効性を評価する。

Proactive and Collaborative Transport System Based on Reinforcement Learning

NAOTO MUKAI TOYOHIDE WATANABE

DEPARTMENT OF SYSTEMS AND SOCIAL INFORMATICS,
GRADUATE SCHOOL OF INFORMATION SCIENCE, NAGOYA UNIVERSITY
FURO-CHO, CHIKUSA-KU, NAGOYA, 464-8603, JAPAN

In these years, on-demand transport systems have been focused. Most of traditional studies related to such on-demand systems are regarded as reactive systems. In fact, transport vehicles and their traveling routes are assigned to customers after the occurrences of their demands. In contrast, we propose a proactive and collaborative transport system based on reinforcement learning technique (Q-Learning). Transport vehicles can acquire proactive traveling routes depending on rewards obtained in their transport experiences. Moreover, transport vehicles acquire individual responsible area depending on the topological features of their service area in addition to the rewards of transportations. Finally, we evaluate the efficiency of our transport system by simulation experiments.

1. はじめに

近年、高度道路交通システム (ITS) に関する研究が盛んである。その一つとして、デマンド・バスやインターネット・タクシーと呼ばれる新しい交通システム^{11),12)}が注目されている。それらのシステムの特徴は、位置情報システム (GPS) を活用して、輸送車両や顧客の位置を特定することで、顧客に最適な輸送車両や走行経路を割り当てることにある。既存研究の多くは、顧客の要求発生後に、顧客満足度を最大にするような輸送車両や走行経路を割り当てるといった反射的な最適化問題を扱っている^{1),2),5),6),9)}。一方、本稿では、強化学習手法 (Q-学習) を用いることで、都市に潜む輸送要求の傾向を経験から学習し、先見のかつ

協調的な輸送サービスを提案する。

先見のとは、過去に得た輸送経験から顧客の発生分布を予測することで、輸送車両が顧客の発生に対して先回りの行動をとることを意味する。例えば、空車状態において、輸送車両は新たな顧客の発生が見込める位置に事前に移動することで、顧客の待機時間を減少させることができる。また、経路選択において、単に顧客の乗降位置間の最短経路を選択するのではなく、他の顧客の乗降が見込める経路を選択することで、効率的な運営が可能となる。本稿では、このような先見の行動を一般化するために、顧客の乗降を輸送車両の報酬としてとらえ、輸送車両は将来得られる期待報酬を最大にする行動を選択する。協調的¹⁰⁾とは、輸送車両間での効果的な負荷分散を意味する。一般的な負

荷分散の方法として、「巡回分散」と「領域分散」が考えられる。前者は、巡回経路や時刻表を設定し、車両間の巡回地点（巡回経路上の一点）への到着のタイムラグを一定にすることで、サービス領域全体をカバーすることである。一方、後者は、サービス領域を分割し、各車両に割り当てることによって、担当領域を明確にすることである。本稿では、後者に注目し、担当領域間をリレー形式に顧客を輸送することを試みる。サービス領域をグラフ構造とみなし、そのトポロジカルな特徴や、学習によって獲得した報酬の期待値に基づき領域をクラスタリング（分割）^{3),4)} する。

本論文の構成は以下である。2章で提案するリレーに基づく輸送システムを形式化する。3章で顧客発生傾向のQ-学習を用いた獲得方法、4章でサービス領域のクラスタリング方法を提案する。5章でシミュレーション実験により提案手法を評価し、6章でまとめと今後の課題を述べる。

2. 形式化

提案システムを以下のように形式化する。サービス領域を式(1)で与える。ノード n は顧客の乗降位置を表し、エッジ e はノード間の経路を表す。ノード数は有限であるとし、その数をグラフの大きさとし $|G|$ と表す。エッジ長は全て同一であるとする（輸送車両がノード間を遷移するために必要な時間は全て等しい）。また、一方通行や車線数等の制限は考慮しない。

$$\begin{cases} G &= (N, E) \\ N &= \{n_1, n_2, \dots\} \\ E &= \{e(n, n') | n, n' \in N\} \end{cases} \quad (1)$$

