Vol.2021-MPS-132 No.1
2021/3/1

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]

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．中でも自動車の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など様々な活用方法が期待されている．本稿では，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，もっと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作成する．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 Kawabe[1]　Taiki Nonoguchi[1]　Yoko Inaba[1]
Takashi Namatame[2]

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, predicting congestion, and taking safety measures. 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 created 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, Optimization

[1] 株式会社 NTT データ
Toyosu Center Building Annex, 3-3-9, Toyosu, Koto-ku, Tokyo 135-8672, Japan
[2] 中央大学理工学部
Faculty of Science and Engineering, Chuo University, 1-13-27, Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan
[a)] Mika.Hasegawa@nttdata.com
[b)] Yuya.kawamata@nttdata.com

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でおり，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．例えば，自動車のカーナビから得られるプ

ⓒ 2021 Information Processing Society of Japan

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．一つは，スマートフォン等の普及により，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ー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，リモートワーカーの増加に伴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]．これに伴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るべきかの提案を，タクシードライバー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」という前提をおく．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を乗せられること」が理想状態であると仮定する．具体的には，利用者を乗せている時間を利益，利用者を乗せずに走行および待機している時間を損失と定義し，一人の利用者に対してもっとも大きな利益が見込める場所を配車目的地とする．また，ドライバーの嗜好情報については，終電前後のドライバー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，各ドライバー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作成し，事業所の全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．乗車需要に関しての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．タクシードライバーに関しては，金月らが [5], [6] ドライバーの行動特性を考慮したシミュレーションを行っており，[5] 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，[6] 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 [8] 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 [9] 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り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一つとしては，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，その時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，各ドライバー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．利用するタクシープローブの情報は，経営科学系研究部会連合協議会主催「令和元年データ解析コンペティション」にて提供されたデータであり，プローブデータ取得時間，客対応状態（実車，空車など），タクシー会社 ID，ドライバー ID，緯度，経度の 6 種類である．また，プローブ間隔の取得間隔がタクシー

会社により異なっていたため，時間計算のためにそれぞれのプローブ取得間隔の平均を計算する．朝 6:00 から翌朝 6:00 までを 1 日として，2017 年 3 月から平日各曜日 4 日ずつ計 20 日分のデータを用いて集計をする．集計時間の単位はタクシーの勤務形態により決定した．[*1]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 か所を用いる．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

効用関数作成にあたって，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捉える．また， タクシープールに入ったタクシーは全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 $\bar{w}_p$: ある 30 分間についてタクシープール p 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取ったものを平均待ち時間と定義する．
- 平均乗車数 $\bar{r}_p$: 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→乗車に変化した数の平均を取ったものを平均乗車数とする．
- 平均利用距離 $\bar{d}_p$: 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．この時，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，作成する効用関数について述べる．タクシープール p の平均利用距離 $\bar{d}_p$ 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 $\bar{w}_p$ 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 $\bar{w}_p$ 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 [10].

また，k はタクシードライバーの選好を反映した変数であり，移動や待ち時間を損失であると捉える度合の大きさにあたる．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．

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

$L(t_i, p_j)$: 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 がとあるタクシープール p_j に向かうかどうかを $x(t_i, p_j) \in \{0, 1\}$ と表すとき，1 台のタクシーが向かうことが出来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，配車上限値の概念が必要になる．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 $C_{p_j} = \frac{\bar{r}_{p_j}}{\sum_j r_{p_j}}$ 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．

以上を踏まえると最大化する目的関数は，

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

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 か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捉える．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 1~4 に示す．この図においては，横軸は右にあるほど直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

[*1] 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．

シーに対して利用者の需要が大きいエリア，すなわち更なるタクシーの供給が見込めるエリアと言える．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言える．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 時-26 時に掛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 1,2,3,4 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 分毎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 つのクラスタに分けた結果を図 5 に示す．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，収益効率の多いエリアと少な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も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 時 ~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．オフィス街に近い且つ居酒屋等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 時 ~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利用者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．

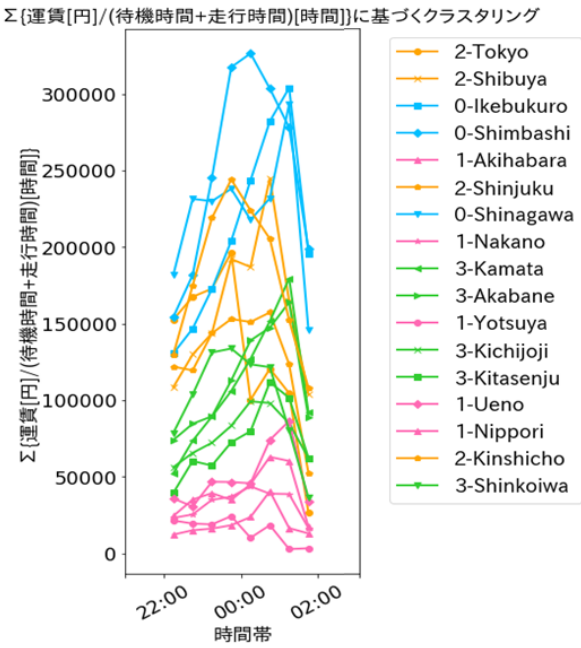

図 5 深夜帯タクシーの収益効率

5. 最適化によるタクシーの目的地設定の提案

5.1 ドライバーの嗜好による推奨エリアの変化

本章では，効用関数を用いた最適化の結果を可視化する．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 時 ~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 6~8 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 6 は，$k = 0.8$ の場合であり，業務時間に余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 8 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，長時間かけて遠いタクシープールに移動しようとしていないことが見て取れる．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くなっている．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視化することが出来た．

