

移動時間と待機時間に対するドライバーの嗜好を考慮した タクシーの配車目的地設定の提案

長谷川 美夏^{1,a)} 川又 裕也^{1,b)} 河部 瞭太¹ 野々口 大幹¹ 稲葉 陽子¹ 生田目 崇²

受付日 2021年1月29日, 再受付日 2021年4月22日,
採録日 2021年5月10日

概要: 近年, あらゆる分野でビッグデータの活用が進んでおり, 様々な形でデータが蓄積されている交通分野においてもデータ活用が期待されている. なかでも自動車のプローブデータは従来の道路にセンサを取り付けた交通システムとは異なり, 各車のリアルタイムの情報を用いることによる交通状況の把握や分析のほか, 混雑予測や安全対策など様々な活用方法が期待されている. 本論文では, 都内エリアのタクシーのプローブデータを用いて, 深夜帯を対象に都内主要駅のタクシープールにおける平均待ち時間, 平均乗車数, 平均利用距離を求めて, タクシーが目的地まで移動する時間と利用者を獲得するまでに発生する待ち時間に対するドライバーの嗜好を考慮し, 最も利益が期待できる配車目的地を提案するための効用関数を設定する. この効用関数に対して, 深夜時間帯の配車の最適化を行い, その結果を地図上に可視化する. 最適化した結果を用いて, とある事業所を対象に深夜帯に売上がないタクシーが推奨目的地に向かい利用者を拾えた場合, 最大でひと月約 19.5%の売上増が見込めることを示唆できた.

キーワード: タクシープローブデータ, 効用関数, 最適化

Recommendation of Taxi Destination Considering Driver Preferences for Moving and Waiting Time

MIKA HASEGAWA^{1,a)} YUYA KAWAMATA^{1,b)} RYOTA KAWABE¹ TAIKI NONOGUCHI¹ YOKO INABA¹
TAKASHI NAMATAME²

Received: January 29, 2021, Revised: April 22, 2021,
Accepted: May 10, 2021

Abstract: In recent years, the use of big data has been advancing in all fields, and the field of transportation is also attracting attention because various data according to vehicle are accumulated. In particular, automobile probe data is expected to be used in a variety of ways, such as for understanding and analyzing traffic conditions using real-time data of each automobile in stead of the existed transportation systems which embedded various sensors on road. In this paper, we use the probe data of cabs in the Tokyo area to determine the average waiting time, average number of rides, and average distance used in cab pools at major stations in Tokyo, focusing on the midnight period. An utility function of taxi driver is settled to suggest the most profitable destination. The utility function is used to optimize the allocation of vehicles during late night hours, and the results are visualized on a map. Using the optimization results, we show that if a cab, which has no sales during the midnight hours, can go to the recommended destination and pick up a user, it could increase sales by about 19.5% per month at most.

Keywords: taxi probe data, utility function, optimization

¹ 株式会社 NTT データ
Toyosu Center Building Annex, Koto, Tokyo 135-8672,
Japan

² 中央大学理工学部
Faculty of Science and Engineering, Chuo University,
Bunkyo, Tokyo 112-8551, Japan

1. はじめに

近年, あらゆる分野においてビッグデータの活用が進ん

^{a)} Mika.Hasegawa@nttdata.com

^{b)} Yuya.kawamata@nttdata.com

であり、交通分野においても注目されている。交通分野では、1990年代後半からプローブデータの活用に関心が向けられている。たとえば、自動車のカーナビから得られるプローブデータとして、時刻、緯度、経度、加速度などの走行や挙動に関わるデータが取得できる。プローブデータは、交通状況の把握や分析のほか、混雑予測や安全対策への活用が期待されている [1]。

さて、昨今の交通分野においてはタクシー業界を取り巻く状況が著しく変化している。1つは、スマートフォンなどの普及により、タクシーの配車アプリが活用される機会が増えたことである。従来、利用者が多い都心部においては「流し」と呼ばれる街中を走りながら利用者を探す営業方式が主流であり、これは開催されているイベントや時間帯により需要を予測できるドライバにとっては大きな利益を生む方式である。また、「付け待ち」と呼ばれるタクシープールなどで待機して利用者を待つ営業方式もある。こちらは、待ち時間のロスはあるものの利用者獲得の確実性が高い方式であり、主に終電前後などにおいて利用者からの需要が高まる傾向にある。流しであれば移動コスト、付け待ちであれば待機コストはタクシー業界目線では売上が発生していない時間であるため、配車アプリによりこれらの時間の削減が期待される。

加えて、2020年以降においては感染症流行の影響により、飲食業界の営業時間短縮や、リモートワークの増加にともない終電時間の変更を検討する鉄道会社が見受けられる。感染拡大後、鉄道利用客は3割程度減少したが、深夜だけに限ると6割以上が減少しており、この需要が回復するのは難しいと見込まれている [2]。これにともないタクシー利用者も減少している一方で、終電後の時間帯については利用者目線では依然としてタクシー以外の交通機関の利用はほぼできない状況である。したがって、公共交通機関が使えない時間帯の利用者からの需要に確実に応えることはタクシー業界にとって重要であると考えられる。

そこで本研究では、都内タクシーのプローブデータを対象に、都内エリアにおいてタクシー事業者が、タクシーをどのエリアに配車すべきかの提案を、タクシードライバの働き方の嗜好情報に基づいて行う。まず、タクシー事業者にとっては「利用者の利用距離が長いほど利益が多い」、「空車時間が短いほど損失が少ない」という前提を置く。したがって、タクシー事業者にとっては「現時点からなるべく短時間で長距離移動が見込める利用者に乗せられること」が理想状態であると仮定する。具体的には、利用者に乗せている時間を利益、利用者に乗せずに走行および待機している時間を損失と定義し、1人の利用者に対して最も大きな利益が見込める場所を配車目的地とする。また、ドライバの嗜好情報については、終電前後のドライバの行動パターンにおいては営業所付近で素早く少額利益を得たい場合と、最終利用者を遠くまで乗せて時間はかかるが大き

な利益を得たい場合があると仮定し、利益を多くすることと損失が少ないことのどちらを重視するかという係数を嗜好情報と定める。以上を用いて、各ドライバの働き方の嗜好を反映した最適なタクシープールを選択するような効用関数を設定し、事業所のすべてのタクシーで効用を最大化するように配車の最適化を行う。また、最適エリアの分布や嗜好の変化による推奨先の違いについて可視化を行う。