顧客の要求発生を以下のようにモデル化する。一般に、顧客の発生分布は一様ではない。出勤時間であれば、住宅街からビジネス街へといった傾向が存在すると考えられる。そこで、サービス領域内に存在する要求発生傾向を式(2)で与える。フロー f は、乗車位置 n_r 、降車位置 n_d 、発生確率 η によって特徴付けられる。サービス領域には、特徴の異なるフローが複数重なり合っており存在しているものとする。

$$\begin{cases} F &= \{f_1, f_2, \dots, f_m\} \\ f &= (n_r, n_d, \eta) \end{cases} \quad (2)$$

台数 k の車両を式(3)で与える。走行速度や最大顧客乗員数は全車両で同一とする（車両の能力は等しい）。よって、担当領域の割り当ては車両の能力に依存しないと考える。つまり、車両が分割されたどの領域を担当したとしても、得られる成果は等しい。

$$V = \{v_1, v_2, \dots, v_k\} \quad (3)$$

ここで、各車両の担当領域を定義するために、サービス領域の分割領域を式(4)で定義する。分割領域は4章で述べるクラスタリング手法によって得られる。サービス領域 G の部分集合が分割領域 SA であり、分割領域 SA 間のオーバーラップ（積）はないものとする。つまり、 $\bigcup SA = G, \bigcap SA = \phi$ である。

$$C = \{SA_1, SA_2, \dots, SA_k\} \quad (4)$$

次に、分割領域の接続領域を定義する。接続領域 $CA_{SA \rightarrow SA'}$ とは、分割領域 SA 内に存在する、隣接する分割領域 SA' に接続しているノード ($n \in SA$) とエッジ ($e(n, n') | n' \in SA'$) の集合である。

定義された分割領域と接続領域から担当領域を式(5)で定義する。車両は担当領域内のみを移動可能であり、もし領域外への輸送要求が発生した場合は、顧客を接続領域で一端降車させ、隣接する領域の担当車両に依頼する。

$$\begin{cases} CA_i &= \bigcup_{SA \in C - SA_i} CA_{SA \rightarrow SA_i} \\ TA_i &= SA_i \cup CA_i \end{cases} \quad (5)$$

具体例を図1に示す。サービス領域が8つのノード (n_1, n_2, \dots, n_8) から構成されているとする。このとき、あるクラスタリング手法によって、2つの領域 ($SA_1 = \{n_1, n_2, n_5, n_6\}, SA_2 = \{n_3, n_4, n_7, n_8\}$) に分割されたとする。両領域間でエッジを共有している部分が接続領域 ($CA_{SA_1 \rightarrow SA_2} = \{n_2, n_6\}, CA_{SA_2 \rightarrow SA_1} = \{n_3, n_7\}$) である。よって、担当領域 TA_1 は $SA_1 \cup CA_{SA_2 \rightarrow SA_1} = \{n_1, n_2, n_3, n_5, n_6, n_7\}$ 、 TA_2 は $SA_2 \cup CA_{SA_1 \rightarrow SA_2} = \{n_2, n_3, n_4, n_6, n_7, n_8\}$ となる。ここで、 n_1 から n_8 に向かうフローが存在した場合、 TA_1 を担当する車両は顧客を一端 n_7 まで輸送する。その後、 TA_2 を担当する車両が顧客を n_8 まで輸送する。以上のように、顧客をリレー形式に輸送することで、サービス領域全体をカバーする。

