

エージェント強化学習によるダイナミックプライシングを用いたシェアサイクル運用システムの研究

○矢嶋耕平 倉橋節也 (筑波大学)

Modeling of Bicycle Sharing operation system with dynamic pricing by Agent Reinforcement Learning

*K. Yashima and S. Kurahashi (University of Tsukuba)

概要— In recent years, many companies have entered the share cycle business in Japan, Europe, China, and other countries. It is also expected to reduce traffic congestion in urban areas. In the operation of the share cycle business, which allows free access between bicycle parking stations, trucks are dispatched to each parking station to eliminate the opportunity loss due to the uneven distribution of bicycles. The purpose of this research is to encourage users to return to a specific station by paying an incentive, and to verify a business operation model that is fundamentally different from the conventional operation based on truck allocation. We employed a reinforcement learning technique for directions to user agents in order to autonomously respond to uncertain events. The results showed that the proposed model reduced the cost compared to the conventional truck operation. It suggests the usefulness of the autonomous system operation by reinforcement learning.

キーワード: Bicycle Sharing, Reinforcement Learning, Agent Simulation

1 はじめに

シェアサイクルは指定されたエリア内であれば、点在するステーションで自由に自転車の貸出・返却を行うことができるサービスであり、タクシーやバスの代わりとなる公共交通の一つとしても注目を浴びている。シェアサイクルの導入はヨーロッパや中国を中心とする海外をはじめ、国内でも複数の企業が参入しており、GPSデータの解析によるリアルタイムなマーケティングデータとしての活用だけでなく、CO2排出量の削減や市街地における渋滞緩和などの効果も期待されている。

1990年代後半頃からシェアサイクル導入が進んでいるフランスでは、地方自治体から街中での広告パネルの設置権を与えられた企業が、広告料収入を元にユーザーにサービスを提供する仕組みとなっており、広告料収入が得られないシェアサイクル事業単独では、実質的にビジネスが成立していないと考えられる。また、国内においてもシェアサイクルの維持管理を利用料収入で賄うことは困難であり、自治体からの支援を受けているケースも多い。

2 研究の目的

シェアサイクルは利用者の需要の変化に伴い、特定のステーションに自転車が偏在してしまう。そのため、シェアサイクル事業の運営で大きな負担となっているのが、1箇所のステーションに自転車が偏在しないよう、運送トラックが巡回して自転車を回収・再配車する人件費や配送費などの販管費である。利用者の需要はオフィス街や観光地といった状況で変わり、また、利用者個々の特徴によって利用行動も変わるため、適切な再配車を行うことが困難である。そこで、本研究では利用者にインセンティブを提示することで、特定のステーションへの返却行動を促し、従来のトラックによる再配車を主体とする運営とは、抜本的に異なる事業運営モデルの検証を目的とする。検証には利用者個々の行動と環境との複雑な相互作用が起るため、エージェント・シミュレーショ

ンを用いることで、効果の測定・検証への効果が期待される。

3 関連研究

3.1 通勤者の交通行動選択における主要因

Heinen et al.¹⁾は、2008年にオランダの複数の大企業に勤める社員に対してインターネットによるアンケート調査を行い、通勤時の自転車利用に関する主要因を探索的因子分析を用いて推定している。アンケートは-2から+2までの5段階のリッカート尺度が用いられ、バイアスを避けるために自転車利用に限らない、交通行動に対してのアンケート調査として実施された。また、骨子となる概念には、社会的な対人関係からの期待を意味する「主観的規範」や、その交通経路を知っているかという「知覚」、行動することの容易さを表す「行動コントロール」といった要素に加え、「習慣」という要素の4つで構成されている。さらに、常に通勤で自転車を利用する者と稀にしか利用しない者とは、天候や荷物の運搬など、日々変化する多くの要因に対して選択行動に与える影響が異なることが言及されている。因子分析の結果を以下のTable 1に示す。

Table 1: 通勤者の交通行動選択に与える因子分析結果
Factor scores of the attitudes towards characteristics of bicycle commuting.

	Factor		
	Direct benefit	Awareness	Safety
Comfortable	0.712		
Flexible	0.658		
Mentally relaxing		-0.717	
Health benefits		-0.774	
Cheap		-0.559	
Suits lifestyle	0.530		
Physically relaxing	0.313	-0.527	
Environmental benefits		-0.822	
Pleasant/nice	0.554	-0.367	
Offers privacy			0.315
Socially safe			0.917
Time-saving	0.815		
Traffic safety			0.746

Values below 0.3 are not reported.