2. 関連研究と本論文の関係

自動車のプローブデータを利用した研究はここ10年で数多く公表されている。乗車需要に関する研究については以下のようなものがある。Yaoら [3] はタクシーの需要予測を深層学習の枠組みで行っており、他の機械学習手法や多変量解析手法よりも高精度な予測ができたことを報告している。落合ら [4] は、乗合型であるデマンド応答型交通を対象にコスト関数を定義したうえでマルチエージェントシミュレーションを用いて、導入効果について検証している。Afianら [5] は、待ち行列理論を応用し、プローブデータから拾い切れていない需要についての予測を行っている。

タクシードライバに関しては、金月ら [6], [7] がドライバの行動特性を考慮したシミュレーションを行っており、文献 [6] では、付け待ちか流しかによる行動の違いを評価しており、文献 [7] では、典型パターンを設定したうえでのシミュレーションを行っている。

最適化に目を向けると、三輪ら [9] は、GPSの誤差を考慮しながら、走行経路の最適化を割当問題として定式化している。

ここで取り上げた論文以外にも数多くの車両プローブデータを用いた研究はされている。

本論文では、ある時点で都内で稼働しているタクシーを対象としている。都内タクシープローブデータを用いた既存研究としてはMukaiら [10] による天候などを考慮したエリア別の需要予測などもあるが、本論文では効率的なタクシープールへの割当てを行うという点でこれらの研究とは異なる視点を持っている。特に、タクシープールまでの移動について、各タクシープールの売上特性とそこまでの移動コストを考慮した効用関数を導入し、全体最適の観点での割当てを行う点は既存の研究との視点の違いがある。本論文の方法の意義の1つとしては、各車両の配車をエリア全体で行うことで、どのタクシープールがどのくらいのエリアを念頭に置けばよいのかを、そのときの状況などから把握できることにある。

3. 分析方針

3.1 分析準備

本研究においては、各ドライバに対して終電後の時間帯に「付け待ち」をするのに適した都内エリアのタクシー

プールを提案することを目的とする。本論文では、経営科学系研究部会連合協議会主催「令和元年度データ解析コンペティション*1」にて提供された、みずほ情報総研株式会社の都内(23区+武蔵野・三鷹)タクシープローブデータを用いる。本分析で用いたデータ項目はプローブデータ取得時間、客対応状態(実車、空車など)、無線協組ID(タクシー事業所ごとに付与されるID)、乗務員コード(ドライバーごとに付与されるID)、緯度、経度の6種類である。データ特性として、プローブの取得間隔は無線協組IDにより約20秒から約50秒と異なっていたため、時間計算の際は考慮している。分析対象は、朝6:00から翌朝6:00までを1日として2017年3月の平日5日間について4週間分の計20日分、約310M行のタクシープローブデータを対象に実施する。タクシー総数はドライバー単位で集計し20日間でのべ約266k人(実際のドライバー数は約23k人)、タクシー事業所は7事業所が対象である。

集計時間の単位はタクシーの勤務形態により決定した*2。対象となるタクシープールは利用者が多い都内エリア主要駅から29カ所を用いる。複数のタクシープールを持つ駅については、それぞれについて集計している。

3.2 分析指標

効用関数を設定するにあたって、各タクシープールについて集計期間の20日間を対象に、30分おきの「平均待ち時間」「平均乗車数」「平均利用距離」を集計する。本研究ではタクシーの状況について次のように定義する。タクシープールの中心座標から半径50m以内を「タクシープールに入っている」とし、タクシープールに入ったタクシーについて半径200m以内で客対応状態が変化した場合、タクシープール内でステータスが変化したととらえる。また、タクシープールに入ったタクシーはすべて利用者を乗せる意志があると仮定している。

- 平均待ち時間 \bar{w}_{p_j} : ある30分間についてタクシープール p_j に入ったタクシーが利用者を乗せるまでにかかった時間を待ち時間と定義する。したがって、タクシープールに入った各タクシーについて、タクシープール範囲内における空車プローブ数と各社ごとのプローブ取得間隔平均を乗算した結果の平均をとったものを平均待ち時間と定義する。このとき、ある30分間の集計範囲に含まれる時間内にタクシープールで待ち始めたが、乗車がなかった場合、タクシープールに入ってから集計範囲の最後の時間までの時間をその集計範囲の待ち時間として計上し、さらに待ち時間が続く場合

は、次の集計範囲の待ち時間として計算している*3。

- 平均乗車数 \bar{r}_{p_j} : 客対応状況が空車 → 乗車に変化した場合「駅のタクシープールから利用者を乗せた」と定義する。したがって、ある30分間についての客対応状況が「空車」から「乗車」に変化した数の平均をとったものを平均乗車数とする。
- 平均利用距離 \bar{d}_{p_j} : ある30分間について利用者を乗せたタクシーが走行した距離の平均を平均利用距離と定義する。このとき、距離計算はプローブごとの座標変化距離の合計により行う。

3.3 効用関数

次に、本研究で使用する効用関数について述べる。タクシープール p_j の平均利用距離 \bar{d}_{p_j} は、利用者を乗せたときに期待される運賃と相関を持つため長いほど利益が増大すると考えられる。一方で、タクシープール p の平均待ち時間 \bar{w}_{p_j} は、利益が得られない時間であるため長いほど損失が増大すると考えられる。今回は平均待ち時間 \bar{w}_{p_j} は、距離換算するためタクシーの平均移動速度に基づき時速20kmを乗算する[11]。

また、 k はタクシードライバーの選好を反映した変数として定義した値であり、移動や待ち時間を損失であるとして与える度合の大きさを表している。

つまり、 k が大きいドライバーほど、次の利用者を早く拾えることを優先し、 k が小さいドライバーほど、より長距離の走行が見込める利用者を拾えることを優先する。これらから、タクシー t_i がタクシープール p_j に向かうときの効用 $u(t_i, p_j)$ を以下のように定式化する。

$$u(t_i, p_j) = \bar{d}_{p_j} - k(\bar{w}_{p_j} \times 20 [\text{km/h}] + L(t_i, p_j)) \quad (1)$$