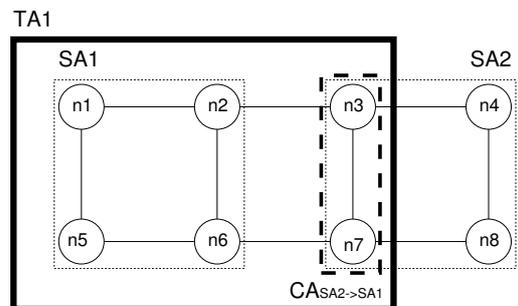


図1 リレー輸送

3. 学 習

担当領域が割り当てられた車両は、領域内の要求発生分布を強化学習手法 (Q-学習)^{7,8)} によって学習し、政策 (どの経路を選択すべきか) を獲得する。本稿で対象とする輸送問題に合わせて、Q-学習に必要な、“状態-行動”、“報酬”、“刑罰” を定義しなくてはならない。本章では、それらの定義を順に述べ、最後に、獲得した報酬の推定値に基づき、どのように政策を導くかを述べる。

3.1 状態-行動

車両の時刻 t における状態 s_t を、現在ノード n_t と、そこに至るまでの長さ δ のノード履歴 $\{n_{t-\delta}, \dots, n_{t-1}\}$ の列で与える。よって、状態 s_t は式 (6) で表される。

$$s_t = (n_{t-\delta}, \dots, n_{t-1}, n_t) \quad (6)$$

ここで、ノード n に接続するノードの集合を $\{L(n) | n' \in L(n), e(n, n') \in E\}$ と表す。状態 s_t における行動は、現在ノード n_t に接続するノード集合 $L(n_t)$ から選択され、式 (7) で表される。

$$a_t = n_{t+1} \in L(n_t) \quad (7)$$

結果的に、履歴長 $\delta = 1$ のとき、状態と行動の組は式 (8) で与えられる。この状態と行動の組に対して、報酬の期待値 (Q 値) が定まることになる。

$$(s_t, a_t) = (n_{t-1}, n_t, n_{t+1}) \quad (8)$$

以上のような状態と行動の定義は、車両の経路選択が現在ノードのみでなく、どの経路を辿ったかに依存することを表している。図 2 に示すように、ノード n_1 を経由して、ノード n_3 に到着した車両と、ノード n_2 を経由して、ノード n_3 に到着した車両とでは、次の行動を選択したときの報酬の期待値が異なる。結果的に、履歴長 δ は、ある輸送要求の乗車位置と降車位置の関連度を表す。履歴長 δ が長いほど、長距離を移動する要求の報酬をより正確に学習できるが、記憶すべき状態-行動の数が膨大となる。

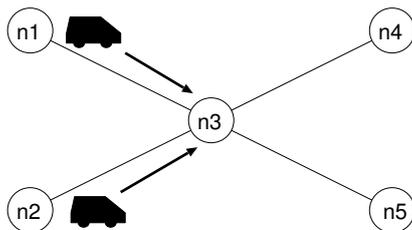


図 2 車両の状態と行動

3.2 報 酬

車両がノード n に到達したときに得られる報酬を式 (9) で与える。ここで、 $|R_r^v(n)|$ はノード n で乗車する顧客の人数であり、 ω_r はその重み係数である。また、 $|R_d^v(n)|$ はノード n で降車する顧客の人数であり、 ω_d はその重み係数である。ここで、重み係数のバランスは、車両の振舞いに強く影響する。例えば、乗車係数 ω_r を大きく、降車係数 ω_d を小さくすると、車両は一度に多くの顧客を乗車させることを好む。逆に、乗車係数 ω_r を小さく、降車係数 ω_d を大きくすると、車両は顧客を頻繁に降車させることを好む。

$$R^v(n) = \omega_r \cdot |R_r^v(n)| + \omega_d \cdot |R_d^v(n)| \quad (9)$$

3.3 刑 罰

車両がノード n に到達すると、ノード n で乗降する顧客の要求は満足される。よって、到達直後のノード n で得られる報酬は 0 である。時間が経過することによって、要求がある確率に従って発生し、本来の期待報酬値に近付くと考える。そこで、ノード n の経過時間 $idle(n)$ に従って、期待報酬の推定値である $Q(s_t, a_t)$ を低く見積もる。ここで、経過時間 $idle(n)$ とは、車両がノード n に到達してから、経過した時間を表す。刑罰関数 $P(a_t)$ は、式 (10) で与えられ、推定値 $Q(s_t, a_t)$ の重み係数として用いられる。 ζ は刑罰を与える最大回数を表している。刑罰関数 $P(a_t)$ は、時間経過と共に 0 から 1 に増加し、最大回数 ζ を越えると常に 1 となる ($P(a_t) = 1$ は本来の期待報酬値を表す)。