自転車に関する主な態度を特定するため、共通性が0.3以下の低い特性は省き、3つの因子について言及されている。1つ目は、快適さや時間の節約が高い点数を持つ

「直接的な利益」としている。2つ目は、環境的な利点、健康上の利点、および精神的なリラックスが高い点数を持つ「意識の高さ」としている。3つ目は、社会的安全性と交通安全性が高い点数を持つ「安全性」としている。これらの分析結果より、通勤で自転車を利用する者は、時間、快適さ、柔軟性の面で直接的な利益に基づいて自転車を選択している可能性が高いことが示されている。

3.2 インセンティブに伴う行動変容

Lu et al.²⁾ は、エージェントベースのモデリングを使用して、台北市の通勤者における自転車共有サービスと他の交通手段の選択行動をシミュレートし、自転車共有によるCO₂・温室効果ガス排出量などを改善するための空間マルチエージェントシミュレーションを行っている。また、エージェントが自転車を利用する際の重要な要因には、Heinen et al. の自転車通勤者の主要な関心事は時間や快適さだという主張に加え、Godavarthy et al.³⁾ の寒さ・豪雨・高湿度・暴風などの悪天候はシェアサイクルの利用を減少させ、地域固有の快適な気温が利用を増加させるという主張にも触れている。モデル全体のフレームワークを以下の Fig.1 に示す。

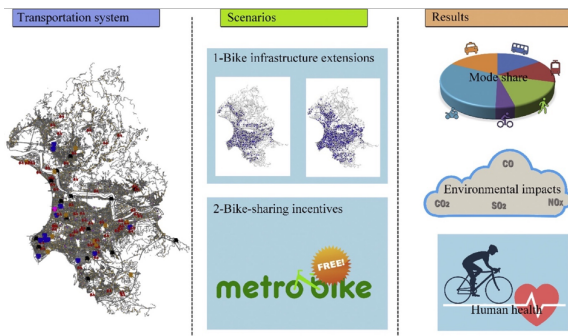


Fig. 1: モデルフレームワーク

通勤のための移動手段として、交通結節点となるバス、メトロの2つに加え、目的地へのラストワンマイルを含む移動が可能な、自転車、徒歩、オートバイ、車の4つを加えた、合計6つの移動手段を想定している。シミュレーションに用いる環境は、バス停やメトロ駅の場所を含む実際の台北市における空間情報が利用されている。

通勤エージェントは、交通手段の選択に影響を与える4つの要因として、移動料金、総移動時間、徒歩か自転車での労力を伴う移動時間、自動車 or バイクの所有権を変数として持っており、各エージェントの社会経済的なステータスから算出される時給を元に時間の価値を定義し、どの移動手段を用いるか決定される設計となっている。

実験のシナリオは3つ設定されており、1つ目は、369箇所のシェアサイクルステーションをバス停の近くに設置する、2つ目は、バスやメトロへの乗り継ぎに利用するシェアサイクル料を無料にする、3つ目は、シェアサイクルの利用料5NTD(約1.66ドル)に対して、利用後に2NTD(約0.66ドル)のクーポンを発行する。いずれも2015年に調査された台北市における通勤時の移動手段の選択割合と比較し、どのシナリオが環境負荷の少ない持続可能な取り組みとなるかを検証している。各シナリオの結果を以下の Table 2 に示す。

Table 2: 台北市のシナリオシミュレーション結果
Simulation results of the two scenarios.

Mode	2015 BAU	Scenario2		
		Scenario1	Infrastructure extensions	Free for transit connection
Bike%	5.40	5.79	6.30	5.60
Walk%	16.40	15.70	20.47	20.00
Motor%	27.30	31.40	24.41	33.60
Car%	16.90	12.40	10.24	12.80
Bus%	17.20	21.49	19.69	14.40
Metro%	16.90	13.22	18.90	13.60

Notes: BAU represents business as usual, and NTD refers to the New Taiwan Dollar.

バス停の近辺に新たにシェアサイクルステーションを設置するシナリオ1では、基準となる2015年の結果と比較すると、シェアサイクルは5.40%から5.79%に増加し、バスの利用が17.20%から21.49%に増加したことにより、バスがメトロのマーケットを奪ったことが示されている。シナリオ2とシナリオ3では、クーポンを発行するシナリオ3より、乗り継ぎ無料のシナリオ2がシェアサイクルの利用を高め、全てのシナリオと比較しても最も環境負荷の低い取り組みとなる可能性が示されている。

本先行研究では、通勤者を対象とした移動手段の選択をシミュレートすることで、環境負荷に対する考察を行っているが、シェアサイクルの再配車問題については言及されていない。また、観光者をモデルとしたシミュレーションへ拡張することに触れており、観光者の時間や費用に対する考え方が通勤者とは異なるため、これらの特徴を踏まえた検討が必要であると言及されている。

4 シミュレーションモデル

本研究では、従来のトラックでの運用とは異なる考案した2つのシナリオについてシミュレーションを行う。シナリオの1つ目は、配車を生業とする個人事業主である事業者に対してのみ配車を依頼する場合である。シナリオの2つ目は、事業者と利用者の2者に対して配車を依頼する場合である。

4.1 ステーション環境

ステーション環境は、オフィス街や観光地といったステーションの立地条件に伴って、様々な利用者の需要パターンが考えられる。本シミュレーションで用いるステーション環境では、特定の状況下における事業運営モデルの検証を行うため、駅から観光・レジャー地域が2km圏内にある立地条件とし、天候は晴れで過ごしやすい外気温の週末を想定した。