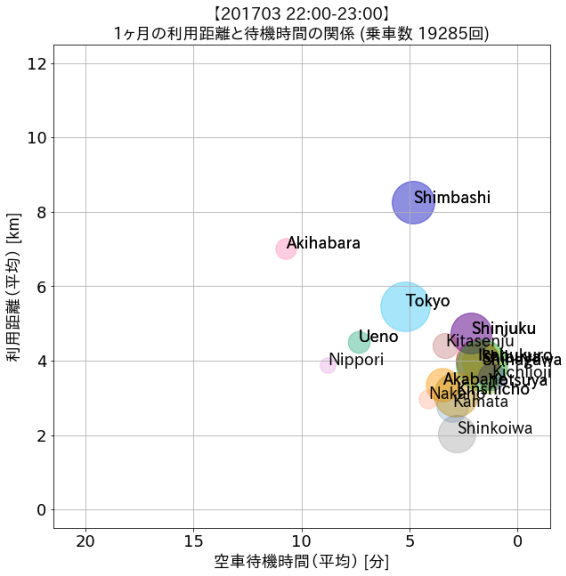

図 1　22 時 ~23 時における駅の特徴

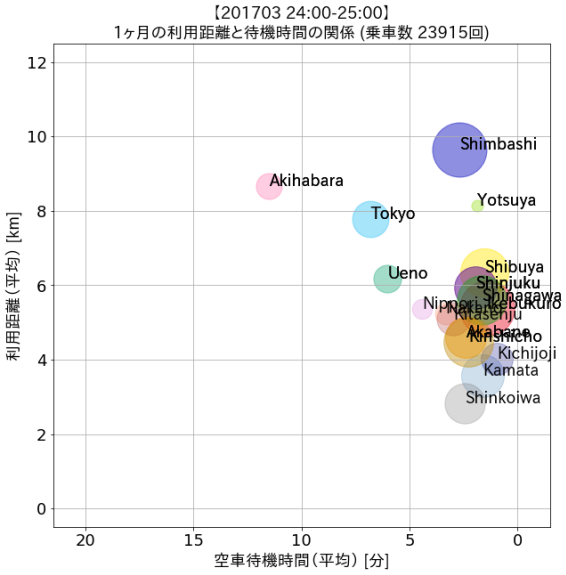

図 2　24 時 ~25 時における駅の特徴

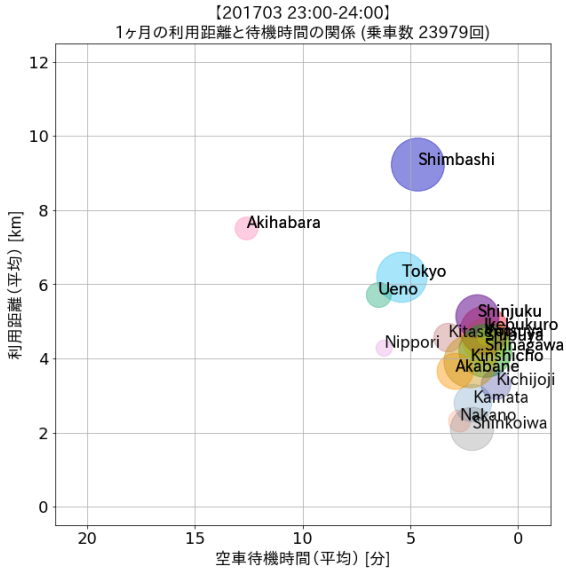

図 3　23 時 ~24 時における駅の特徴

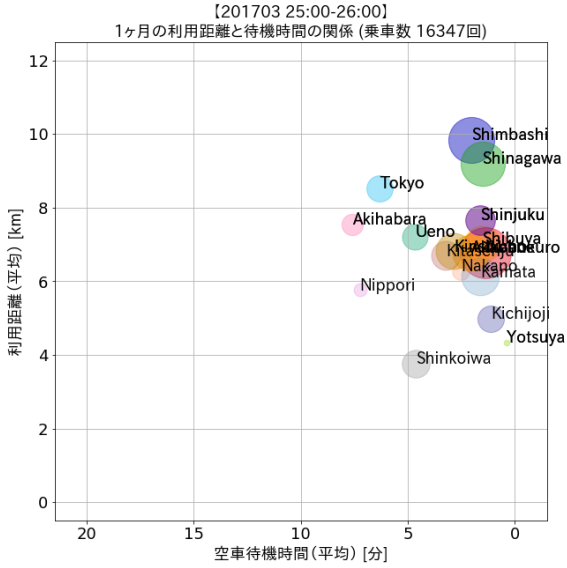

図 4　25 時 ~26 時における駅の特徴

ただし，実際はドライバーごとに働き方の選好は異なっていると考える方が自然である．そこで，ドライバーごとに k の値をランダムに付与し，ドライバーごとに最適な目的地を推奨した結果を図 9 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り当てが多いが，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 9 の場合は、池袋（赤）への推奨が多い．