$$P^v(a_t) = \begin{cases} \frac{idle(n_{t+1})}{\zeta} & (if \ idle(n_{t+1}) < \zeta) \\ 1 & (else) \end{cases} \quad (10)$$

3.4 政 策

以上のような定義に基づいて得られた期待報酬の推定値 (Q 値) から、車両の政策を ϵ -ルーレット手法 ($0 < \epsilon < 1$) で与える。すなわち、確率 ϵ で、車両はランダムに行動を選択し、確率 $1 - \epsilon$ で、車両は式 (11) で示されるルーレット手法で行動を選択する。

$$Pr^v(s_t, a_t) = \frac{P^v(a_t) \cdot Q(s_t, a_t)}{\sum_{a'_t \in L(n_t)} P^v(a'_t) \cdot Q(s_t, a'_t)} \quad (11)$$

4. 領 域 分 散

本章では、サービス領域を分割するための、クラスターリング方法を提案する。最初に、クラスターリングの概要を述べ、次に、クラスターリングの基準となる適合

関数について述べる。

4.1 クラスタリング

クラスタリングによって、サービス領域 G を分割し、その部分集合である分割領域 C を得る。クラスタリングは2ステップ(ボトム・アップ, 最適化)で構成される。ここで、分割領域を評価するための適合関数 $fit(C)$ を考える。適合関数 $fit(C)$ はその値が小さいほど形成された分割領域 C が適していることを表す。

ボトム・アップでは、分割領域の結合を繰り返すことによって、車両台数 k と同数の分割領域を得る。まず、初期状態では、図 3(a) に示すように、ノードを1つだけ含む分割領域を生成する。次に、図 3(b) に示すように、分割領域を結合することによって、適合関数 $fit(C)$ の値が最も小さくなる組を選択し、分割領域の結合を繰り返していく。分割領域の数が、車両台数 k と等しくなったとき、結合を終了する。

最適化では、山登り法に従い、ボトム・アップによって得られた分割領域間でノードを交換し、適合関数 $fit(C)$ の値を最小にする。ボトム・アップにより、図 3(c) に示すような、車両台数 k と同数の分割領域が得られる。次に、図 3(d) に示すように、適合関数 $fit(C)$ の値が最も小さくなるように、分割領域間でノードの交換を繰り返す。ただし、交換するノードは以下の2つの条件を満足する必要がある。分割領域 SA から SA' に交換されるノード n は接続領域 $CA_{SA \rightarrow SA'}$ に含まれていなければならない。また、交換されるノード n は分割領域 SA の連結性を破壊しない。連結性とは、領域内のいずれのノードからも、エッジを辿ることで、その他の全てのノードに辿り着けることを意味する。適合関数 $fit(C)$ の値が改善されなくなったとき、ノードの交換を終了する。

4.2 適合関数

本節では、クラスタリングの基準となる、サービス領域のトポロジカルな特徴に基づく適合関数と、学習によって獲得した報酬の期待値(Q値)に基づく適合関数について述べる。

