

実車率の向上を目的としたタクシー経路の可視化に関する検討

田口 智樹[†] 関根 祐輔[†] 中野 滉斗[‡] 野澤 諒平[‡] 大塚 真吾^{†‡}

[†]{[†][‡]}神奈川県立大学情報学部 〒243-0292 神奈川県厚木市下荻野 1030

E-mail: [†]{s1621043,1621095,1621020,1621056}@cce.kanagawa-it.ac.jp, [‡]otsuka@ic.kanagawa-it.ac.jp

あらまし タクシー業界は基本給に歩合制を加えた給与体系であることが多く、基本給のみでは他の産業と比べて一般的に低収入という傾向がある。そのためタクシー乗務員は給与の向上のために実車率(顧客を乗せている時間の割合)を多くしなくてはならない。一方、タクシー会社は実車率を向上させることにより業務の効率化に繋がると考えられる。実車率を向上させるためには、顧客を乗せているタクシーの経路に合わせてタクシーを配置すれば良いと考える。そこで本研究では、タクシーの実車率を向上させるためにタクシーのプロープデータを利用し、タクシーの経路の可視化を行う。

キーワード タクシー プロープデータ データ解析

1. はじめに

近年、タクシー運転手数が減少している現状にある。図1は国土交通省の調べのもと、縦軸を法人タクシー事業者数、横軸を年度とするグラフであり、平成21年度から減少傾向にあることがわかる。その理由として、タクシー乗務員は一般的に低収入・長時間労働の傾向があり、若年入職者が減少していることが挙げられる[1]。図2, 3は厚生労働省の調べのもと、タクシー乗務員と全産業の月間労働時間と月間給与額を比較したグラフであり、タクシー乗務員は一般的に全産業と比べて労働時間が長く基本給のみでは給与が低い傾向にあることがわかる[2]。

また、顧客を乗せている時間の割合が多くなればその分歩合制であるタクシー乗務員の給与は向上し、タクシー会社は業務の効率化に繋がると予想される[3]。文献[1]によると、2019年度3月の都内タクシーの実車率(総走行距離に占める実車距離の割合)は武三地区(東京都の武蔵野市、三鷹市の2地区)で47.9%、多摩地区で49.0%となっており、毎年50%に満たない実車率となっている。タクシーの台数が減少傾向にあり、若年入職者の減少による人手不足が問題視されているタクシー業界にとっては如何に限られたタクシーの台数で効率良く顧客を運送することができるかが今後の重要なポイントになってくると考える。

そこで本稿では、タクシーの実車率向上のために、実車率の高いタクシーと空車率(顧客を乗せていない時間の割合)の高いタクシーの相違点を発見することを目的とする。

図1. タクシー運転手の推移(法人)



19

図1. タクシー運転手の推移

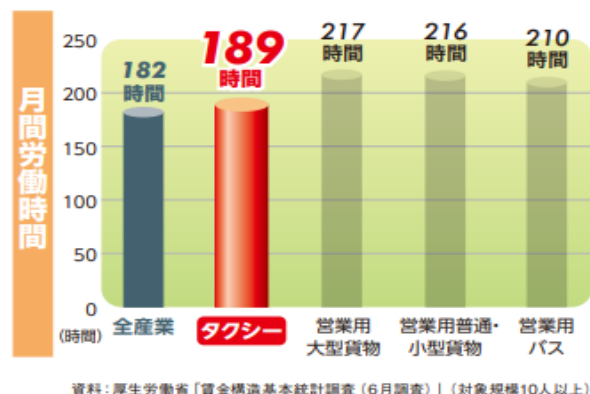


図2. タクシー乗務員の労働時間(平成29年度)

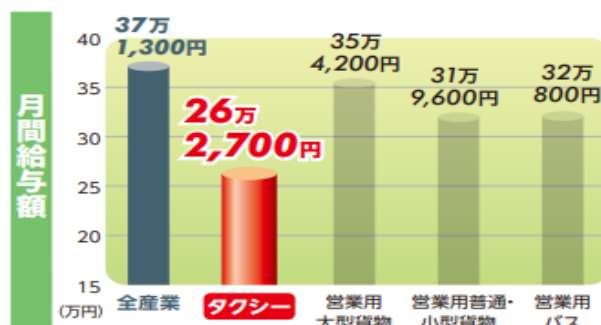


図3. タクシー乗務員の月間給与額(平成29年度)