$L(t_i, p_j)$: タクシー t_i とタクシープール p_j の距離

タクシー t_i がタクシープール p_j に向かうかどうかを $x(t_i, p_j) \in \{0, 1\}$ と表すとき、それぞれのタクシーが向かうことができるタクシープールは最大で1つであるという制約条件が課される。また、タクシープールにおいて乗車したい利用者の数は有限であるため、配車するタクシーには上限を設ける必要がある。そこで、過去の乗車数・各駅乗車比を元に計算した、各駅の配車上限数をキャパシティ $C_{p_j} = \frac{\bar{r}_{p_j}}{\sum_j \bar{r}_{p_j}}$ としたとき、それぞれのタクシープールで客を拾えるタクシーの数はキャパシティ C_{p_j} 以下になるという制約条件が課される。

以上をふまえると最大化する目的関数は、

$$\sum_{i,j} u(t_i, p_j) x(t_i, p_j) \quad (2)$$

*1 <https://jasmac-j.jimdofree.com/>

*2 タクシードライバーの勤務形態には、朝から夕方まで勤務する昼日勤、夕方から深夜まで勤務する夜日勤、丸1日勤務し翌日は休暇となる隔日勤務の3パターンがある。勤務時間の目安は昼日勤は7:00~16:00、夜日勤は17:00~26:00、隔日勤務は8:00~26:00である。

*3 22:25にタクシープールで待ち始め、22:40に乗車客を乗せたタクシーがあった場合、22:00~22:30において待ち時間が5分のタクシーが1台、22:30~23:00において待ち時間が10分のタクシーが1台とカウントしています。

となり、制約条件は、

$$\sum_i x(t_i, p_j) \leq 1 \wedge x(t_i, p_j) \in \{0, 1\} \quad (3)$$

$$\sum_j x(t_i, p_j) \leq C_{p_j} \quad (4)$$

となる。

4. 分析結果

定義した分析指標により、駅およびタクシープールに特徴が現れるかどうかを集計する。まず、29カ所のタクシープールに対して、そのエリアの特徴の分析を行った。次に、複数のタクシープールがある場合の結果を集計し、駅ごとの大きな特徴をとらえる。本研究では深夜帯のタクシーの動向に着目するため、22時から26時（午前2時のこと、本稿では適宜30時間制表記を用いる）の時間帯について分析する。

4.1 時間帯におけるタクシープールの需要の推移

横軸を平均待ち時間、縦軸を平均利用距離、プロットの大きさを乗車数として時間帯ごとの各エリアの特徴を図1、図2、図3、図4に示す。プロットは大きいほど乗車客が多いことを示す。この図においては、横軸は右にあるほどすぐに利用者を獲得できていることから、供給されているタクシーに対して利用者の需要が大きいエリア、すなわちさらなるタクシーの供給が見込めるエリアといえる。また、縦軸は上に行くほど利用距離が大きいことから、一度の乗車で大きな売上を期待できるエリアといえる。プロットは大きいほど乗車客が多いことから、よりタクシー利用者の多いエリアであるといえる。以上より、右上に大きなプロットがあるほど大きな収益効率が期待できるエリアで

あると考えられる。

図を見るとほとんどのエリアは22~26時にかけて右上に推移していくが、プロットのサイズは23~25時がピークであることが可視化されている。以上より、右上にあるエリアほどタクシーの追加供給先として期待が高いと考えられる。

4.2 収益効率を元にしたタクシープールの傾向分析

さて、図1から図4に示された結果からエリアを分類するために、各エリアの収益効率の推移を考える。各エリアについて30分ごとに乗車状態になったタクシーに関する、待ち時間と走行時間に対する運賃の総計を収益効率と定義

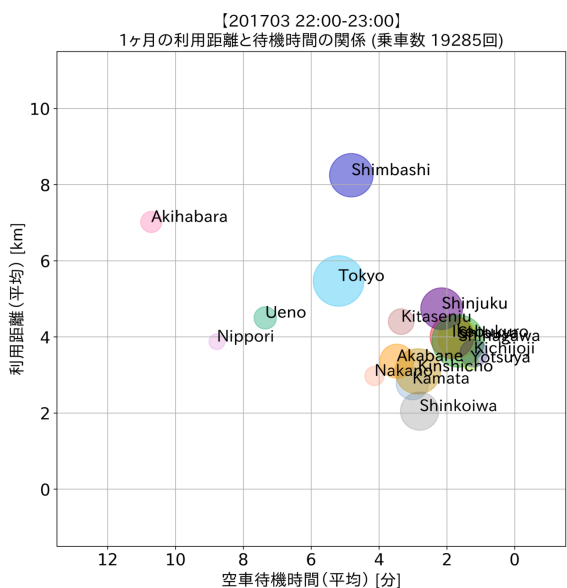


図1 22~23時における駅の特徴
Fig. 1 Features of the station at 22:00-23:00.

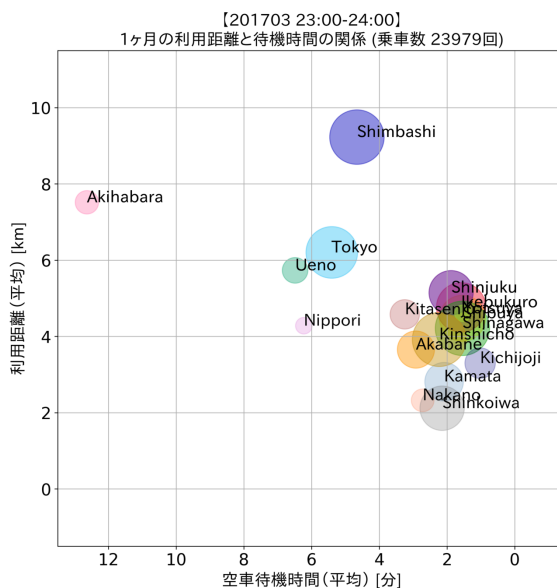


図2 23~24時における駅の特徴
Fig. 2 Features of the station at 23:00-24:00.