シミュレーションを行う環境の概略図として、以下の Fig.2 に配置ステーションの位置関係を示す。環境には4つのステーションが存在し、中央がなだらかな丘となる1辺1kmの正四角形の四隅の位置に配置され、初期状態では各ステーションに10台ずつの自転車が設置されている状況を想定した。また、朝晩の需要がダイナミックに移り変わる環境を模擬するため、ステーションAは鉄道駅、ステーションBは立ち寄り施設1、ステーションCは立ち寄り施設2、ステーションDは観光エリアと仮定し、朝は鉄道駅から観光エリアへ人の移動があり、夜は観光エリアから鉄道駅へ逆向きに移動する需要を想定した。

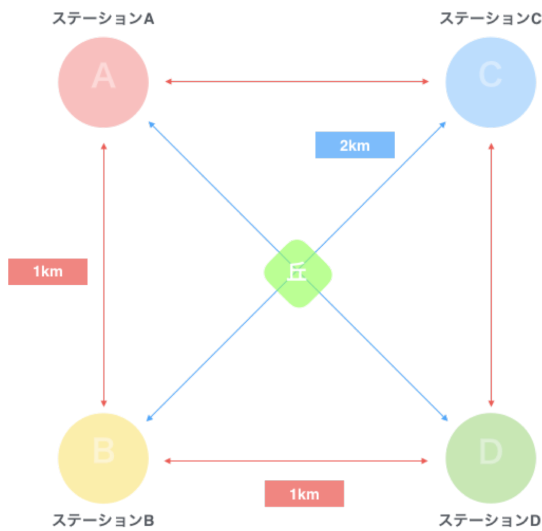


Fig. 2: 配置ステーション概略図.

4.2 需要曲線

ステーションの需要は、朝夕で需要がダイナミックに移り変わり、さらに A と C のステーション残台数が 0 になるよう、各ステーションで AM7:00 から PM23:00 (1 時間毎 17 コマ) までの需要を擬似的に作成した. 各ステーションの 1 時間毎の残台数の推移を以下の Fig.3 に示す.(ただし、残台数を上回る需要があった場合には、マイナスの値として記録している)



Fig. 3: 各ステーションの 1 時間毎の残台数推移.

4.3 全体俯瞰エージェントの行動選択モデル

シミュレーションモデルの全体概要を以下の Fig.4 に示す. 本論文では、利用者や専業者に対して配車依頼を行うエージェントのことを全体俯瞰エージェントと呼称する. 全体俯瞰エージェントには、ある状態 s における行動 a を選択する価値 $Q(s, a)$ を探索的に学習する強化学習を用いる. 本シミュレーションにおいて、状態とは環境から取得される毎時間の各ステーションの自転車残台数であり、行動とは全体俯瞰エージェントが行う「特定のステーションから異なる特定のステーションへの配車依頼」「インセンティブ額」「1 台または 3 台の依頼台数」の 3 つの要素が組み合わされた、利用者または専業者に対する複数パターンの配車依頼のことを意味している.

価値 $Q(s, a)$ の算出には階層型ニューラルネットワークによる関数近似の手法を用いる. 階層型ニューラル

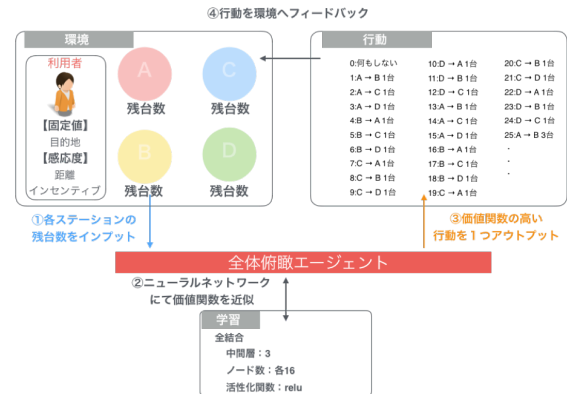


Fig. 4: シミュレーションモデルの全体概要.

ネットワークには、全結合の中間層が 3 層、ノード数を 16 とするネットワーク構成を恣意的に定めた. シェアサイクルのようなステーションの状態と指示する行動の組み合わせが膨大となる場合、全探索するような手法では現実的な時間で計算を終えることができず、学習にはニューラルネットワークを用いた関数近似の手法が有効だと考えられる. 取得された状態 s はニューラルネットワークの入力層への入力値となり、中間層を経て最終的に出力層から出力値として各行動に対する価値 $Q(s, a_1 \dots a_n)$ の見積もりが算出される. 中間層の各ノードでは、出力値を y 、入力値を x_i とした以下の式で順方向に出力が伝播されるようになっており、出力層で出力された価値 $Q(s, a)$ の見積もりと実際の行動から得られた価値 $Q^*(s, a)$ の誤差を少なくするよう、パラメータである重み w_i とバイアス項 b を調整し、実際の価値 $Q^*(s, a)$ を表現できるような関数に近づける処理を行う.

$$y = \varphi\left(\sum_{i=1}^m w_i x_i + b\right)$$

φ は活性化関数であり、本シミュレーションでは以下の式で表される $ReLU$ (Rectified Linear Unit) を活性化関数に用いた.

$$ReLU : \varphi(x) = \max(0, x)$$

5 シミュレーションシナリオ

5.1 シナリオ 1 専業者への配車依頼

利用者とは異なる配車を行うエージェントとして、飲食宅配代行サービス UberEats の個人宅配事業者のような、配車依頼をステーションで待機する存在を想定した. UberEats の個人宅配事業者は、レストランから配達先までの距離 1km あたり 150 円の報酬が設定されている. 本論文のシミュレーションでは、専業者がシェアサイクルの返却行動を行った後、元のステーションまで戻るには徒歩での移動が想定されるため、配車後の帰路にも 1km あたり 150 円のインセンティブを付与し、1km あたり合計 300 円のインセンティブを付与する設計とした.