2. 関連研究

プローブデータとは、センサを積載した車体を走らせて得られた位置情報や車体にかかった加速度などのデータを指す。その中でも今回用いたタクシーのプローブデータは位置情報や加速度の他にも、顧客の有無やタクシーの種類、所属しているタクシー会社などの情報が記載されている。タクシーのプローブデータを用いた基礎的な分析やタクシードライバーの運転行動の分析はすでに行われている[4][5]。

その中でもタクシーのプローブデータを用いるデータ解析において、配車履歴を可視化し、顧客の需要パターンや配車位置の問題点などを明らかにする研究などもすでに行われている[6]。関連研究では、タクシーの利用率に関係する可能性が高い気候データと共に茨城県つくば市近辺のタクシーの配車状況を地図上に可視化し、配車上の問題点の考察や改善案の提案を行っている。しかし、解析に使用したタクシーのデータは1社からのみのものであり、状況分析を行った場所も特定の部分のみのため、場所と乗車率との関係性が掴み切れていないと考える。我々は、複数のタクシー会社のデータを利用していることや人の流れ、タクシーの台数が多い都内タクシーのプローブデータを利用して配車の状況を考察しており、関連研究とは異なる。

3. 研究内容

3.1 今回使用したデータについて

本稿では、データ解析コンペティションで提供されたタクシーのプローブデータをもとにデータ解析を行っている。提供されたデータは2016年4月から2018年3月までのデータであるが、本研究ではまず平均的なタクシーの実車率、空車率について検討するため顧客の需要が比較的落ち着いた6月のデータを用いて実車率、空車率を算出した。

今回の実験では主に VehicleStatus, DriverNumber, およびタクシーの座標データ(緯度, 経度)を使用した。VehicleStatus とはタクシーの現在の状態を表すデータとなっており、顧客を乗せた状態である「実車」、顧客を乗せていない状態である「空車」などがある。また、DriverNumber とは各タクシードライバーに1つずつ割り振られている番号であり、この値からタクシードライバーを一意に定めることが可能となっている。

3.2 実験方法

実車率の高いタクシーと空車率の高いタクシーの相違点を発見するために、我々は実車率の高いタクシーと空車率の高いタクシーの経路を可視化し比較することによって相違点の調査を行った。本来実車率とは総走行距離に占める実車距離の割合を示すが、データ処理の関係上、本研究ではタクシードライバーのデータ数に占める実車回数の割合に着目した。また、空車

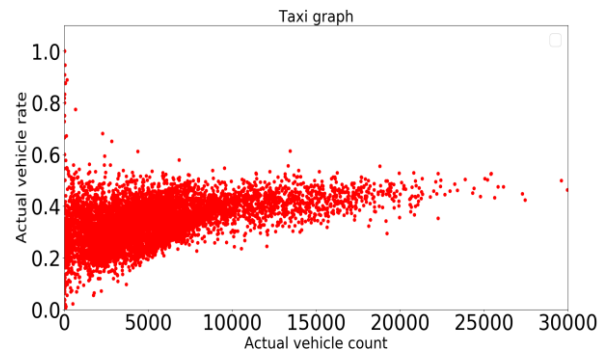


図 4. 実車率と実車回数についてのグラフ

率とはタクシードライバーのデータ数に占める空車回数の割合を示す。DriverNumber ごとで実車率と実車回数との間に関連性がないかを調べるために作成した縦軸を実車率、横軸を実車回数とする散布図を図4に示す。図4では、実車回数が多くなるにつれてデータ数は減少しており、実車率に相関関係が表れていることがわかった。また、データ数が他の DriverNumber と比較して特に低い DriverNumber のデータには相関関係はなく信頼性に欠けていると判断したため、ある一定のデータ数を受け取っている DriverNumber を解析の対象とした。DriverNumber の信頼性のあるデータ数の基準を図るために、縦軸を実車回数とする箱ひげ図を作成した。その結果より、第3四分位数である実車回数4,242以上の DriverNumber(4,396名)を対象に解析を進めていく。

4. 解析の流れ

今回の実験では、提供されたデータを MySQL に格納し、Python を用いて MySQL からのデータの取り出しやマップの作成、座標データのプロットなどのデータ処理の実装を行った。さらに、実車率の高いタクシーと空車率の高いタクシーを算出し、今回用いた2017年6月の実データからマップ上に最も実車率の高かったタクシーと最も空車率の高かったタクシーの座標データのプロットを行った。最も実車率の高いタクシーと最も空車率の高いタクシーの1か月分の座標データを図5, 6, 7に示す。図5は図6の拡大図であり、図6は図7に縮尺を合わせた図となっている。