4.2.1 トポロジーに基づく適合関数

分割領域のトポロジカルな特徴量として、“ノード数平均”、“ノード数標準偏差”、“ノード接続度”の3つを定義する。ノード数平均は式(12)、ノード数標準偏差は式(13)で与えられる。また、ノード接続度は、分割領域間の接続領域のノード数平均を表し、式(14)で与えられる。

$$med_{top}(C) = \frac{1}{|C|} \sum_{SA \in C} |SA| \quad (12)$$

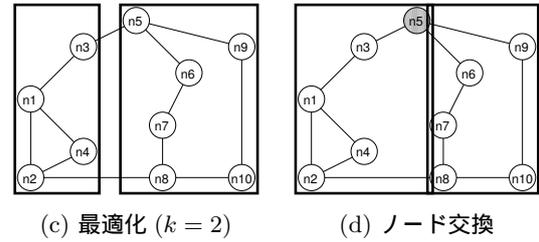
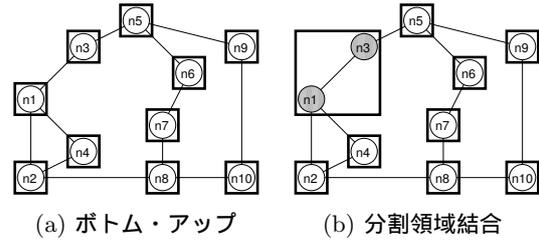


図3 クラスタリング

$$dev_{top}(C) = \frac{1}{|C|} \sum_{SA \in C} \sqrt{(|SA| - med_{top}(C))^2}$$

$$con_{top}(C) = \frac{1}{(|C|^2 - |C|)} \sum_{SA, SA' \in C} |CA_{SA \rightarrow SA'}|$$

以上の3つの特徴量から、トポロジーに基づく適合関数を式(15)で定義する。この適合関数を最小にすることは、分割領域の大きさを均等化し、かつ、リレー輸送の中継位置となる接続領域を最小にすることに等しい。

$$fit_{top}(C) = \frac{dev_{top}(C) + con_{top}(C)}{med_{top}(C)} \quad (15)$$

4.2.2 期待報酬値に基づく適合関数

車両の状態と行動の組に対して定まる期待報酬値 $Q(s_t, a_t)$ を用いて、分割領域の期待報酬値 $Q(SA)$ を、式(16)に示すように、分割領域内のノードで構成可能な、状態と行動の組の期待報酬値の総和で定義する。

$$Q(SA) = \sum_{n, n', n'' \in SA} Q((n, n'), n'') \quad (16)$$

分割領域の期待報酬値 $Q(SA)$ を用いて、“報酬平均”、“報酬標準偏差”、“報酬接続度”の3つの特徴量を定義する。報酬平均は式(17)、報酬標準偏差は式(18)、報酬接続度は式(19)でそれぞれ与えられる。

$$med_{exp}(C) = \frac{1}{|C|} \sum_{SA \in C} Q(SA) \quad (17)$$

$$dev_{exp}(C) = \frac{1}{|C|} \sum_{SA \in C} \sqrt{(Q(SA) - med_{exp}(C))^2}$$

$$con_{exp}(C) = \frac{1}{(|C|^2 - |C|)} \sum_{SA, SA' \in C} Q(CA_{SA \rightarrow SA'})$$

以上の3つの特徴量から、期待報酬値に基づく適合

関数を式 (20) で定義する．この適合関数を最小にすることは，分割領域内を遷移することで得られる報酬を最大にし，かつ，分割領域で得られる報酬を均等化し，さらに，リレー輸送の中継位置となる接続領域で得られる報酬を最小にすることに等しい．

$$fit_{exp}(C) = \frac{dev_{exp}(C) + con_{exp}(C)}{med_{exp}(C)} \quad (20)$$

5. 実験

本章では，2 種類の評価実験における結果を報告する．最初に，1 台の車両のみを用いて，要求分布の学習によって獲得した政策に従って行動する車両と，領域内を最短経路で巡回する車両を比較する．次に，複数台の車両を用いて，領域分割によって車両間でリレー輸送する場合と，領域分割しないで各車両が独立に輸送する場合を比較する．両実験における共通のパラメータ設定を表 1 に示す．

学習率 α		0.1	
割引率 γ	0.5	降車係数 ω_d	1
履歴長 δ	1	刑罰回数 ζ	9
乗車係数 ω_r	1	ランダム行動率 ϵ	0.1
		総要求発生確率 $\sum \eta$	0.3