5.2 シナリオ 2 専門家と利用者への配車依頼

専門家エージェントとは別に、利用料の一部を割引くインセンティブを与えることで、指定のステーションへ返却行動を起こす配車エージェントの1つとして利用者エージェントを想定した。各々の利用者エージェントは、利用を開始する起点ステーションとは異なる、いずれかのステーションを返却する終点ステーションとして持っている。利用者が返却する予定であった終点ステーションと、全体俯瞰エージェントから依頼される終点ステーションの距離や提示インセンティブ額に応じて、利用者の依頼に対する応諾可否が決まる。また、利用者へ支払うインセンティブ額と指定ステーションへの返却依頼に対する応諾率の関係は未知のため、シナリオの2つ目では利用者の応諾率パラメータを変えた2パターンを試行し、どのように挙動が変化するか確認することを目的とした。

利用者エージェントへのインセンティブ額に関しては、ドコモ・バイクシェア社のシェアサイクルの利用料金が30分150円であり、利用者に50円または100円の割引インセンティブを付与することで目的地とは異なるステーションへの返却行動を促すことを想定した。

全体俯瞰エージェントからの依頼に対する利用者エージェントの応諾に関しては、起点となるステーションに需要がある場合に限り、需要の台数内で応諾を行うことが可能であり、利用者エージェントの応諾ロジックをAlgorithm1として以下に示す。

```

Algorithm 1 利用者エージェント応諾ロジック
if 指示する起点ステーションからの貸出需要 > 0 then
  if 利用者の目的地 = 全体俯瞰エージェントが指示する終点ステーション then
    100円ならば100%の確率で応諾
    50円ならば100%の確率で応諾
  else if 利用者の目的地と指示する終点ステーションまでの距離 = 1km then
    100円ならばw%の確率で応諾
    50円ならばx%の確率で応諾
  else 利用者の目的地と指示する終点ステーションまでの距離 = 2km
    100円ならばy%の確率で応諾
    50円ならばz%の確率で応諾
  end if
else 応諾しない
end if

```

利用者エージェントの距離やインセンティブ額に対する応諾率となる、パラメータ w, x, y, z は未知のため、以下の2種類のパターンを用いて実験することで、挙動の変化を確認する。パターンAをパターンBと比べると、距離が1kmでインセンティブ額が50円の時の応諾率は高くなり、距離が2kmでインセンティブ額が100円の時の応諾率は低くなる。したがって、パターンAの応諾率に寄与する選好度は「距離 > インセンティブ額」となり、逆にパターンBは「距離 < インセンティブ額」となるため、両者から選好度の差異による挙動の変化が見られることを期待してパターンAの設計を行った。

パターンA: $w = 80, x = 60, y = 40, z = 30$

パターンB: $w = 80, x = 40, y = 60, z = 30$

6 シミュレーション実験

全ての実験は全体俯瞰エージェントが毎時0分に1つの行動を選択した結果を1ステップとし、AM7:00からPM23:00まで順に17回の行動を選択することが可

能であるため、最大ステップ数は17となる。ただし、いずれかのステーションの残台数が0台になった時点でステップ数と環境をリセットするため、ステップ数が17に届かずリセットされる場合も存在し、ステップ数17とはPM23:00まで残台数が0台にならない状態を全ての時間帯で維持できたことを意味する。また、いずれかのステーション残台数が0台になるか、ステップ数が17になることを1エピソードとし、1エピソード経過する際にはステップ数と環境がリセットされ、新たに次のエピソードとして開始される。本論文では、学習が収束されるであろう10万ステップを1回の実験とし、実験毎の学習のバラツキを考慮するため1シナリオにつき5回の実験を行い、全体俯瞰エージェントが支払った合計インセンティブ額の平均などについて結果を示し考察を行う。

6.1 専門家シナリオの実験結果と考察

全体俯瞰エージェントが専門家へ配車依頼を行う際、起点となるステーションと返却ステーションに加えて、1台か3台の台数指定を行う。また、専門家はインセンティブを得ることを生業としているため、本シミュレーションではステーションの残台数がある範囲内で、専門家は必ず依頼に応諾する想定で実験を行う。

全体俯瞰エージェントが1時間に1つの行動を行った結果を1ステップとし、10万ステップまでシミュレーションを実施した。エピソード毎のステップ数とインセンティブ総額の推移を以下のFig.5に示す。