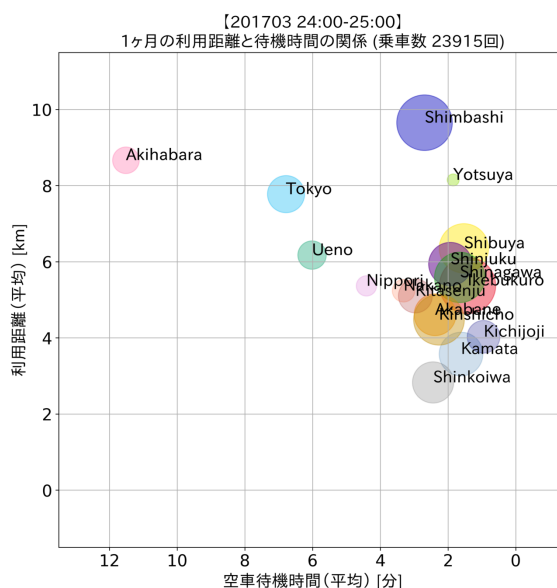


図3 24~25時における駅の特徴
Fig. 3 Features of the station at 24:00-25:00.

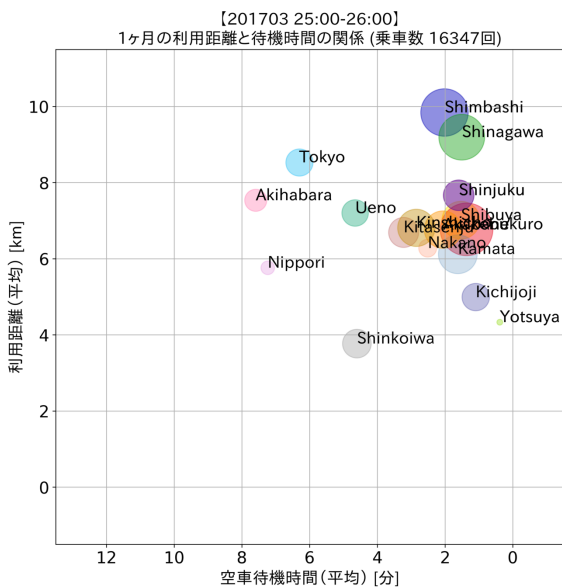


図 4 25～26 時における駅の特徴
Fig. 4 Features of the station at 25:00–26:00.

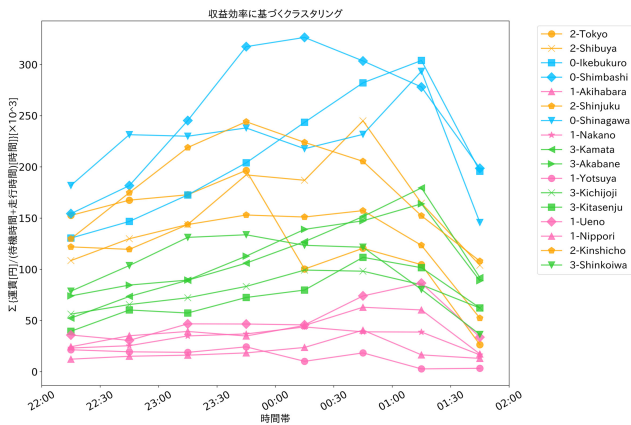


図 5 深夜帯タクシーの収益効率
Fig. 5 Profit efficiency of taxis at midnight.

する。深夜帯における収益効率の推移を分類するため、各タクシープールの収益効率特徴量を 22 時から 26 時まで 30 分刻みの 8 次元の配列として定義する。この収益効率について深夜帯の推移を k-means 法で 4 つのクラスに分類した結果を図 5 に示す。k の値については、オフィス街、住宅街などエリアの特徴について仮説立てを行い、 $k = 2 \sim 5$ 程度で分類した結果、最も解釈性が良いと判断した値を採用している。これより、全体を通して深夜帯のどこかで収益効率は増加したあと急激に低下する傾向があること、収益効率の高いエリアと低いエリアがあること、ピークを迎える時間帯が異なるエリアがあることが可視化されている。クラスタリングの結果、以下のような特徴を持つエリアに分類された。それぞれについて考察する。

- クラス 0 (需要大繁華街エリア)：24～26 時あたりの深夜帯に収益効率が最大値を発揮するエリア。オフィス街に近かつ居酒屋などの多い繁華街であり、新橋、品川については神奈川方面、池袋については埼玉方面

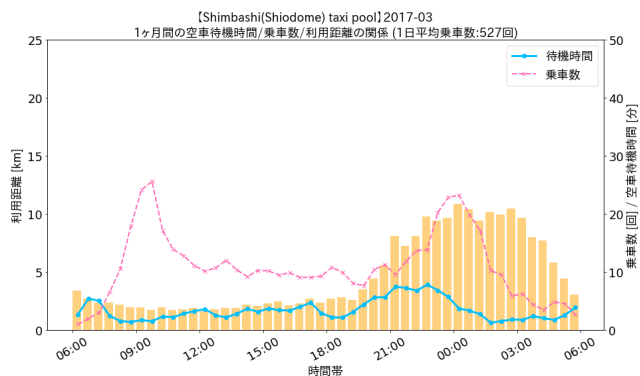


図 6 新橋(汐留)の 1 日の待ち時間・乗車数・利用距離の推移
Fig. 6 Daily waiting time, number of rides, and ride distance in Shimbashi (Shiodome).

の長距離輸送路線があることから郊外在住の利用者が多いことが見込める。

- クラス 1 (都心部近距離利用エリア)：25 時あたりで収益効率が最大値となるが、全体を通して大きな需要はないエリア。秋葉原、中野、四ツ谷、上野、日暮里はどれも山手線内の駅であり、比較的短距離の利用者が多い。
- クラス 2 (都心部繁華街エリア)：24～25 時あたりの深夜帯に収益効率が最大値を発揮するが、25 時以降急激に収益効率が低下するエリア。東京、渋谷、新宿、錦糸町と都心部の繁華街であり、終電付近での利用者の増加が見込める。
- クラス 3 (郊外エリア)：25 時あたりで収益効率が最大値となるエリア。蒲田、赤羽、北千住、新小岩など、山手線の外側で比較的用户の多いエリアである。