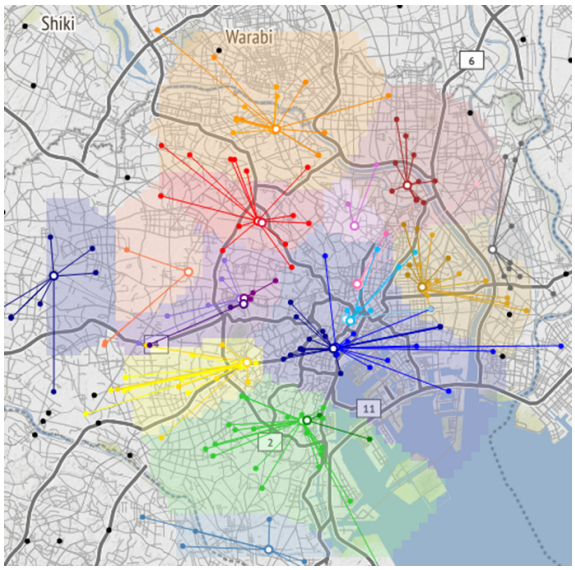

図 9　ドライバーの選好をランダムにした場合の全体最適化

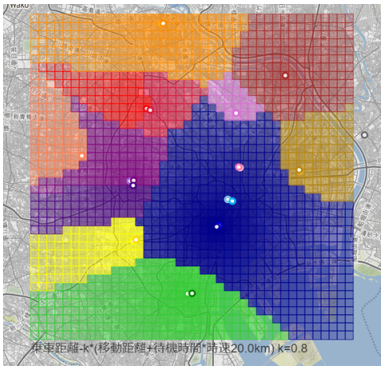

図 6 $k = 0.8$ における推奨エリア

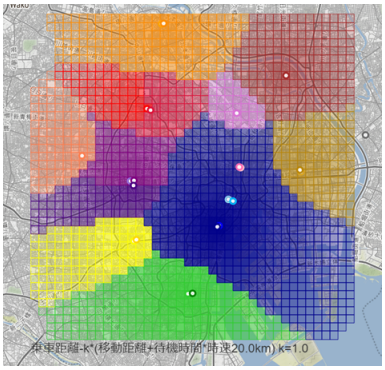

図 7 $k = 1.0$ における推奨エリア

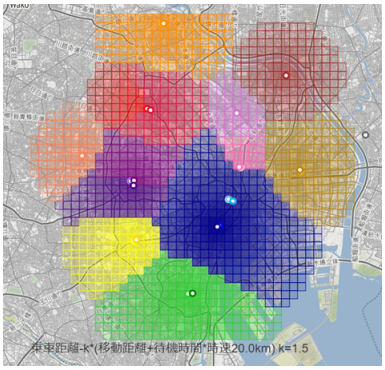

図 8 $k = 1.5$ における推奨エリア

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．対象とするタクシーは 1 つの事業所を対象に 25 時以降に半分以上のプローブデータが空車状態のまま売上 0 円で業務時間終了まで走行を続けていたタクシーとする．この事業所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．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一日の売上平均が約 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 万円程度まで一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，最適化によって一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．

6. おわりに

本研究では都内主要タクシープール 29 か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いての駅エリアの特徴を 30 分毎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作成し，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，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”, `https://www3.nhk.or.jp/news/html/20201022/k10012675891000.html`（2020 年 12 月 1 日閲覧）

[3] H. Yao, F. Wu, J. Ke, X. Tang, Y. Jia, S. Lu, P. Gong, J. Ye, and Z. Li, “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, 1328–1334, 2016

[5] 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, 1-4, 2015

[6] 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, 34(3), C-IA2_1–9, 2019

[7] 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）132(10), 1645–1655, 2012

[8] 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， 65(4), 465–479, 2009

[9] N. Mukai and N. Yoden, “Taxi Demand Forecasting Based on Taxi Probe Data by Neural Network in *Intelligent Interactive Multimedia: Systems and Services* (eds. T. Watanabe, J. Watada, N. Takahashi, R. J. Howlett and L. C. Jain) 589-597, 2012

[10] 総務省，東京都区内走行調査, 2008 `https://www.soumu.go.jp/main_content/000041633.pdf`（2020 年 12 月 1 日閲覧）