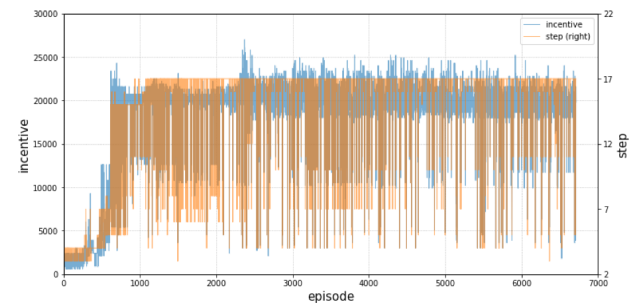


Fig. 5: 専門家シナリオ_ステップ数と支払ったインセンティブ総額の推移。

約500エピソードあたりまではステップ数が7以下となっており、AM7:00からPM13:00までの7時間以内で、いずれかのステーション残台数が0台になり、次エピソードへ移行したことを示している。一方、約1,000エピソード目あたりからはステップ数が17となる場合が頻出しており、学習した全体俯瞰エージェントが上手く配車依頼を出すことで、1日を通して各ステーションの残台数が0台にならないような状態を維持できることが示された。

続いて、各実験の最終100エピソードのうち、残台数が0台となったエピソードを除いた平均値や標準偏差をTable3とFig.6に示す。

Table 3: 専門家シナリオ_実験結果一覧

	実験1回目	実験2回目	実験3回目	実験4回目	実験5回目	全実験合計
データ数	86個	95個	88個	80個	94個	443個
平均	17,892円	20,100円	19,793円	21,450円	19,962円	19,825円
標準偏差	1,030円	1,438円	777円	1,276円	1,102円	1,592円
最小値	15,300円	17,400円	17,700円	18,900円	17,400円	15,300円
最大値	21,300円	23,100円	21,900円	26,100円	23,400円	26,100円

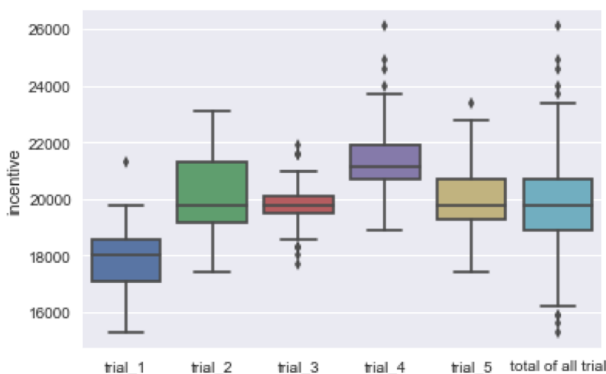


Fig. 6: 専業者シナリオ_箱ひげ図.

全実験結果のインセンティブ額の平均は19,825円、標準偏差は1,592円となる結果であった。また、実験3回目は標準偏差が小さく、実験1回目は最小値が他の実験に比べると15%ほど低い結果となり、最終17ステップ目に到達する割合に関しても実験毎に違いが見られた。

専業者シナリオの実験結果より、各実験で支払うインセンティブ額にばらつきが生じたのは、行動を確率的に選択していることが理由として考えられる。実際の運用時には、ある程度学習が完了した段階で、確率的ではない行動を選択するように変更することで、本実験のような結果のばらつきは起きないと考えられる。したがって、本シミュレーションに用いた需要においては、ステーションの残台数が0台にならない強化学習による運用が実現されており、自律的なシステムでの運用の有用性が示唆された。

6.2 専業者&利用者シナリオの実験結果と考察

専業者エージェントへの依頼は、ステーションに残台数がある限りは確実に指定のステーションへの返却行動を行うが、利用者への依頼に比べると支払うインセンティブ額が高くなってしまふ。一方、利用者エージェントへの依頼は、起点となるステーションに需要がある場合のみ配車を依頼することができるため、ステーション間で需要に偏りがある環境においては、利用者エージェントへの配車依頼のみで残台数を0台にする運用は不可能である。したがって、専業者に加えて、配車エージェントに利用者も加わる想定にて実験を行う。以降、本論文では専業者と利用者の混在シナリオのことを「混在シナリオ」と呼称する。全体俯瞰エージェントが利用者へ配車依頼を行う際、起点となるステーションの利用者に対してのみ依頼し、利用者はパターンA・パターンBの応諾率パラメータに応じて応諾可否を判断する。

パターンAのエピソード毎のステップ数と支払ったインセンティブ総額の推移を以下のFig.7に示す。同様にパターンBについてもFig.8に示す。

パターンA,B共に約2,500エピソード以降から、1日を通して各ステーションの残台数が0台にならないことを意味する、ステップ数が17となるエピソードが頻出しており、専業者モデルと同様に1日を通して残台数が0台にならないような状態を維持できることが示された。また、約1,000エピソード以降から1日を通して残台数が0台にならないような状態が頻出し始めた専業者モデルと比較すると、約2.5倍の学習時間がかかっ

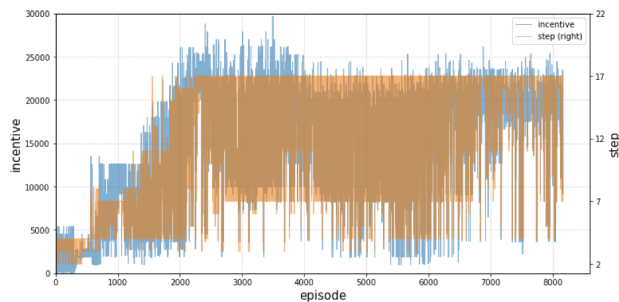


Fig. 7: 混在シナリオ_ステップ数と支払ったインセンティブ総額の推移パターンA.