収益効率が良い時間帯は、利益が見込めるため配車を増やすことで効率化が期待できる。

このうち以下では近隣でありながらクラスの異なる新橋(汐留)タクシープールと東京(八重洲口)タクシープールについて比較する。

4.2.1 クラス 0：新橋(汐留)タクシープール

クラス 0 は深夜帯の利益が最も見込めるクラスであり、そのなかでも特に特徴が顕著であるのは、図 6 に示す新橋駅(汐留)タクシープールであった。全時間帯を通じた結果を見ると、他タクシープールに比べて平均利用距離が長く、平均待ち時間が短いという傾向が見られた。特に 23～25 時においては平均乗車数、平均利用距離が双方伸びており、1 日のうちで大きな利益が見込める時間帯といえる。一方で、20～24 時にかけては平均待ち時間が多く、需要に対してタクシーはやや過剰に共有されていると考えられる。乗車数が増え始める 23 時ごろと比較して 24 時ごろになると平均待ち時間は減少していることから、タクシーの需要がコンスタントに見込めると判断できる。このことより、新橋(汐留)エリアにおいては平均利用距離と平均乗

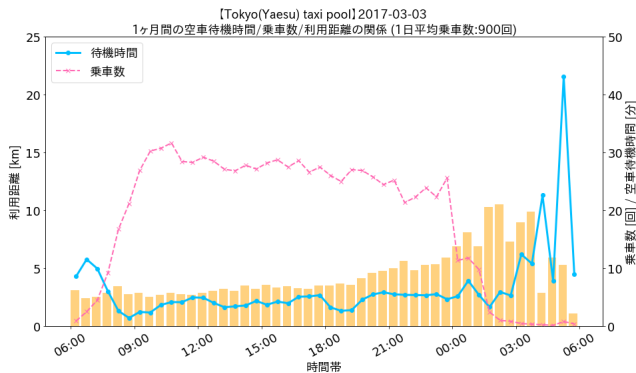


図 7 東京（八重洲口）の1日の待ち時間・乗車数・利用距離の推移
 Fig. 7 Daily waiting time, number of rides, and ride distance in Tokyo (Yaesu).

車数が増加しているかつ待ち時間が少ない比較的活発な24時以降の時間帯に配車を行うことで、利用者の需要にこたえることができると考えられる。

4.2.2 クラス2：東京（八重洲口）タクシープール

クラス2は深夜帯の利益が最も見込めるものの、クラス0よりも早い時間帯である25時付近から急激に需要が落ち込む特徴にあるエリアである。最も顕著な特徴を持つのは、図7に示す東京（八重洲口）タクシープールであった。全時間帯を通した結果を見ると、他タクシープールに比べて日中の利用者が多く、安定傾向にある。一方で日中の利用距離は5km以内と比較的短距離であり、1回の乗車あたりの利益が見込めるのは24~26時付近である。東京駅に関していうと、遠方からの新幹線利用客とオフィス利用客の異なる需要が重なり合った結果であると考えられる。したがって、深夜帯には日中ほどの乗客数を見込むことはできない。このことより、東京（八重洲口）エリアにおいては日中と深夜帯では配車の戦略を変更するべきであると考えられる。

5. 最適化によるタクシーの目的地設定の提案

5.1 ドライバの嗜好による推奨エリアの変化

本節では、式(2)~(4)による効用関数を用いた分枝限定法による厳密解法での最適化を実施し、結果を可視化する。本研究においては、各ドライバーに対して全体的に利用者が多い終電後の時間帯（25~26時と定義）に「付け待ち」をするのに適した都内エリアのタクシープールを提案することを目的とする。そのために、都内エリアを500m四方のメッシュでエリアに区切り、ドライバーの嗜好を全体で固定して、それぞれ最適化を行った結果を図8、図9、図10に可視化する。図はタクシープールごとに色分けされており、塗りつぶされていない範囲は利益の見込みがマイナスになるため目的地として推奨されるタクシープールがない状態を示す。また、色が濃いほど大きな利益が見込めることを示す。図8は、 $k = 0.8$ の場合であり、業務時間に余

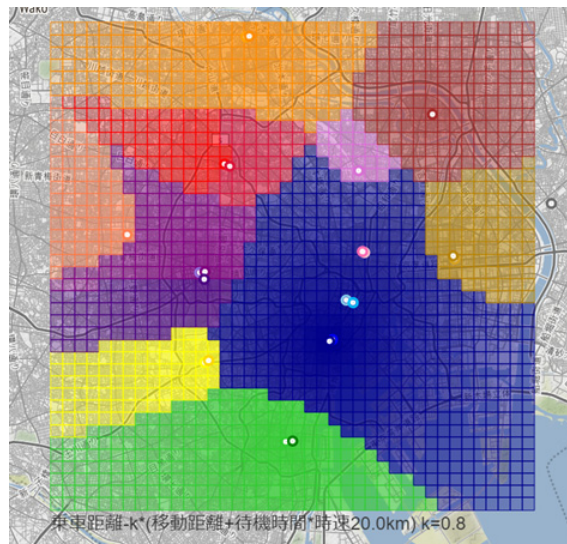


図 8 $k = 0.8$ における推奨エリア
 Fig. 8 Recommended area at $k = 0.8$.

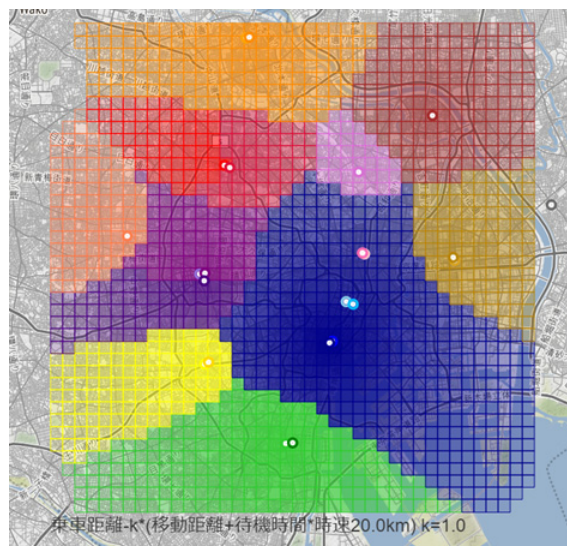


図 9 $k = 1.0$ における推奨エリア
 Fig. 9 Recommended area at $k = 1.0$.