5.1 学習経路 vs. 巡回経路

ノード数 10，エッジ数 15 で構成されるサービス領域を生成し，分布が異なる 10 パターンのフロー集合 $(F_1, F_2, \dots, F_{10})$ をランダムに発生させた．シミュレーション・サイクルを 20000t に設定し，その前半 10000t を学習に，後半 10000t を評価に用いた．

図 4 は顧客の平均待機時間（要求発生から車両が到着するまでの時間），図 5 は顧客の平均乗車時間（車両の到着から目的地に到着するまでの時間），図 6 は平均総時間をそれぞれ表している．巡回経路は，待機時間のばらつきは小さいのに対し，乗車時間のばらつきが大きい．これは，車両が一定間隔でノードを巡るため，いかなるフローパターンにおいても待機時間は安定するのに対し，乗車時間はノードを巡る順序に大きく影響を受けてしまうからである．一方，学習経路は，フローパターンに合わせて，ノードの到達間隔や順序を変えるため，待機時間，乗車時間の両方においてばらつきは大きくなった．また，総時間に関しては，10 パターン中の 8 パターンにおいて，学習経路が優位性を示した．これは，多くの報酬（顧客の乗降車数）が見込める経路を選択することで，システムの効率向上が期待できることを示している．

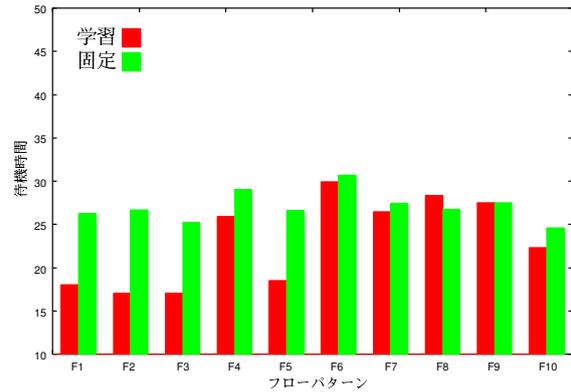


図 4 実験 1: 平均待機時間

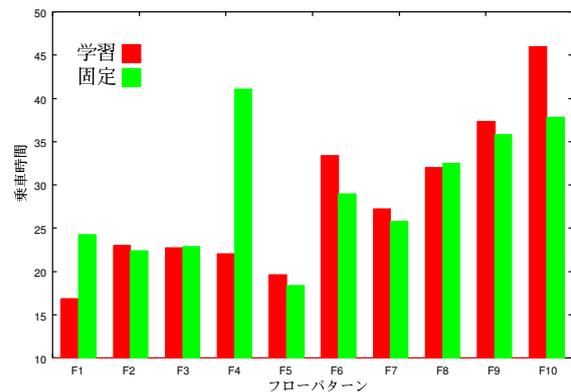


図 5 実験 1: 平均乗車時間

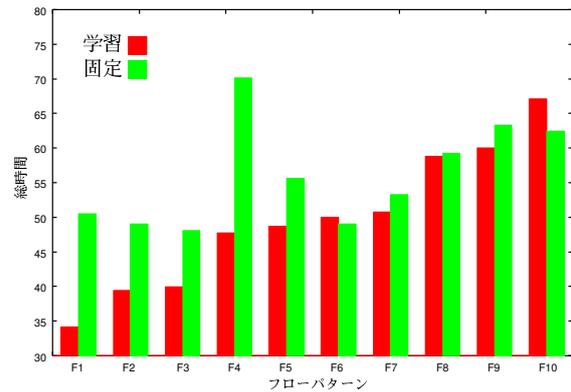


図 6 実験 1: 平均総時間 (平均待機時間+平均乗車時間)

5.2 リレー輸送 vs. 独立輸送

5×5 で構成されるグリッド状のサービス領域を生成し，4 台の車両を配置した．フロー数 ($|F| = 10, \dots, 20$) が異なるフロー集合をランダムに発生させた．シミュレーション・サイクルを 15000t に設定した．最初の 5000t では，領域を分割することなく領域全体を学習させる．次の 5000t では，先に得られた