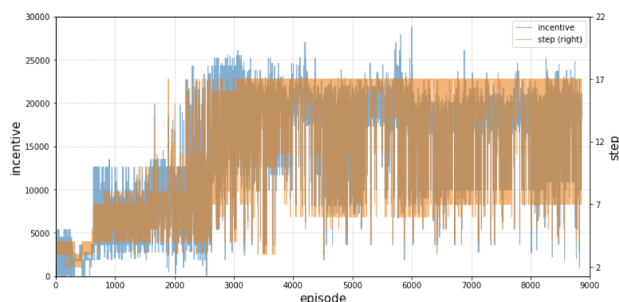


Fig. 8: 混在シナリオ_ステップ数と支払ったインセンティブ総額の推移パターンB.

ていることになるが、専業者モデルの行動集合に利用者モデルの行動集合が増えた分だけ学習に時間を要した結果と考えられる。また、残台数が0になるステーションAやステーションCに対しては、確実に配車される専業者に指示して3台の配車依頼を行っているが、残台数が0台になるリスクの少ない18時以降に関しては、単価の低い利用者エージェントを利用して配車する行動が見られた。

続いて、各実験の最終100エピソードのうち、残台数が0台となったエピソードを除いた平均値や標準偏差をパターンAに関してはTable 4とFig.9に、パターンBに関してはTable 5とFig.10に示す。

Table 4: 混在シナリオ_実験結果一覧パターンA

	実験1回目	実験2回目	実験3回目	実験4回目	実験5回目	全実験合計
データ数	87個	81個	88個	91個	85個	432個
平均	19,641円	20,045円	16,937円	19,236円	17,536円	18,666円
標準偏差	2,095円	1,698円	1,851円	1,132円	1,499円	2,071円
最小値	14,400円	16,600円	14,100円	16,150円	14,150円	14,100円
最大値	24,450円	22,950円	22,150円	22,000円	21,300円	24,450円

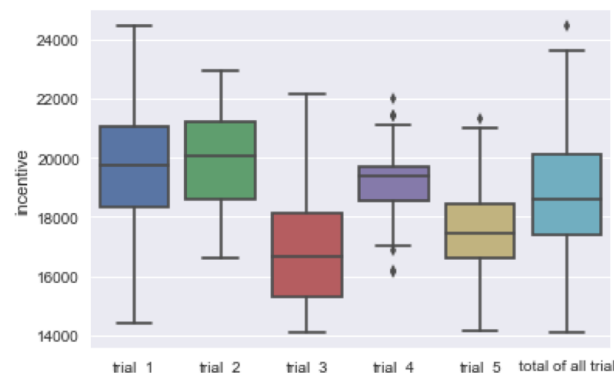


Fig. 9: 混在シナリオ_箱ひげ図パターンA.

Table 5: 混在シナリオ_実験結果一覧_パターン B

	実験 1 回目	実験 2 回目	実験 3 回目	実験 4 回目	実験 5 回目	全実験合計
データ数	85 個	83 個	74 個	87 個	78 個	407 個
平均	17,266 円	16,816 円	18,106 円	17,708 円	19,569 円	17,863 円
標準偏差	1,516 円	1,466 円	1,159 円	1,629 円	1,102 円	1,679 円
最小値	14,100 円	13,200 円	16,550 円	14,700 円	17,650 円	13,200 円
最大値	22,650 円	21,000 円	21,050 円	21,850 円	22,750 円	22,750 円

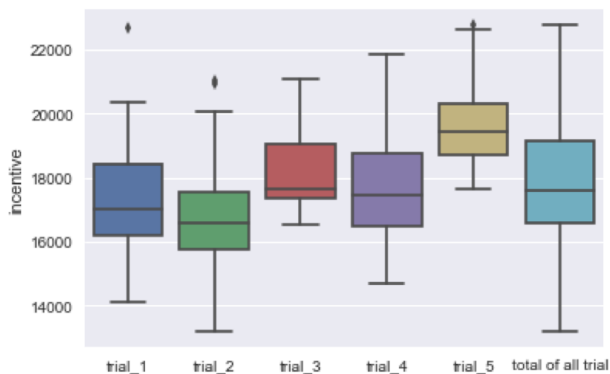


Fig. 10: 混在シナリオ_箱ひげ図_パターン B.