裕があったり長距離利用者の見込める目的地を優先するという選好が反映され、新橋（青）、品川（緑）などクラス0のエリアを推奨する範囲が大きく出ている。図10は、 $k = 1.5$ の場合であり、なるべく利用者を早く獲得することを優先するという選好が反映され、近隣のエリアに集中するようになっている。 k の値が小さい場合に比べてそれぞれの推奨エリアが円に近く、長時間かけて移動する必要がある遠方のタクシープールは推奨されにくいということが分かる。特に、新宿（紫）と池袋（赤）の間のエリアに注目すると、推奨エリアはおよそ二地点の中間部を境界として分割されているが、中野（左部オレンジ）と新宿（紫）は、新宿の方が利益率が圧倒的に良いため新宿を推奨するエリアのほうが広がっている。以上より、ドライバーの働き方の嗜好を反映した最適なタクシープールのエリアを可

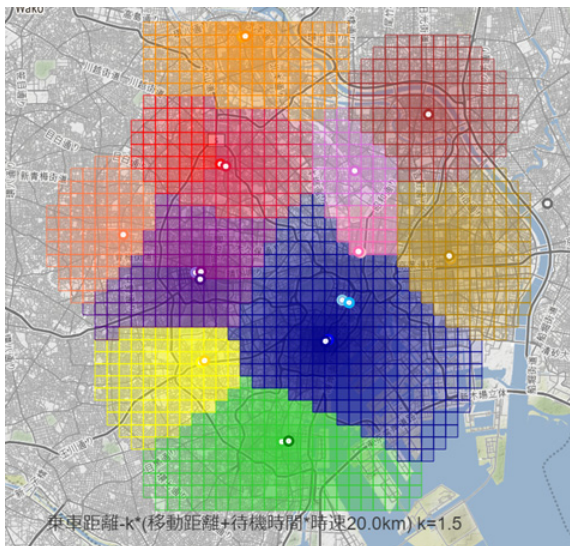


図 10 $k = 1.5$ における推奨エリア
 Fig. 10 Recommended area at $k = 1.5$.

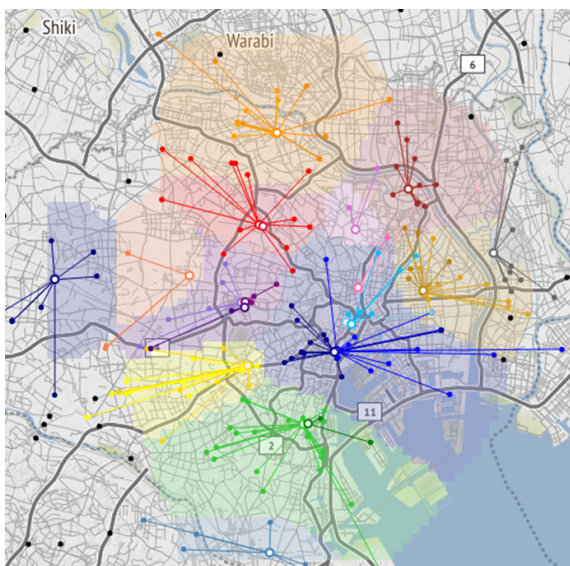


図 11 ドライバの選好をランダムにした場合の全体最適化
 Fig. 11 Total optimization when driver preferences (random).
 視化することができた。

ただし、実際はドライバーは複数存在しており、かつ、ドライバーごとに k の値、すなわち働き方の選好は異なっていると考える方が自然である。また、タクシープールにもキャパシティが存在しているため、1カ所に推奨が集中してしまう場合は、より効用が高いタクシーを採用し、効用が低いタクシーについては次点となるような推奨先に割り振る方が望ましい。したがって、効用が低いタクシーは図 8, 図 9, 図 10 に示すような、推奨エリアの通りの推奨先を得られない可能性がある。そこで、25 時時点で空車のドライバーに対し、ドライバーごとに k の値をランダムに付与し、最適な目的地を推奨した結果を図 11 に可視化する。すると、推奨エリアをまたいで移動するドライバーを確認できる。それぞれのドライバーの選好にフィットするかつ、全

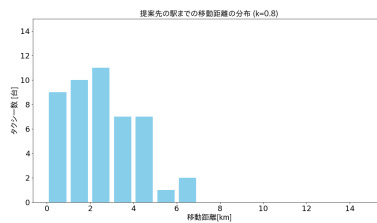


図 12 $k = 0.8$ における推奨先との距離
 Fig. 12 Distance to the target at $k = 0.8$.

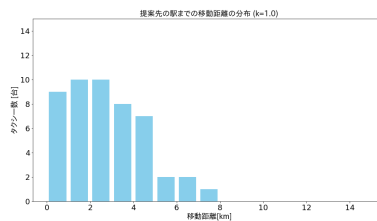


図 13 $k = 1.0$ における推奨先との距離
 Fig. 13 Distance to the target at $k = 1.0$.

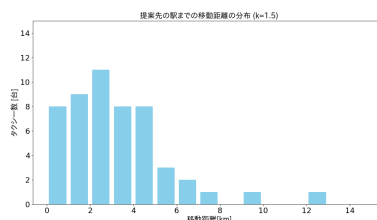


図 14 $k = 1.5$ における推奨先との距離
 Fig. 14 Distance to the target at $k = 1.5$.

体を最適化するような目的地推奨結果が可視化された。新橋 (青)、品川 (緑) などは推奨エリア内のタクシーの割当てが多いが、赤羽 (オレンジ) のようにキャパシティが小さいタクシープール付近では、タクシーが密集していたとしても別のタクシープールを推奨できる。図 11 の場合は、池袋 (赤) への推奨が多い。

また、各車両の振舞いについて比較する。

図 12, 図 13, 図 14, 図 15, 図 16, 図 17, 図 18, 図 19, 図 20 はそれぞれが左から $k = 0.8, 1.0, 1.5$ と変化させたときの、モデルで提案された行先のタクシープールまでの距離、総移動距離、運賃収入である。外れ値の影響を軽減するため、集計範囲である 20 日分を対象とし、累計を集計した。

k の値が大きくなると、相対的にタクシープールまでの距離に対する効用が低くてもよくなるため、推奨先への移動距離のビンは $k = 0.8, 1.0$ と比較すると長距離を示す階級に出現している。それにともない、総移動距離に関しても、 k を大きくするとヒストグラムのビンは $k = 0.8, 1.0$ と比較すると長距離を示す階級に出現する。運賃収入については大きな差はないように見えるものの、最適な推奨先が存在したタクシー台数の増減の影響が表れている。

また、図 21, 図 22, 図 23 はドライバーに対して $k = 0.8$

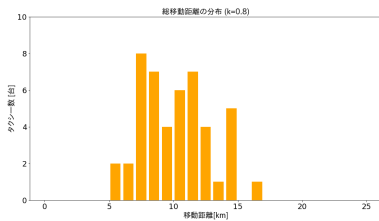


図 15 $k = 0.8$ における総移動距離
Fig. 15 Total move distance at $k = 0.8$.