期待報酬値 (又はトポロジー) に基づきクラスタリングし、分割された領域内を学習させる。最後の 5000t を評価に用いた。また、領域全体の学習時には、割引率 γ を 0.8, ランダム行動率 ϵ を 0.3 とした。

図 7 は顧客の平均待機時間、図 8 は顧客の平均乗車時間、図 9 は平均総時間をそれぞれ表している。領域分割 (クラスタリング) しないで独立輸送する場合、待機時間は小さくなるのに対し、乗車時間は大きくなった。これは、担当領域が広いために、安定して報酬が見込める乗車報酬ばかり学習し、降車報酬を学習できていないことが原因であると考えられる。一方、領域分割によるリレー輸送する場合、両報酬をほぼ平等に学習できるため、待機時間と乗車時間の差は小さい。また、総時間に関しては、リレー輸送がいずれにおいても独立輸送に勝る結果となった。これは、リレーするという負担が発生したとしても、領域分割による協調的輸送が効果的であることを示している。

また、トポロジーに基づく場合と、期待報酬に基づく場合を比較すると、いずれにおいても、トポロジーに基づく方が優位を示した。これは、期待報酬に基づく場合、報酬は各車両に均等に分散することができるが、担当領域の形状が極端に一方に伸びてしまうことが発生し、リレー回数が増加してしまうことが原因と考えられる。さらに、学習における状態履歴が短いため、乗降位置の関連付けが弱く、乗車報酬のみの分割領域や、逆に、降車報酬のみの分割領域が形成されてしまうことも原因と考えられる。今後は、両適合関数を組み合わせることによってリレー回数を減少させたり、乗降車の組を一つの分割領域に含む場合は評価にボーナスを加えるなど、改良を加える予定である。

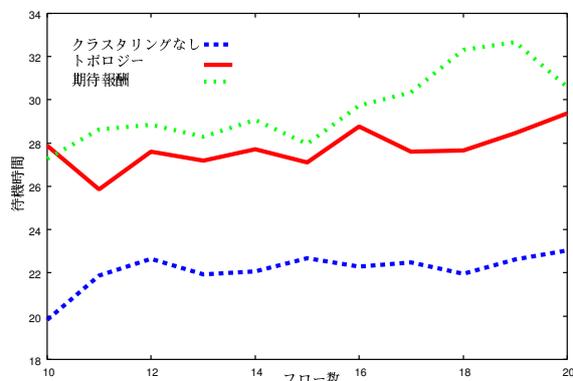


図 7 実験 2: 平均待機時間

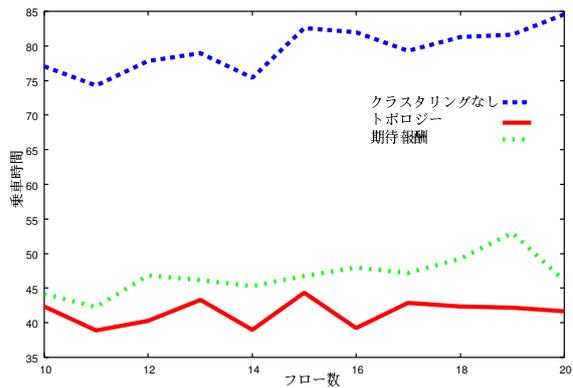


図 8 実験 2: 平均乗車時間

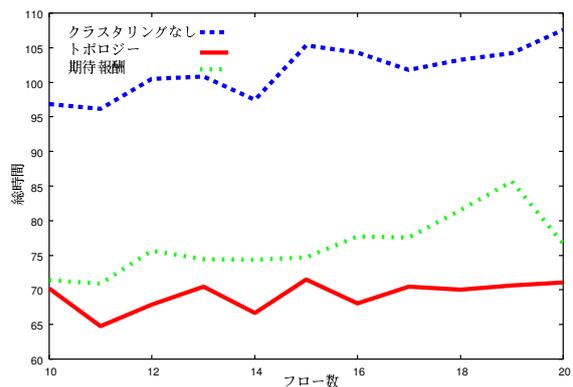


図 9 実験 2: 平均総時間 (待機時間+乗車時間)