また、VehicleStatus が実車ならば赤、空車ならば青、その他であれば緑に色を変えることでタクシーの状況を経路とともに可視化した。最も実車率の高かったタクシーでは実車率が61.29%、最も空車率の高かったタクシーでは実車率が24.31%であった。図6, 7を比較すると、実車率の高いタクシーの方が空車率の高いタクシーよりも狭い範囲を行動していることがわかる。この結果から、実車率の高いタクシーと空車率の高いタクシーには行動範囲の広さに相違点があると推測できる。

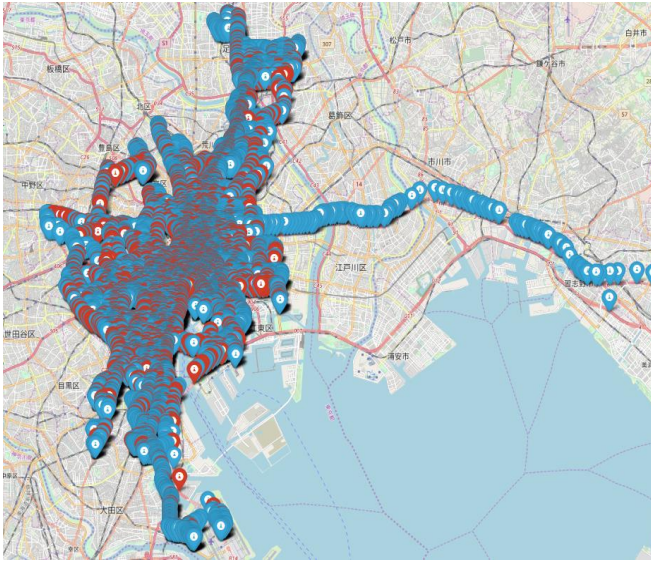


図 5. 最も実車率の高いタクシードライバーの
1 か月分の座標データ

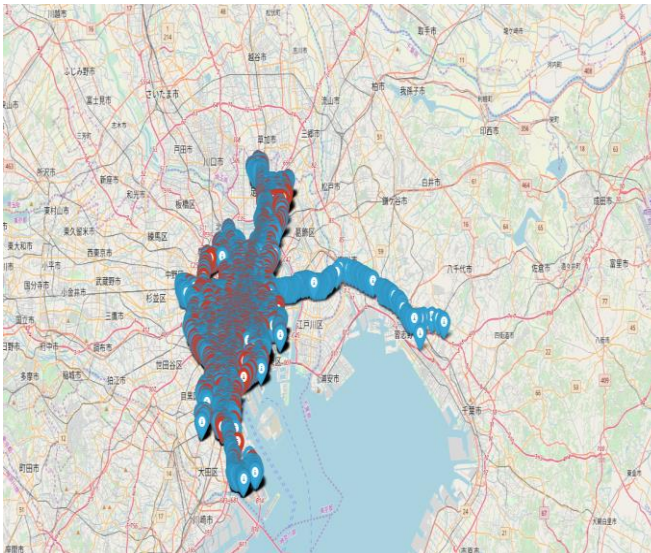


図 6. 最も実車率の高いタクシードライバーの
1 か月分の座標データ

5. 実車率の高いタクシーと空車率の高いタクシーの行動範囲に関する相違点

4 章で述べた実車率の高いタクシーの方が空車率の高いタクシーよりも狭い範囲を活動しているという推測がある程度の数のデータに適しているかを確認するため、実車率の高いタクシーと空車率の高いタクシーの座標データの分散をそれぞれ上位 30 名ずつ算出し横軸を緯度、縦軸を経度とする散布図を作成した。作成した散布図を以下の図 8, 9 に示す。本稿で算出した分散とは座標データの散らばりを表しており、分散の値が小さければ小さいほどより狭い範囲で活動していることを示す。