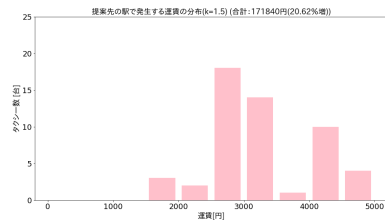


図 20 $k = 1.5$ における運賃
Fig. 20 Fares at $k = 1.5$.

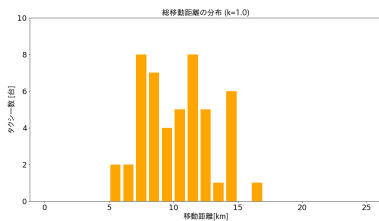


図 16 $k = 1.0$ における総移動距離
Fig. 16 Total move distance at $k = 1.0$.

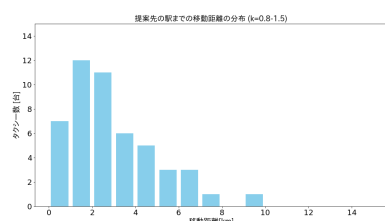


図 21 $k = 0.8 \sim 1.5$ における提案先との距離
Fig. 21 Distance to the target at $k = 0.8 \sim 1.5$.

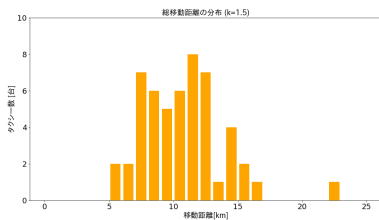


図 17 $k = 1.5$ における総移動距離
Fig. 17 Total move distance at $k = 1.5$.

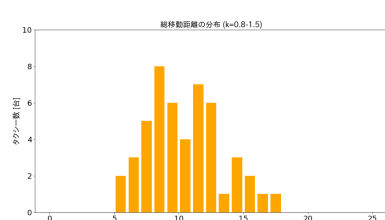


図 22 $k = 0.8 \sim 1.5$ における総移動距離
Fig. 22 Total move distance at $k = 0.8 \sim 1.5$.

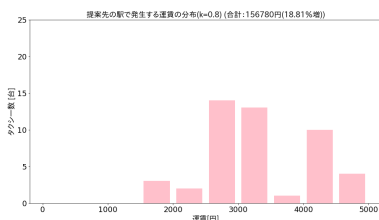


図 18 $k = 0.8$ における運賃
Fig. 18 Fares at $k = 0.8$.

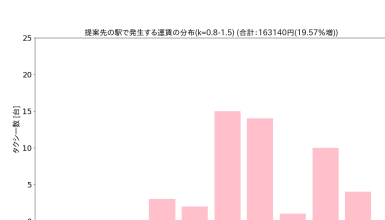


図 23 $k = 0.8 \sim 1.5$ における運賃
Fig. 23 Fares at $k = 0.8 \sim 1.5$.

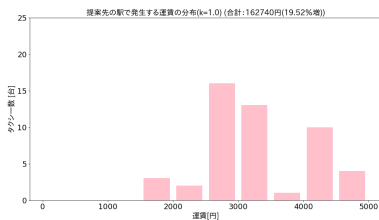


図 19 $k = 1.0$ における運賃
Fig. 19 Fares at $k = 1.0$.

と 1.5 をランダムに付与した場合の結果である。移動距離については選好がランダムである方が、分布の形は正規分布に近いならかな分布となり、移動距離に対する選好が異なるドライバーにおいてもそれぞれのニーズに合致した行き先提案ができていることが見て取れる。運賃収入につい

ては上述のとおり配車数の影響を受けていると考えられる。

5.2 最適化によるタクシー事業者への影響

最適化を実施した際に見込める売上を計算し、その効果を評価する。運賃の導出は都内深夜料金の計算方法に則り、走行距離 842m 以下を初乗り 420 円とし、それ以降を 178m ごとに 80 円加算とする。今回は、7つあるうちの 1つの事業所を対象に、25 時以降に売上 0 円で業務時間終了まで走行を続けていたドライバーを最適化対象とする。具体的には、対象事業所に所属しており 25 時以降半分以上のプロープデータが空車かつ一度も乗車ステータスになっていないまま走行していたタクシーを対象とする。タクシーのステータスが、回送、休憩、閉局など乗車客を受け入れ

表 1 1つの事業所の最適化効果

Table 1 Optimization effect of a taxi business office.

	日付	最適化数	最適化対象数	全ドライバー数	売上向上率
平均	—	49.15	67.2	373.3	18.62%
最大	3/20	59	79	297	37.9%
中央値	3/14, 28	54	73.5	380	19.47%
最小	3/24	12	27	448	2.86%

られない状態になっている場合は売上を得る意思がないと判断できるため、今回は対象にしていない。

この事業所の規模は対象の時間帯に稼働しているタクシーのうち約28%のドライバー数を占めている。これはこの時間帯においては最も大きい規模の事業所である。対象のタクシーを最適化計算先に分配し、平均利用距離を元に予想売上を計算する。このとき、仮にすべてのタクシーが目的地で利用者を獲得できた場合最適化を行わない場合の1日の売上平均が約95万円であったのに対して、最適化結果に従って配置した場合は売上が0のタクシーが利用者を獲得でき約111万円程度まで1日の売上平均を伸ばせるという計算結果が得られた。したがって、最適化によって1カ月の売上を最大で約19.5%増量することが見込めた。最適化の対象となるドライバーの数は平均68.4名で、そのうち実際に現在地から提案可能なタクシープールが存在していたドライバーは平均49.2名であった。これより、提案先がないドライバー、すなわちどのタクシープールに向かっても移動コストの方が期待できる売上よりも大きくなってしまいうドライバーは早期に業務を終了させることでガソリンの節約を視野に入れられる。プローブデータのみから実際に乗車に至らなかった需要数であるキャパシティの情報を得るのは難しいため、本論文では図4に基づき、多くのエリアでタクシーが待機している時間が5分以内と比較的短いことから各タクシープールには深夜帯はほとんど乗車待ちの客がいるという前提のもと平均乗車数の割合から期待できるキャパシティを計算している。仮に、正確なキャパシティを用いて最適化を行うことができればより確信度の高い結果が得られる。