パターン A の全実験結果のインセンティブ額の平均は 18,666 円で標準偏差は 2,071 円となる結果であった。パターン B の全実験結果のインセンティブ額の平均は 17,863 円で標準偏差は 1,679 円となる結果であった。最小値・平均値共にパターン B の方が良い結果となった。

混在シナリオの実験結果より、本シミュレーション環境における需要に関しては、パターン A とパターン B を比較した場合、パターン B の方がインセンティブの最小値・平均値共に低いことが示された。インセンティブによる応諾率が高い (距離 < インセンティブ) パターン B においては、パターン A よりインセンティブ総額が高くなると思われたが、パターン A の全体俯瞰エージェントは応諾率が低い利用者へ依頼して残台数が 0 になるリスクを背負うより、インセンティブが高額でも確実な配車が可能な事業者へ依頼することを学習した結果であると考えられる。また、事業者シナリオと比較した際には、最小値・平均値のいずれにおいてもパターン A とパターン B が良い結果となっており、事業者だけではなく、利用者も加わった混在シナリオが実ビジネスには有用であることが明らかになった。

7 従来のトラック運用との比較

従来のトラックによる運用コストと比較するため、事業者シナリオと混在シナリオのパターン A, B をまとめて比較した結果を以下の Table 6 と Fig.11 に示す。混在モデルのパターン B が最も運用コストが低い可能性があることがわかる。また、トラックによる運用コストの想定を以下に記載する。

Table 6: 全シナリオ_実験結果一覧

	事業者シナリオ	混在シナリオ_パターン A	混在シナリオ_パターン B
データ数	443 個	432 個	407 個
平均	19,825 円	18,666 円	17,863 円
標準偏差	1,592 円	2,071 円	1,679 円
最小値	15,300 円	14,100 円	13,200 円
最大値	26,100 円	24,450 円	22,750 円

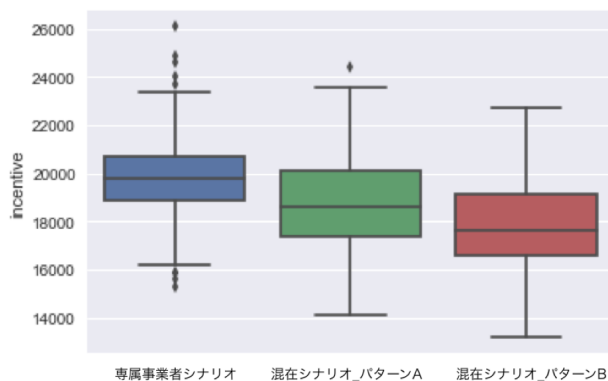


Fig. 11: 全シナリオ_箱ひげ図

トラック運用コスト想定

- ガソリン代金：約 146 円/日
- 配車人件費：16,000 円/日 (作業人員 2 人)
- 自動車維持費用：約 230 円/日 (各種保険等)

トラック運用コストは、1 日あたりに換算すると、合計 16,376 円となり、従来のトラックによる運用と比較した場合、平均値との比較では大きな金額のメリットを提示することはできなかったが、最小値のエピソードでは 13,200 円という結果もあり、従来のモデルより運用コストを下げられる可能性が示された。また、事業者へのインセンティブ額が少なくなることで、事業者が撤退することを想定されるが、その場合は事業者が事業者の変わりとなる運用を想定することも可能である。

8 まとめ

本研究の目的は、従来の人による勘や経験に基づくトラックでの配車を主体とする運営とは、抜本的に異なる事業運営モデルの検証を行うことであった。検証項目の 1 つ目は、事業者や利用者に対する配車依頼を強化学習エージェントが行うことで、ステーションの残台数が 0 台にならない運用が可能であることを確認することであった。2 つ目は、事業者や利用者へ配車を依頼することで、従来の運用と比較してコストの削減や挙動がどのように変化するかを確認することであった。

実験結果より、ステーションの貸出需要が朝夕で変わる本研究のような需要においては、利用者に対する配車依頼だけでは残台数を 0 台にすることはできないが、事業者の存在により、ステーションの残台数が 0 台にならない運用が強化学習を用いて実現できる可能性が示された。また、利用者に対してもインセンティブを付与して配車を依頼することで、事業者のみの運用と比較してもコストを抑えられる可能性が示された。

9 今後の課題と展望

今後の課題として、1 つ目に、ニューラルネットワークの中間層の数やパラメータに加え、強化学習エージェントに対する報酬設計が恣意的になっており、より適切なパラメータや報酬設計を行うことで、さらにコストやステーション残台数が 0 台になる確率を下げる事が期待される。報酬設計にはエキスパートの知見を活かす逆シミュレーションのような手法を用いることも考えられるが、まだ世にない事業モデルにおいて、どのようにエキスパートの知見を得るか検討が必要である。ま

た,全体俯瞰エージェントは同一の需要環境にて学習を行っているため過学習の可能性がある.需要は立地条件と曜日や天候が一緒であれば,一定の分散の範囲内で大まかな需要は変わらないため,本研究の需要で学習した全体俯瞰エージェントを用いて,分散のある需要環境でシミュレーションすることが適切な評価に繋がると考えられる.