6. まとめ

本稿では、強化学習手法 (Q-学習) を用いることで、従来の考え方とは異なる、先見のかつ協調的な輸送サービスを提案した。先見的とは、輸送経験から顧客の発生分布を予測し、将来獲得する報酬が大きくなるような経路選択を行うことである。また、協調的とは、輸送車両間での効果的な負荷分散を行うことである。最後に評価実験を行い、提案手法が輸送システムの効率を向上させることを示した。

今後の課題は、車両の速度や最大乗車数等の能力差を考慮したクラスタリングを導入することである。また、先に述べた巡回分散と領域分散の共存が必要である。

参考文献

- 1) Gambardella, L. M., Éric Taillard and Agazzi, G.: MACS-VRPTW: A Multiple Ant Colony System for Vehicle Routing Problems with Time Windows, *New Ideas in Optimization*

- (Corne, D., Dorigo, M. and Glover, F.(eds.)), McGraw-Hill, pp. 63–76 (1999).
- 2) Louis, S. J., Yin, X. and Yuan, Z. Y.: Multiple Vehicle Routing with Time Windows Using Genetic Algorithms, *Proceedings of the Congress on Evolutionary Computation* (Angeline, P. J., Michalewicz, Z., Schoenauer, M., Yao, X. and Zalzal, A.(eds.)), Vol. 3, Mayflower Hotel, Washington D.C., USA, IEEE Press, pp. 1804–1808 (1999).
 - 3) Mannor, S., Menache, I., Hoze, A. and Klein, U.: Dynamic abstraction in reinforcement learning via clustering., *ICML* (2004).
 - 4) McGovern, A., Sutton, R. S. and Fagg, A. H.: Roles of Macro-Actions in Accelerating Reinforcement Learning, *Proceedings of the 1997 Grace Hopper Celebration of Women in Computing*, pp. 13–18 (1997).
 - 5) Montemanni, R., Gambardella, L., Rizzoli, A. and Donati, A.: A new algorithm for a Dynamic Vehicle Routing Problem based on Ant Colony System, *Proceedings of ODYSSEUS 2003: Second International Workshop on Freight Transportation and Logistics*, Palermo, Italy, pp. 27–30 (2003).
 - 6) Potvin, J.-Y. and Bengio, S.: The vehicle routing problem with time windows — Part II: Genetic search, *INFORMS Journal on Computing*, Vol. 8, pp. 165–172 (1996).
 - 7) Santana, H., Ramalho, G., Corruble, V. and Ratitch, B.: Multi-Agent Patrolling with Reinforcement Learning, *Proceedings of International Conference on Autonomous Agents and Multi-Agents Systems*, pp. 1120–1127 (2004).
 - 8) Sutton, R. S. and Barto, A. G.: *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, MA (1998). A Bradford Book.
 - 9) Thangiah, S.: Vehicle Routing with Time Windows using Genetic Algorithms, *Application Handbook of Genetic Algorithms: New Frontiers, Volume II. Lance Chambers (Ed.)*, CRC Press, pp. 253–277 (1995).
 - 10) 原野貴裕, 石川孝: 協調型デマンドバスの有効性について, 技術報告 2004-ITS-19-18, 情報処理学会研究報告 (2004).
 - 11) 太田正幸, 篠田孝祐, 野田五十樹, 車谷浩一, 中島秀之: 都市型フルデマンドバスの実用性, 技術報告 2002-ITS-11-33, 情報処理学会研究報告 (2002).
 - 12) 野田五十樹, 太田正幸, 篠田孝祐, 熊田陽一郎, 中島秀之: デマンドバスはペイするか?, 技術報告 2003-ICS-131, 情報処理学会研究報告 (2003).