6. おわりに

本研究では都内主要タクシープール29カ所を対象に待ち時間・利用距離・乗車数の3つに着目して分析を行った。分析結果を深夜帯における駅エリアの特徴を30分ごとのドライバー時給の総計推移により分類した結果、ピークのはやい繁華街エリア、ピークの遅い繁華街エリア、山手線内エリア、山手線外エリアに分類されるということを可視化した。また、待ち時間・利用距離・乗車数の3つに着目すると終電前後の時間帯に需要のピークがあることが可視化できた。待ち時間・利用距離・乗車数を元に駅のキャパ

シティとドライバーの選好を考慮した効用関数を設定し、配車の最適化を行い、その結果を可視化した。結果、効用関数を用いて、25時以降収益を得ていないタクシーについて最適化を行うと、売上を最大約19.5%増やすことが可能という結果が得られた。

以上より、着目する評価指標として待ち時間・利用距離・乗車数を用いることは、エリアを特徴づける際に有効であることが確認できた。本研究では2017年3月のデータを用いているが、終電時間や人流は日々変化していくものであるため、新しいデータを用いて同様の分析をすることで柔軟な予測や傾向を知る手掛かりになると考えられる。

参考文献

- [1] 渡部大輔, 佐治秀剛, 田中良寛, 鹿野島秀行, 牧野浩志: 道路プローブデータの分析による有効な活用方法の検討, 土木計画学研究・講演集, Vol.50, 8 pages (2014).
- [2] NHK news web: 利用客66%減の衝撃! 相次ぐ終電繰り上げの影響は, 入手先 (<https://www3.nhk.or.jp/news/html/20201022/k10012675891000.html>) (参照 2020-12-01).
- [3] Yao, H., Wu, F., Ke, J., Tang, X., Jia, Y., Lu, S., Gong, P., Ye, J. and Li, Z.: Deep multi-view spatialtemporal network for taxi demand prediction, 2018 AAAI Conference on Artificial Intelligence (AAAI '18) (2018).
- [4] 落合純一, 金森 亮, 松島裕康, 野田五十樹, 中島秀之: タクシー配車データを用いたリアルタイムデマンド交通システムの実用性評価, 第53回土木計画学研究発表会・講演集, pp.1328-1334 (2016).
- [5] Afian, A., Odoni, A. and Rus, D.: Inferring unmet demand from taxi probe data, Proc. 2015 IEEE 18th International Conference on Intelligent Transportation Systems, pp.861-868 (2015).
- [6] 金月寛彰, 服部宏充: プローブカーデータを利用したタクシードライバーの個人特性の分析とモデル化, 第2015年度人工知能学会全国大会 (第29回), セッション1N4-4, pp.1-4 (2015).
- [7] 金月寛彰, 服部宏充: マルチエージェントシミュレーションによるタクシー営業戦略の改善シナリオの提案, 人工知能学会論文誌, Vol.34, No.3, pp.C-IA2.1-9 (2019).
- [8] 大原 誠, 玉置 久: 客獲得数向上のためのタクシー車両配車問題のモデル化と目標エリア決定ルールの構成, 電気学会論文誌C (電子・情報・システム部門誌), Vol.132, No.10, pp.1645-1655 (2012).
- [9] 三輪富生, 石黒洋介, 山本俊行, 森川高行: 情報の信頼性と収集頻度を考慮したタクシープローブカーの確率論的最適割当計画, 土木学会論文誌, Vol.65, No.4, pp.465-479 (2009).
- [10] Mukai, N. and Yoden, N.: Taxi Demand Forecasting Based on Taxi Probe Data by Neural Network in Intelligent Interactive Multimedia: Systems and Services, Watanabe, T., Watada, J., Takahashi, N., Howlett, R.J. and Jain, L.C. (Eds.), pp.589-597 (2012).
- [11] 総務省: 東京都区内走行調査 (2008), 入手先 (https://www.soumu.go.jp/main_content/000041633.pdf) (参照 2020-12-01).



長谷川 美夏

2019年早稲田大学大学院基幹理工学研究科情報理工・情報通信専攻修士課程修了。同年株式会社エヌ・ティ・ティ・データ。自然言語や画像を対象としたAI技術の研究開発に従事。



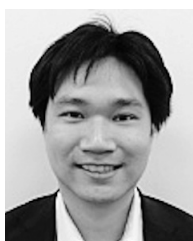
生田目 崇

1999年東京理科大学大学院工学研究科経営工学専攻博士後期課程修了，博士（工学）。2013年より中央大学理工学部経営システム工学科教授。マーケティング・サイエンス，経営科学の研究に従事。



川又 裕也

2019年東京大学大学院新領域創成科学研究科複雑理工学専攻修士課程修了。同年株式会社エヌ・ティ・ティ・データ。量子コンピュータの実活用に向けた研究開発に従事。



河部 瞭太

2019年早稲田大学大学院創造理工学研究科経営システム工学専攻修士課程修了。同年株式会社エヌ・ティ・ティ・データ。商用サービス向けアーキテクチャ提案活動に従事。



野々口 大幹

2019年徳島大学大学院先端技術科学教育部システム創生工学専攻修士課程修了。同年株式会社エヌ・ティ・ティ・データ。次世代クラウド会計システム開発に従事。



稲葉 陽子

2001年東京大学理学系研究科地球惑星科学専攻修士課程修了。同年株式会社エヌ・ティ・ティ・データ。データアナリティクス・AI等の研究開発，および，顧客向けのコンサルティングに従事。