2 つ目に,本研究では利用者の応諾率に関するパラメータを距離と金額で表現しているが,天候や高低差に伴う労力が考慮されておらず,限定的な状況でのシミュレーションとなっている.また,需要環境が異なるオフィス街などでは,朝方の通勤時は利用者のインセンティブに対する応諾率が下がることも想定され,時間帯毎に応諾率のパラメータを変更するモデルの拡張も考えられる.いずれにしても,より現実 に即したシミュレーションを行うためには,さらにパラメータを増やした実験が必要であり,パラメータパターンの組み合わせを複数試行することで,現実で想定される挙動を捉え,適切なサービス設計を検討していきたい.

最後に,現実社会においては従来の運用でも残台数が0台になるステーションは存在しており,今後も規模が拡大するステーションの全てをトラックで配車することは現実的とは言えず,特別大きな需要が発生する地域以外は放置されているのが現状である.しかし,専門家や利用者の協力を仰ぐことにより,距離的・時間的な観点からもトラックと比較して柔軟な運用を行える可能性があり,本研究の検証結果が交通のラストワンマイル問題の解消に貢献できる取り組みとなることを期待する.

参考文献

- 1) Heinen, E., Maat, K., & Van Wee, B.:The role of attitudes toward characteristics of bicycle commuting on the choice to cycle to work over various distances, *Transportation Research Part D: Transport and Environment*, **16-2**, 102/109(2011)
- 2) MiaoJia Lu & Shu-Chien Hsu & Pi-Cheng Chen & Wan-Yu Lee:Improving the sustainability of integrated transportation system with bike-sharing: A spatial agent-based approach, *Sustainable Cities and Society*, **41**, 44/51(2018)
- 3) Godavarthi, R. P., & Rahim Taleqani, A.:Winter bikesharing in US: User willingness, and operator's challenges and best practices, *Sustainable Cities and Society*, **30**, 254/262(2017)
- 4) 清水, 赤井, 西野:Modeling and Multi-agent Simulation of Bicycle Sharing, *Serviceology for Services*, **39/46**(2014)
- 5) 前田, 福田:バイクシェアリングにおける自転車再配置車両の経路決定手法に関する研究, *オペレーションズ・リサーチ*, **62-11**,743/755(2017)
- 6) 岸田:日本の自転車交通の現状と改善への取り組み,(財) 国土技術研究センター (2009)
- 7) 永井, 倉橋:賑わいが街を変える- 立ち寄り施設と都市動態モデリング-, *人工知能学会論文誌*, **32-1**,D-G26.1/10(2017)
- 8) 谷口, 高橋:交通行動の居住地選択行動への影響を仮定した 都市動態のマルチエージェントシミュレーション, *計測自動制御学会論文集*,**47-11**,571/580(2011)
- 9) 山分, 黒江, 飯間:マルチエージェントタスクに対する群強化学習法, *計測自動制御学会論文集* **49-3**, 370/377(2013)
- 10) 荒井:逆強化学習によるマルチエージェント系の報酬設定, *計測と制御*,**52-6**,534/539(2013)
- 11) M. Dell'Amico & E. Hadjicostantinou & M. Iori & S. Novellani: The bike sharing rebalancing problem: Mathematical formulations and benchmark instances,*Omega*,**45**, 7/19(2014)
- 12) 伊藤:マルチエージェントの自動交渉モデルとその応用, *知能と情報*, **22-3**,295/302(2010)
- 13) 伊藤:エージェント技術に基づく新しい社会システムの創成に向けて, *市民共創知研究会資料/Proceedings of SIG-CCI*,**1-8**,1/6(2016)
- 14) 山本, 中山, 北村:再配車を用いない複数ステーション型自動車共同利用システムの挙動に関するシミュレーション分析, *土木学会論文集*,**786-67**,11/20(2005)
- 15) 池田, 小川, 末田, 新田:契約ネットプロトコルにおける入札ルールの学習, *情報処理学会第 68 回全国大会講演論文集*,229/230(2006)
- 16) 丸山, 松田, 長谷川, 有村:データマイニングアプローチによるコミュニティサイクルの利用動態の抽出, *土木計画学研究・論文集 第 31 巻 (特集)*, **70-5**,671/680(2014)
- 17) 池田, 伊庭:n 人繰り返し囚人のジレンマゲーム戦略の GA による進化, *情報処理学会研究報告*,**127**,191/198(2002)
- 18) 諏訪, 高見, 大森, 原田:バイクシェアリングシステムの計画・評価手法に関する一考察-パリの Velib' における検討事例を踏まえて-, *土木計画学研究・講演集*,(2009)
- 19) 木村, 宮崎, 小林:強化学習システムの設計指針, *計測と制御*,**38-10**,618/623(1999)
- 20) 荒井:マルチエージェント強化学習: 実用化に向けての課題・理論・諸技術との融合, *人工知能学会誌*,**16-4**,476/481(2001)
- 21) 森山:囚人のジレンマゲームにおける Q 学習による協調の維持, *コンピュータソフトウェア*,**25-4**,145/153(2008)

- 22) Jen Te Pai&Shih Ying Pai:User Behaviour Analysis of the Public Bike System in Taipei,International Review for Spatial Planning and Sustainable Development,**3-2**, 39/52(2014)
- 23) Tal Raviv& Ofer Kolka:Optimal inventory management of a bike-sharing station,IEE Transactions, **45-10**,1077/1093(2013)