図 8 からは、ほとんどのタクシーが緯度、経度ともに低い値を取っていることがわかるが、図 9 を見てみ

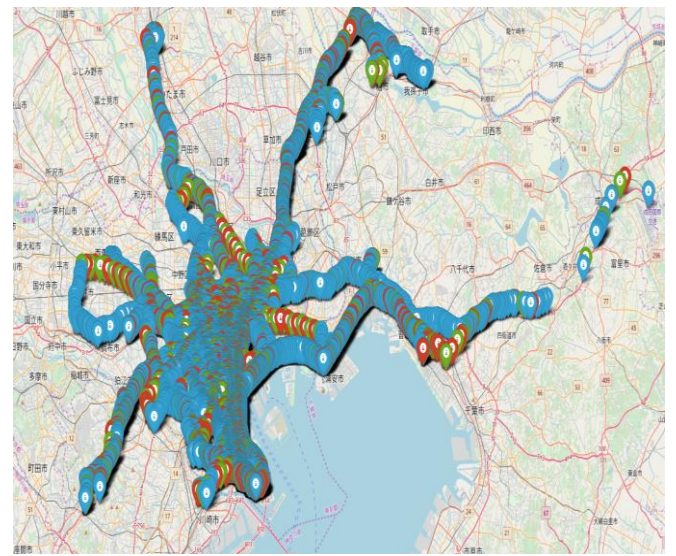


図 7. 最も空車率の高いタクシードライバーの
1 か月分の座標データ

ると、経度は比較的低い値を取っているタクシーが多いものの緯度のばらつきは多いことがわかる。この結果から、実車率の高いタクシーの方が空車率の高いタクシーよりも狭い範囲で効率的に収益を上げていることがわかる。

6. おわりに

我々はタクシーの実車率の向上のために、実車率の高いタクシーと空車率の高いタクシーの経路の可視化を行った。その結果、実車率の高いタクシーの方が空車率の高いタクシーよりも狭い範囲の中で活動しているという相違点を発見することができた。

本稿の課題として、今回データ解析コンペティションで提供されたデータはタクシーによってデータ数やデータの取得速度にばらつきがあるということから、全てのタクシーが同等の条件ではない点や顧客の乗車位置や降車位置が不明瞭な点などから正確な値が出せない現状にあることが挙げられる。そのため、タクシー業界に貢献するためにはより精度の高い座標データの可視化が必要であると考え、さらに、タクシーの利用率に影響が出るであろう天候やタクシーの活動範囲周辺のイベント情報などと組み合わせることによって新たな相違点の発見に繋がると予測される。

謝辞

本件で使用したデータは経営科学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で提供されたものである。ここに記して謹んで感謝の意を表する。

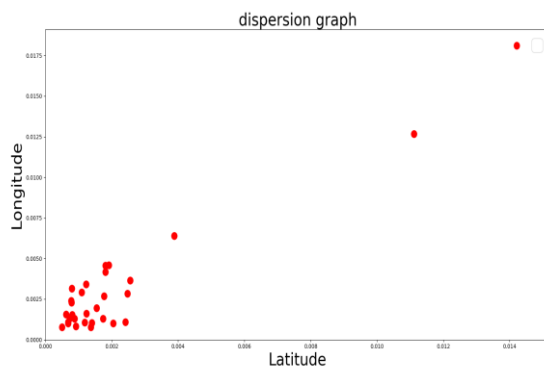


図 8. 実車率の高いタクシーの 1 か月分の
座標データの分散 (30 名)

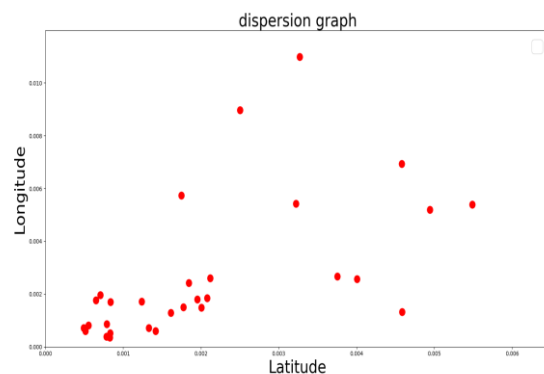


図 9. 空車率の高いタクシーの 1 か月分の
座標データの分散 (30 名)

[参考文献]

- [1]”タクシー事業の現状について”, 国土交通省
< mlit.go.jp/common/001083859.pdf>
- [2] 藤原 廣彦,” 東京のタクシー2019”, 東京ハイヤー・タクシー協会, p.p.15, p.p.37-39,(2019).
- [3]”タクシー運転者の労働時間等の改善基準のポイント”, 厚生労働省労働基準局
<<https://jsite.mhlw.go.jp/tokyo-roudoukyoku/content/contents/000428496.pdf>>
- [4] Toshiyuki Yamamoto , KaiLIU , TakaMORIKAWA,” Preliminary Analysis on Taxi Dispatching Data as Potential Probe Data” ,土木計画学研究,23,4,(2006)
- [5] 金月 寛彰, 服部 宏充,” Analysing and Modeling Individualities of Taxi Drivers’ Driving Behaviors using Probe-Car Data ” ,人口知能学会全国大会論文集,JSAI2015,(2015)
- [6] 宮原直紀, 張勇兵, 繁野麻衣子, 塚本 一也,” Visualization and Analysis of Taxi Dispatch Records” , 情報処理学会報告, 2012-IS-119(2), p.p.1-8,(2012)